

ANHITZ:
Itzulpenean laguntzeko hiztegi-sistema
eleanitza

Xabier Arregi Iparragirre



Udako Euskal Unibertsitatea

© Xabier Arregi Iparragirre
© Udako Euskal Unibertsitatea

ISBN: 84-8438-059-9

Argitaratzailea: UEU. Erribera 14, 1.D, 48005 Bilbo. <http://www.ueu.org>

URL: http://www.inguma.org/tesiak/Arregi_Iparragirre_1995.pdf

Nola aipatu argitalpen hau:

Arregi Iparragirre, Xabier. *ANHITZ: itzulpenean laguntzeko hiztegi-sistema eleanitza* [linean]. [Bilbo: Udako Euskal Unibertsitatea], 2004. http://www.inguma.org/tesiak/Arregi_Iparragirre_1995.pdf [kontsulta: Urtea/hilea/eguna]

Edizio honen ontzailea:



EUSKAL KOMUNITATE ZIENTIFIKOAREN DATU-BASEA

<http://www.inguma.org>

inguma@ueu.org

OHARRA: Galarazita dago dokumentu honen kopia egitea, osoa nahiz zatikakoa, edozein modutara delarik ere, Copyright-jabearen baimenik gabe. Dokumentu honen erabilera bakarra Jabego Intelektualaren legeak 31. eta 32. artikuluetan jasota dakarrena izango da.

LENGOAIA ETA SISTEMA INFORMATIKOEN SAILA



ANHITZ: ITZULPENEAN LAGUNTZEKO HIZTEGI-SISTEMA ELEANITZA

Xabier Arregi Iparragirrek Arantza Díaz de Ilarraza eta Kepa Sarasolaren zuzendaritzapean egindako tesiaren txostena, Euskal Herriko Unibertsitatean Informatikan Doktore titulua eskuratzeko aurkeztua.

Donostia, 1995eko maiatza.



*Joxe Luis Zugastikoari eta Joxepa Postuetxekoari
(koa zuei zor)*

*"Hainbeste bider naiz irakurria eta hainbeste hizkuntzatan
dago*

*idatzia non euskara bera asko den iadanik bihar osoa
ikasiaz izateko*

*Entzun inozoen proposamen zoroa Orain mozten ahal dugu
bizarra"*

(K. Izagirre, Balizko Erroten Erresuma. Susa, 1989).

Eskerrak Mendikoetxeko tandemari, zuekin aldaparik ez,
Begoñako dantzariari, kiribil biribil, txantxak egin eta alai ibil,
Eibarko abeslariari, urretxindorrak *ini-ini*, kukuak *kou-kou*,
Holandako bi arrantzaleei, zuek bai planta, zuek bai sareak,
baina *Arraia Posteris*-ek ihes,
Hiteshi,

eskerrak hiri, kaskarin doktoresari, animoak hirea din izena, mantsotasunak hi hau etsai,
eskerrak hiri, kaskarin doktorgai, animoak hirea dik izaera, hi haiz hi, akuilua

ezkerrak xuxenkideeri, zenbat erakutzi didazuen!

eskerrak itzultzaileoi, Antxon, Arantza, Bittori, Eneko, Itziar, Miriam, Xabier, eta
bereziki Edorta eta Rubeni,

eskerrik asko fakultateko kideei,

eskerrak zuei, Xabier eta Nagore, Sarasola anai arrebak, hain iradokizun interesgarriak
egiteagatik

Lan hau Gipuzkoako Diputazioak lagundutako proiektu baten testuinguruan kokatzen da.

Egileak Gipuzkoako Kuxaren diru-laguntza jaso du.

AURKIBIDEA

I.	ANHITZ: motibazioak, aurkezpena.....	13
I.1	Giza itzulpenean laguntzeko sistemez.....	14
I.1.1	Hiztegien erabilpena.....	18
I.2	ANHITZ sistemaren aurkezpen orokorra.....	19
I.3	Diseinu-metodologia: KADS.....	21
I.4	Txostenaren nondik norakoak.....	24
II.	Hiztegien erabilpenaren modelizazioa itzulpenean.....	27
II.1	Hiztegien erabilpena aztergai.....	27
II.1.1	Aurrekariak.....	28
II.1.2	Hiztegi elektronikoen erabilera.....	32
II.2	Gure azterketaren deskribapena.....	33
II.3	Modelizaziorantz: kontsiderazio orokorrak.....	39
II.4	Itzulpen-prozesuaren modelizazioa.....	46
II.4.1	Modelizazio-metodologiaz.....	46
II.4.2	Modelizazio-lengoaiaz.....	49
II.4.3	ANHITZ ereduaren deskribapena.....	54
II.4.4	Ataza primitiboak.....	68
III.	Hiztegi-ezagutzaren errepresentazioa.....	89
III.1	Domeinu-ezagutzaren errepresentazioa KADS metodologian.....	89
III.2	Hiztegi-ezagutzaren errepresentazio-ereduak.....	91
III.2.1	Lexiko eleanitzen egituraketa.....	93
III.2.2	Itzulpenerako lexikoak.....	98
III.3	ANHITZ Hiztegi-Ezagutza Baserako aurremugapen eta printzipioak.....	102
III.4	ANHITZ HEBaren deskribapena.....	108
III.4.1.	S-STRUCTURES Ezagutza-Basea.....	109
III.4.2.	Ingurune elebakarrak.....	110
III.4.3.	Ingurune elebiduna.....	117
III.4.4.	Itzulpenera zuzendutako ezagutza.....	120
III.4.5.	Inferentzi geruza.....	122
III.4.6.	ANHITZ HEBaren egungo prototipoaren egoera.....	124
III.5	Gogoeta eta oharpen zenbait.....	126

IV. ANHITZ Sistemaren Geruza Estrategikoa eta Arkitektura	131
IV.1 Atazen Kudeaketarako Arkitekturak	132
IV.1.1 Estrategi geruza KADSen. Arkitektura-eredu bat.	133
IV.2 ANHITZ Sistemaren Arkitektura	137
IV.2.1 Datu-moduluak.	139
IV.2.2 Modulu Funtzionalak	147
IV.3 Inplementazio-egoera eta landu beharrak.....	155
V. Ikuspegi funtzionala.	159
V.1 ANHITZ sistemaren portaera.	159
V.1.1 Laguntza-estrategiak.....	162
V.2 Sistemaren barne-funtzionamendua.	164
V.3 Interfazearen xehetasunak.	168
V.4 Erabilera-adibideak.....	173
VI. Ondorioak eta etorkizuneko ikerlerroak.	179
VI.1 Ondorioak.	179
VI.2 Aurrera begirako lan-lerroak.	181
VI.2.1 Hiztegien erabilera-azterketak.	181
VI.2.2 Hiztegi-ezagutza osatu, egokitu eta aberasteko aukerak.	181
VI.2.3 Laguntza emateko beste erak.	183
VI.2.4 Baliozkotasuna eta ebaluazioa.	183
VI.2.5 Sistemaren berrerabilgarritasuna.	184
BIBLIOGRAFIA.....	186

I. ANHITZ: motibazioak, aurkezpena.

I. ANHITZ: motibazioak, aurkezpena.

Askoren interesa eta gogoia ernatu du itzulpen automatikoaren arloak urteetan. Ez da harritzekoa, beraz, lengoia naturalaren prozesamenduan (LNP) itzulpen automatikoarena izatea gairik harrotuena. Argitu behar da, ordea, erabat makinaz egindako itzulpenari zail iritzita, gizakiaren parte hartzea aintzat edukitzera jo dela arlo honen bilakaeran. Dela gizakiak lagundutako ordenadore-itzulpena, dela ordenadorez lagundutako giza itzulpena, elkarrekintzak aparteko garrantzia du gaur egun.

Gauzak horrela, ordenadorea eta gizakia elkarlanean jarri nahiak ekimen ugari eragin du. Ekimen horietan, Isabelle-k dioenez (92:77), apal antzean landu nahi izan da itzulpenaren automatizatzea:

Rather than start from inadequate systems and ask translators to compensate for their flaws, one starts from human translation and looks for ways, however modest, to make machines helpful.

Hurbilpen horren ildoak proposatuak zituen Martin Kay-k urte batzuk lehenago (80:11):

I want to advocate a view of the problem in which machines are gradually, almost imperceptibly, allowed to take over certain functions in the overall translation process. First they will take over functions not essentially related to translation. Then, little by little, they will approach translation itself. The keynote will be modesty. At each stage, we will do only what we know we can do reliably. Little steps for little feet!

Idatzi horietan, ematen den ikuspegi praktikoaz gain, nabarmentzekoa da orobat giza itzultzailearen arieran jartzen den arreta. Hor bada, guk uste, ohikoa izan ez den tratamendu baten aldeko jarrera, 80an iradokia bai, baina berritzailetzat har daitekeena gaur egun ere. Bestela, nekez uler daiteke Steiner-ek dioena (93:282):

Translators have had surprisingly little contact with the world of MT research so far.

Itzulpen automatikoaren eta giza itzulpenaren arteko erlazioa hizpide hartuta, C. Hauenschield-ek hiru baieztapenetan oinarritzen du bere ustea (93):

- Gaur arte itzulpen automatikoak ez du arreta handirik jarri giza itzulpenean, nahiz eta helburu duen gizakiaren ordezkio izatea.
- Erlazio-ezak eragin kaltegarria izan du itzulpen automatikoaren teorian eta praktikan.

- Itzulpen automatikoko ikerkuntza eta itzulpen-teoria elkarren aberasgarri izan daitezke eta, are, arlo horien konbinazioa oinarri gisa har daiteke itzulpen automatikoko aplikazioei soluzioak emateko.

Giza itzultzaileei behatu eta haiengandik ikasi beharra azpimarratzen dute Johnson eta Whitelock-ek. Itzultzaileen portaera simulatzeaz ere mintzo zaizkigu autore hauek (87:136):

*Since an ideal MT system will probably be expected by consumers of translations to exhibit the functional input-output behavior of an ideal human translator, it is not unreasonable to look to translators as a primary source of information about the problems of MT.*¹

Itzulpen automatikorako hurbilpen horrek, praktikotasuna eta giza itzultzaileen behaketa azpimarratzen duenak, badu gure iritzirako zimendarri aski sendorik aztergai erakargarria eta oparoa izateko.

I.1 Giza itzulpenean laguntzeko sistemez.

Ordenadorez lagundutako giza itzulpenean gizakia da itzultzailea, hark eta ez bestek hartzen du erantzukizuna; ordenadorea, berriz, laguntza eskaintzera mugatzen da². Molde horretan, zehazki gizakiaren eta ordenadorearen arteko elkarrekintzan, hiru estadio bereizi ohi dira:

- Aurre-edizioa: sorburu-testuaren prestaketa lantzen da fase honetan, besteak beste, testua erabilerraztu, aurrikus daitezkeen arazoak markatu eta horiek ebazpidean jartzen dira. Helburu horiek iristeko iharduera automatiko ohikoenak hauexek dira:
 - Terminoen edo hitz berezituen identifikazioa: ataza hau glosategien osaketarako aurrepausotzat har daiteke.
 - Testua puskatan antolatzea ("bracketing"): bereziki testuko egiturak — parrafoak, esaldiak, perpausak, sintagmak — banatzea edo nabarmenaraztea.
 - "Aurre-itzulpena": itzulpen-aukera bakarra duten sorburu-hitzak automatikoki itzultzean datza. Prozesu hau baliagarria gertatzen da terminologia itzultzean, baliokide sendo eta anbiguotasunik gabeak nahi izaten baitira horrelakoetan. Azken batean, aurre-itzulpen automatikoa giza itzultzaileek egin ohi dutenaren mekanizazioa da.

¹ Horrek ez du esan nahi, autoreen ustetan, giza itzultzaileen portaeraren aspektu guzti-guztiak modelizatu behar direnik itzulpen automatikoko sistema ideala diseinatzeko.

² Eredu horri kontrajarriz, gizakiak lagundutako ordenadore-itzulpenean ordenadoreak hartzen du itzulpenaren ardura eta gizakiari ordenadoreak beharra duenean laguntzea dagokio.

- Post-edizioa: xede-testuaren bertsio bat lortutakoan abiatzen da post-edizioa. Fase honetako eginkizunik aipagarrienek egiaztapenarekin eta zuzenketarekin dute zerikusia. Ortografia, sintaxia eta estiloa jorratzen dira zuzenketa-saioetan.
- Edizioa: itzultzailea bete-betean itzultzen ari dela gertatzen da edizioko elkarrekintza. Etapa honetan testua prozesatzeko aukerak eta informazio-kontsultak dira gehien ustiatzen direnak.

Azken urteotan langintza horietarako tresna informatiko berezituak garatu dira, eta tresna horien integrazioak *itzultzailearen lanpostuak* ("translator's workbench") erakarri ditu. Hona hemen ondoan lanpostu horietan biltzen diren tresnetako batzuk:

- *Testu-prozesadoreak*: bistan denez, ez dago esaterik testu-prozesadoreak itzultzailearentzako tresna berezitua direnik, askoz erabilera zabalagoa baitute. Badira, ordea, itzultzaileari begira diseinatutako testu-prozesadore linguistikoki adimentsuak, eta hauek bai, hauek badituzte laguntza berezituak: gai dira, konparazio batera, zenbait anbigutasun antzemateko, egitura alternatiboen sorkuntza errazteko edo zenbait komunztadura automatikoki kontrolatzeko.
- *Egiaztatzaileak eta zuzentzaileak*: ezagun eta hedatuenak egiaztatzaile-zuzentzaile ortografikoak dira³. Horiek baino sofistikatuago, eta ez hain ezagun, dira gramatika-eta estilo-egiaztatzaileak. Errakuntza gramatikalak —komunztadura-ezak, deklinatze-akatsak, aditz-joko desegokiak, etab.— dira gramatika-egiaztatzaileen gaia. Estilo-egiaztatzaileak, berriz, joskera edo klitxeen erabilera bezalako estilo kontuez arduratzen dira.
- *On-line kontsultarako baliabideak*: esan gabe doa hiztegien kontsultak baduela garrantzia itzulpengintzan. Hiztegi elebakarrak, elebidunak, terminologi hiztegiak, norbere hiztegiak, tesauruak, mota askotako hiztegiak nahi izaten ditu itzultzaileak. Baina badira, hiztegiak ezezik, beste iturburu batzuk ere: entziklopediak, korpusak, aurretik egindako itzulpenak, testuari atxikitutako glosategiak eta datu-base terminologikoak⁴, besteak beste. Horien guztien kontsulta azkar, eroso eta aberatserako aukera ematea da on-line tresna hauen helburua.
- *Testu-lerrokatzaileak*: sorburu- eta xede-testuen arteko korrespondentziak edo parekatzeak ezartzeari deitzen zaio testu-lerrokatzea. Maila desberdineko

³ Euskarak badu berea, XUXEN izenekoa (Agirre *et al.*, 92).

⁴ Datu-base terminologikoen kontsultak badu berezko garrantzia. (Hutchins & Somers, 92:150)n emandako estimazio baten arabera, itzultzaile teknikoek %60an kontsulta terminologikoan omen datza.

korrespondentziak ezar litezke: hitzen mailatik atalen mailaraino, sintagmak, esaldiak eta paragrafoak lerrokatu ohi dira.

Corpus elebidunen tratamenduan (Gale & Church, 93) eta itzulpen automatikoan (Isabelle, 92) hedatu da, batez ere, testuak lerrokatzeko interesa. P. Isabelle garbi mintzo da lerrokatzeak itzultzaileen lanpostuetan izan dezakeen eraginaz (76):

The functions taken over by existing translators' workstations are rather peripheral with respect to the core aspects of the translation task. However, recent developments show that it is possible to automatically produce explicit (partial) representations of the translation correspondences that link pairs of source and target texts.

Eta dio, segidan, errepresentazio elebidun horiek —*bi-testu* deitzen die— oinarri gisa erabil daitezkeela beste zenbait tresna diseinatzeko: adibidez, itzulpen-memoriak, komunztadura elebidunak jaso, edo zuzentzaileak, korrespondentzi errakuntzen trataera aberastuz.

Idea horiek eraginda itzultzaileen lanpostu desberdinak garatu dira azken hamarkadan. Aitzindarietan bat BYU-TAS proiektua da (Melby, 87a; 87b). BYU-TAS lanpostua —lehenengo bertsioa 81ean programatu zen— hiru mailatan dago diseinaturik. Lehen mailan testu-prozesadorea, hiztegi elebidunen atzipenerako tresna eta on-line komunikazioetarako aukerak integratzen dira. Bigarrenean, komunztadura gramatikalak eta interfazearen diseinua lantzen dira, itzultzaileari errazago gerta dakion lehen mailako aukerez baliatzea. Hirugarrenean, azkenik, testuaren segmentazioa eta segmentuen itzulpen automatikoa gauzatzen da, segmentu-bikoteak post-ediziorako prest utziz.

(Macklovitch, 89)n CWARD lanpostua deskribatzen da. Tresna horretan itzultzaileari eskain dakizkiokeen aplikazio batzuk integratzen dira, gehientsuak nahikoa oinarritzkoak: testu-prozesadorea, tesaurua, egiaztatzaile ortografikoa, informazio gramatikala, datu-bankuak atzitzeko euskarriak, etab.

TransActive™ sistemaren berri ematen zaigu (Hutchins and Somers, 92)n. Elkarrekintzazko itzulpen automatikoko pakete honek baditu, besteak beste, testu-prozesadore eleanitzak eta lineako glosategiak. Paketaren erabilerak sortarazten dituen gorabeherak aipatzen dituzte autoreek eta, guk uste, mota horretako edozein sistemaren garapenean kontuan hartzekoak dira⁵. Ondoko pasarte hau bederen gogoan edukitzekoa da (153-154):

⁵ Argitu beharrekoa da TransActive™ paketea gizakiak lagundutako itzulpen automatikora lerratzen dela. Hala ere, diferentziak diferentzia, ondorioak baliagarriak litezke ordenadorez lagundutako giza itzulpenarekin ere.

The difficulty with this kind of interactive system is that there are often so many interactions for a single sentence that it might have been quicker for the user to translate from scratch. Furthermore, it is frustrating and irritating to have constant repetitions of the same interactions with the same lexical items.

Pisan garatutako itzultzailearen lanpostuak (Picchi *et al.*, 92) hiztegi datu-baseak eta artxibategi elebidunak ditu oinarrizko osagai. Horien gainean, kontsultarako tresna azkar, malgu, erabilterraz eta garraiaagarriak eraiki nahi izan dira. Kontsultatzeko ezezik, itzultzaileak aukera du aldaketak egin eta bere datuak barneratzeko artxibategi horietan.

TWS ingurunean —"Translator's Workstation", (Nirenburg *et al.*, 92)— itzultzailearen zereginetarako interfaze berezitu bat eratu da. Interfazearen bitartez eskura jartzen zaizkio itzultzaileari testu-editore propioa, hiztegiak eta glosategiak kontsultatzeko edo eguneratzeko baliabideak eta ONTOS sistema. TWS editoreak badu hainbat aukera testuak lerrokatzeko edo maiztasun-azterketak egiteko. ONTOS sistemak ontologia bat kudeatzen du.

Ingelesetik hizkuntza eslaviarretarako itzulpenetan laguntzeko SPSS sistema garatu da (Strossa, 94). SPSSk hiztegi datu-basea eta testu-editorea integratzen ditu. Lan honetan, ordenadorez lagundutako itzulpenerako sistemek, bereziki ezaugarri hain desberdineko hizkuntzak lantzen dituztenek⁶, zein gutxienezko ezagutza linguistikoa behar duten eta zein fenomeno gramatikal tratatu beharko luketen azaltzen da. SPSSn arreta handia jartzen da lematizazioan, hitz anitzeko termino edo adierazpenetan eta hutsune lexikaletan⁷.

GETA taldeko ARIANE-78 sistemaren inguruan, itzultzaileen beharrei erantzuteko asmotan, ATLAS —hiztegi kodetuak maneiatzeko laguntza—, THAM —testu-editoreari lotutako funtzio-paketea— eta VISULEX —datu-base linguistikoaren erabilpenerako interfazea— erremintak egin ziren (Bachut & Verastegui, 84). Geroztik, *itzulpen automatiko pertsonalaren* ideia mamitu da LIDIA proiektuaren markoan (Boitet, 90), eta ideia horri, heldutakoan, "Elkarriketan Oinarritutako Itzulpen Automatiko" (DBMT) deitu zaio (Blanchon, 94; Boitet & Blanchon, 94). LIDIA sisteman itzultzailearekiko elkarrekintzak aspektu desberdinak biltzen ditu, besteak beste, testuaren tipologia xehetzea, zuzenketa ortografikoa, terminologi estandarizazioa, estandarizazio estilistikoa, formula finkatuen detekzio eta tratamendua, desanbiguatzea, sorburutik xederako transferentzia, etab.

TWB ("The Translator's Workbench") Esprit proiektuan (Kugler *et al.*, 95) giza itzulpenari dagozkion prozesu kognitiboen azterketa burutu da, itzultzaileek dituzten

⁶ Hizkuntza eslaviarrak oso flexiboak dira eta hitz-ordena nahikoa askea dute.

⁷ Hala deitu zaie itzulpen zuzenekorik ez duten sorburu-hitzei.

beharrak ezagutu eta horien neurriko software-tresnak eraikitze⁸. TWBn aztertu, integratu eta ebaluatu diren erremintak hiru multzotan sailkatzen dira:

- Itzulpen aurreprozesukoak: eskuarki, hiztegi, datu-banku, corpus eta, oro har, era guztietako dokumentu eta hizkuntz baliabideak atzitzearekin zerikusia duten gorabeherak lantzen dituzte multzo honetako tresnek. Formato estandarren erabilerari aparteko garrantzia ematen zaio.
- Itzulpen prozesukoak: itzulpen automatikoko sistemen erabilpena aipatzen da, zehazki, METAL sistemarena. Horrez gain, *itzulpen-memoriaren* ideia garatzen da, testu eta hitzen gordailu soila baino askoz gehiago dena. TWB itzulpen-memoria sistema oso bat da, datu estatistikoak gorde eta aplikatzeko gai dena, eta sorburu- eta xede-hizkuntzen eredu estokastikoak sortzen dituena.
- Itzulpen post-prozesukoak: sail honetakoak dira egiaztatzaile-zuzentzaileak. Testuingurua aintzat hartzen ez duten zuzentzaile ortografikoez gain, gramatika-, sintaxi- eta estilo-zuzentzaileak ere biltzen dira.

I.1.1 Hiztegien erabilpena.

Orainokoan nabarmendu denez, hiztegien kontsulta eta erabilpena alderdi inportantea da itzultzaileari laguntzeko helburua duten software-erreminten artean. Sistemok, halabeharrez, itzulpenaren eta lexikografiaren arloak elkarrengana biltzen dituzte.

Konputagailuen ahalbideek eraginda, lexikografiaren arloan bide berriak urratzeko saioak egin dira, hiztegien edukia eta erabilera hobetzeko gogoz eta afanez. "Etorkizuneko hiztegia" aipatzen da, aro berri baten atarian gaudela nabarmentzeko. Etorkizunari begira egon gabe, egun, paperean ezezik euskarri elektronikoetan ere badaude hiztegiak: konparazio baterako, CD-ROM hiztegiak produktu komertzial arruntak dira. Baina, kontu, lexikografia konputazionalak euskarri-aldaketa soiletik areago joan behar du, esan nahi baita, euskarri horietara egokitu eta baliabide berriak, orain arte ezagutu ez diren bezalakoak, sortu behar dituela. Argigarriak dira Fillmore eta Atkins-en hitzok (94:376):

At best —and some fall far short of this— dictionaries on CD-ROM represent an enhanced version of a work which already exists in print form. (It should be noted here that the labour and costs involved in producing a good CD-ROM dictionary in this way are not trivial.)

However, if dictionary-users are to benefit from the new medium (and if successful natural language processing is to become a reality) the traditional lexicographical

⁸ Erabiltzaileek ere parte hartu dute softwarearen sortze-prozesuan.

approach to word meaning must, we believe, change in response to the demands of and opportunities offered by the new technology.

Jakina, produktu berri horietara heltzeko beharrezkoa da, lexikografia klasikoarekin batera, informazioaren errepresentazio- eta berreskuratze-teknikak landu eta integratzea. Etorkizuneko hiztegiara begira dauden hurbilpen eta ikuspegi desberdinen artean, guri bereziki erakargarria iruditzen zaigu hiztegia tresna *aktibo* gisa planteatzen dutenena. W. Martin-ek (92:200)en, "... *they [dictionaries] should function as semantic problem solvers on word level*" idatzi zuen. Ideia horren oinarria, (90:394)n formulatu zuten Martin-ek eta B.P.F. Al-ek:

the use of a dictionary can be seen as a typical problem-solving activity, and user-orientation should involve both (static) knowledge and dynamic features (strategies, aims, needs) of the intended user.

Hots, hiztegiak, giza erabiltzaileari zuzendutako tresna izanda ere, "ezaugarri dinamikoak" hartuko lituzke ordenadorearen bitartez. Ez luke, hartara, soluziora heltzeko ardura guztia erabiltzailearen eskuetan utziko.

Hiztegi dinamiko edo aktiboaren kontzepzio hori aplikagarria da, guk uste, ordenadorez lagundutako itzulpenaren arloan, batez ere hitzen tratamenduan. Sorburu-hitzen ulerkuntzan bezala, hitz horien baliokideak bilatzean nahiz xede-hitzen sorkuntzan erabiltzen ahal da hiztegi aktiboa. Hiztegi mota horretan, esanahiez gain, informazio gramatikalak eta adibideei, azpikategorizazioei nahiz kokakidetzei buruzko xehetasunek lekua behar dute izan. Hitzak edo, hobeto, adierak *testuinguruan* noiz eta nola txertatu, hori ere *erakutsi* behar dute hiztegiek. Eta horretarako ez da basta testuingurutik independentea izan ohi den definizioa ematearekin.

I.2 ANHITZ sistemaren aurkezpen orokorra.

Aurreko atalean esandakoek bi aspektu azaleratu dituzte: batetik, giza itzulpenari ordenadorez laguntzeko dauden joerak, eta bestetik, laguntza horietan hiztegiak —konputazionalak— joka dezakeen rol aktiboa. Bi alderdi horiek jaso nahi ditu ANHITZek, itzulpenean laguntzeko hiztegi-sistema eleanitzak.

ANHITZen sorrera ez da hutsetik abiatu, ordea. Testu-ulerkuntzan laguntzeko xedea zuen HIZTSUA hiztegi-sistema elebakarra (Artola, 93) hartu behar da ANHITZen aurrekaritzat. HIZTSUAk portaera adimentsua erakusten du erabiltzaileak frantseseko hiztegi elebakarra kontsultatzen duenean. Dedukzio-gaitasuna eta funtzionalitatearen

aberastasuna dira HIZTSUAren kanpora begirako ezaugarriak nabariak (Agirre *et al.*, 94e).

HIZTSUA ez bezala, ANHITZ eleanitza da —egungo prototipoan elebiduna⁹— eta, ulerkuntzan ezezik, itzulpeneko beste zereginetan ere, hala baliokide-bilaketan nola sorkuntzan, erabiltzeko dago pentsatua, beti ere maila lexikalean.

ANHITZek giza itzulpenen hiztegi-euskarria izateko bokazioa du. Hori bakarrik eta hori guztia. Hori bakarrik diogu, giza itzulpenean hiztegiak kontsultatzea ataza lagungarria baino ez baita. Baina, hori guztia esan dugu orobat, hiztegiek, lexiko-sistemek, informazio asko gorde eta garrantzi handia eduki baitezakete, gure ustetan, tresna lagungarri nagusia izaterainoko garrantzia.

Ohargarria da, bestalde, ANHITZ ez dagoela itzultzaile profesionalei zuzendua. Horiek baino gehiago, arituak izan arren adituak ez diren itzultzaileak, edo are halamoduzkoak, dirateke ANHITZen erabiltzaile, azken hauek hiztegia maizago behar baitute.

Baina bada kontu bat hemen argitu beharrekoa: zergatik ez ditugu ahaleginak itzulpen lexikal automatikorako tresna bat sortzera zuzendu? Azken finean, ez al lioke bada itzulpen lexikal zuzenekoak laguntza oparoagoa emango itzultzaileari? Arrazoi batek bultzatu gaitu bide hori baztertera: itzulpen automatikoan itzulpen-urratsak transparenteak dira, ikustezinak, hau da, itzultzailearen aurrean arazoak ezkutatu egiten dira, eta hori dela eta, sasi-ziurtasuna —zalantza baino askoz okerragoa— sortarazten da¹⁰. Aitzitik, ANHITZen, lexikoaren tratamendua duen zailtasun eta dimentsio osoan landu nahi da, arazoak ezkutatu beharrean azaleratuz eta itzultzaileari topatutako arazoen berri emanez uneoro. Dinamika horretan, hiztegiak noiz, nola eta zertarako kontsultatzen diren azaltzen da, urratsez urrats azaldu ere, itzultzailea pentsatzera eta erabakiak hartzera behartuz. Hartara, itzultzaileak bere gain hartzen du beti erabakitzea. Hiztegia lanabes aktiboa da; itzultzailea, berriz, subjektu aktibo nagusia. Hori da bide bakarra itzulpen lexikalak duen konplexutasunaz jabetu eta zailtasun horiei kontzienteki aurre egiteko.

Helburu nagusiak azalduta, tarte bat egingo diegu ondoan ANHITZen nolakoei:

- Itzultzailearen lanpostuan integratzeko beste tresna bat da, ez bakarra haatik. Giza itzultzaileari hitzak edo unitate lexikalak itzultzean laguntzen dio, hiztegiatako informazioa eskura jarritz.

⁹ Euskara eta frantsesa landu dira gaurdaino.

¹⁰ Ñabardura, aitzaki eta salbuespen asko jar dakizkioke baieztapen horri, baina automatizatzeak badu gutxi edo gehiago hor esandakotik.

- ANHITZ ezagutzen oinarritutako sistema da.

Badu hiztegi-ezagutza, hiztegi batetik baino gehiagotik erauzia. Hiztegianitza izateaz gain eleanitza ere bada. Bestetik, itzulpen-prozesuaz eta hiztegi-erabilpenaz dakiena da. Bi ezagutza-osagai horietan oinarritzen da bere ariera.

- ANHITZ: erabakitze euskarri-sistema ("decision support system", DSS).

Itzultzaileak sorburu-hitz baten baliokidea sortu behar du xede-hizkuntzan. Hori du erabakitze, eta horretan du ANHITZ euskarri. Ikuspegi honetatik DSSen ezaugarrietara hurbiltzen da (Radermacher *et al.*, 94).

- ANHITZ: laguntza-sistema ("help system").

Sistema hauek tresna informatikoen erabileran eman ohi dute laguntza (Houghton, 84; Fischer *et al.*, 84; 85; Kemke, 86; Wilensky *et al.*, 88; Breuker, 90; Krause *et al.*, 93). ANHITZek bere burua —hiztegi-sistema— erabiltzeko laguntza ematen du, ez baitugu ahantzi behar hiztegi-erabilpenaz ere badakiena. Sistema berean hiztegi ezagutza-baseak eta horiek erabiltzeko laguntzak daude integratuta.

- Funtzionamenduari dagokionean¹¹:
 - Arina. Hiztegi-kontsultak erraza, eroso eta azkarra behar du izan.
 - Elkarrenergilea. Itzultzailearekin "joan-etorriko" informazio-fluxua duena.
 - Batzuetan pasiboa —itzultzailearen eskariei erantzuten die—, hurrena berriz, aktiboa —eskatu gabe ere laguntzen badaki—.
 - Adeitsua. Itzultzaileari beti amore eman, eta gogaikarria ez dena.
 - Moldagarria. Egoerak kontuan izaten dituena itzultzaileari laguntza ematean.
 - Adimentsua. Dedukziorako gai dena eta itzultzaileari aurrea hartzeko joera duena.

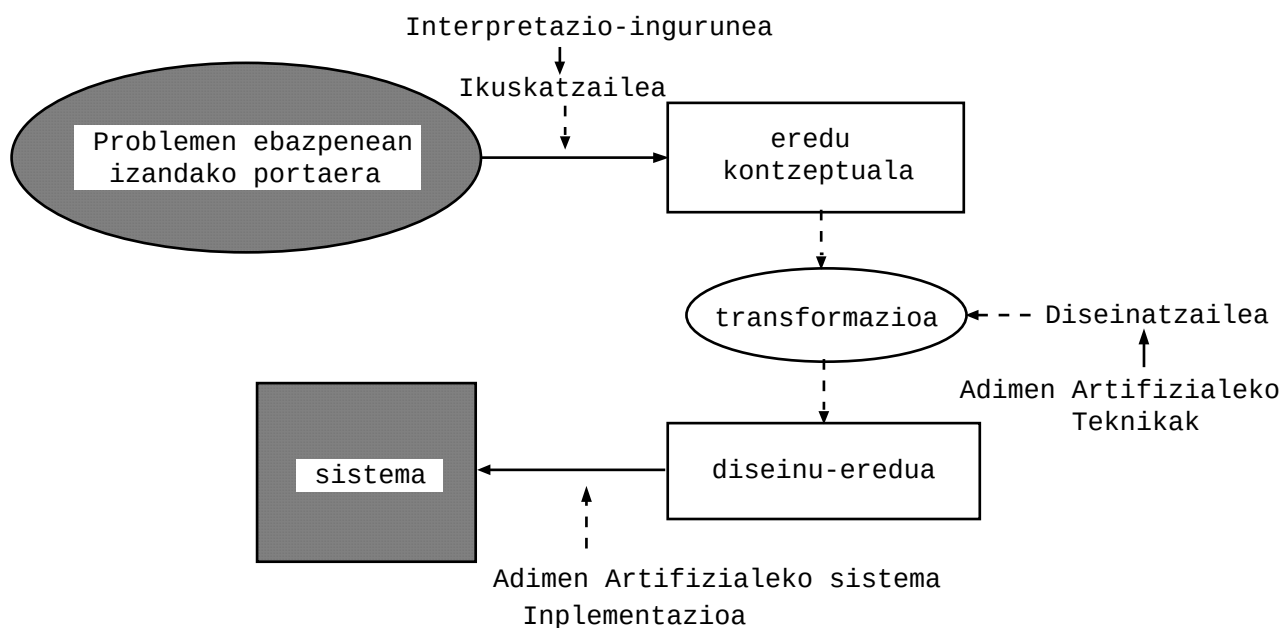
I.3 Diseinu-metodologia: KADS.

ANHITZ sistema nolakoa nahi dugun gogoan hartuta, itzulpenerako hiztegi-sistema sortu eta eraikitze *atazei zuzendutako hurbilpena* baliatu dugu. Hurbilpen horretan, sistemak atazei buruzko ezagutzez eta horiek maneiatzeko gaitasunaz hornitzen dira. Ataza-ezagutza gune gisakoa da: horren pean leudeke objektu oinarritzkoak, horren gainean atazak konbinatzeko baliabideak.

¹¹ Ikuspegi funtzionalari heltzen diogunean, § Ven, xeheago aztertuko ditugu ezaugarriok.

KADS ("Knowledge Acquisition and Documentation Structuring") atazei zuzendutako diseinuak garatzeko metodologia egokia suertatu da (Schreiber *et al.*, 93). Horren frogagarria da 93. urterako bazirela berrogeita hamarren bat Ezagutzan Oinarritutako Sistema —edo proiektu— KADS metodologia erabiltzen zutenak. Egia esan, ez dago esaterik denek KADS "aratza" erabiltzen zutenik, baina bai metodologiaren lerro eta ideia nagusiei atxiki zitzazkiela. Konparazio batera, Esprit VITAL proiektua (Rouge *et al.*, 93), Ezagutzan Oinarritutako Sistemen garapenerako ingurunea eta metodologia sortzeko xedea duen proiektua, KADSen dago oinarritua, baina KADSen gaineko aldaketak eta egokitzapenak proposatzen ditu halaber. Oro har, metodologia honek Europan izan duen zabalkundea eta arrakasta ikusita, esan liteke *de facto* estandar bihurtzen ari dela. KADS metodologian ezagutza eskuratzea, funtsean, modelizazio-saioa da. Zentzu horretan, Ezagutzan Oinarritutako Sistema, ezagutzaren gordailu soil baino areago, portaera desiratu bat erakusten duen eredu operazionala da. I.1 irudiak erakusten du gaingiroki nola erakusten eta esleitzen zaion sistemari adituaren portaera. Eredu kontzeptuala eraiki nahi den sistemaren portaera-espezifikazioa da, inplementaziotik independentea. Diseinu-ereduan, berriz, portaera hori aurrera eramateko errepresentazio- eta konputazio-teknikak zehazten dira. Bereizketa horren komenientzia markatuz, horrela diote (Schreiber *et al.*, 93:9)n:

The main advantage lies in the fact that the knowledge engineer is not biased during conceptual modelling by the restrictions of a computational framework.



I.1 irudia.- Ezagutzaren eskuratzeko prozesua KADS metodologian, (Akkermans *et al.*, 93a:174)tik jaso eta euskaratua.

Ereduranzko bidea, hots, modelizazioa¹² burututakoan KADS jakinduri eredua ("model of expertise") lau maila edo geruza desberdinek osatzen dute:

- *Domeinu-geruza*: kategoria honetakoa da domeinuari buruzko ezagutza estatikoa. Domeinuko kontzeptuak eta horien atributuak, egitateak, egiturak eta erlazioak errepresentatzen dira maila honetan. Domeinuko entitateen errepresentazioa —ahal den neutroena— eta ezagutza horren ustiapena bereiztea garrantzizkoa da, hartara lortzen baita domeinua berrerabilgarria izatea, hau da, testuinguru edo problema desberdinen ebazpenean usatzea.

Gurean, hiztegi-ezagutzaz hornitu dugu domeinu-geruza. Hiztegietako entitate eta erlazio desberdinak ezagutza base batean antolatu ditugu, teoria lexikal-semantikoek erakutsitakoaren bidetik.

- *Inferentzi geruza*: maila honetan inferentziak biltzen dira. Inferentzia, KADSen, arrazonamendu-prozesuko oinarritzko urratsa da, dagoen ezagutzatik informazio berria eratortzen duena.

ANHITZen domeinu-geruza eta inferentzi geruza batera bildu dira, ezagutza-base berean. Bildu bai, elkarrekin bildu dira, baina ez nahasi ordea, ezagutza-base berean egon arren azpi-modulu desberdinetan baitaude. Horrela jokatzek badu zerikusia hiztegiez eta hiztegi-erabilpenaz dugun ikuspegiarekin. Hiztegiak tresna aktibotzat hartuz gero, logikoa da oinarritzko inferentziak, hiztegi-informazioari estuki lotuta daudenak, hiztegien partetzat hartzea. Hiztegiek beren egiten dituzte inferentzia horiek, eta, horrela, inferentziak ere berrerabilgarriak dira testuinguru desberdinetan.

- *Ataza-geruza*: helburu jakin bat erdiesteko inferentziak nola konbinatu behar diren xedatzen da geruza honetan. Normalean atazak hierarkikoki antolatzen dira, goi-atazen helburuak lortzeko azpiatazak konbinatu eta egikaritzen direlarik. Bi molde erabili ohi dira ataza-ezagutza definitzeko: bata, estatikoa, atazen kontrol-egiturak ohiko kontrol-primitiboez adieraziz; bestea, dinamikoa, inferentzi ezagutza interpretatu eta planifikatuz.

Gure ereduaren ataza-geruzan *itzulpen-unitatea itzultzea* da gain-ataza. Atazak eredu estatikoaren arabera definitu dira, eta itzulpen-unitatea itzultzean hiztegia noiz eta nola kontsultatzen den islatzen dute.

¹² Modelizazio-hurbilpen, metodo eta lengoaiei buruzko kontuak § II.4.1 eta § II.4.2n azaltzen dira zabalago.

- Estrategi geruza: problema jakin baten ebazpenean zein bide urratu behar den finkatzen da, uneoroko egoerak kontuan edukiz eta tarteko helburu aproposenak zein diren zehaztuz.

ANHITZen strategi geruzan, molde dinamikoa erabiliz, ataza-ereduaren kudeaketa egiten da, atazen exekuzioa egoeraren arabera kontrolatu eta planifikatuz.

Esan behar da ANHITZ sistema eraikitzeke ez dugula KADS era hertsian erabili. KADS, guretzat, oinarria izan da, ildo nagusiak markatu dituen metodologia. ANHITZ prototipoaren garapenean ez dugu KADS —edo Common KADS— ereduko software-ingurune bereziturik erabili, KEE ("Knowledge Engineering Environment") baizik.

I.4 Txostenaren nondik norakoak.

Sarrerako honi beste bost kapitulu gehitu zaizkio txosten honetan.

Bigarren kapitulan ANHITZ eredu kontzeptuala deskribatuko da. Itzultzaileen portaera jaso da eredu horretan, itzultzaileak hitzak itzultzen jarri eta erakutsi dituzten joerak lengoaia formal batean adieraziz. Eredu horrek atazak ditu ardatz, eta konputagarria bada ere, kapitulu honetan espezifikazioaren ikuspuntutik aurkeztuko da.

Hiztegi-ezagutzaren errepresentazioari helduko zaio III. kapitulan. Hiztegi-ezagutzak, esan dugunez, bi alde ditu gurean: domeinu-geruza eta inferentzi geruza. Bereziki landuko dira kapitulu honetan ezagutza hiztegianitzaren eta eleanitzaren tratamenduak dakartzan gorabeherak.

IV. kapitulan ANHITZ eredua, jadanik espezifikatua, exekutagarri bihurtzeko antolamendua azalduko da, bai strategi geruza deskribatuz, bai sistemaren diseinua xehetuz. ANHITZ eredua ANHITZ sistema bilakatuko da kapitulu horretan.

Sistemaren funtzionamendua azalduko dugu hurrena, V. kapitulan. Modulu guztiak inplementatuta egon ez arren, ikuspegi funtzionala emango da, hala itzultzailearen kanpo-ikuspegia, nola barne-moduluen arteko elkarlana eta koordinazioa.

Azkenik, VI. kapitulan, lanari buruzko hausnarketa egingo da. Ondorio azpimarragarrienak, eta egindakoak izan dezakeen etorkizun eta jarraipena bilduko ditugu bertara.

II. Hiztegien erabilpenaren modelizazioa itzulpenean.

*"... Eusquerara biurtu nai bada, ori ondo biurtceco,
traducio onaren legueak gorde beardira:
eta ez itzac berac, ezpada oriek esan nai dutena,
edo orien sentidua ondo pisatu, ta Eusqueraz nola ondo esangodan
contuz consideratu, ta ori escribitu bearda."*

A. Cardaberaz
Eusqueraren berri onac (1761)

II. Hiztegien erabilpenaren modelizazioa itzulpenean.

Egin dezagun kontu hiztegi-informazio base bat daukagula eta informazio hori baliatu nahi dugula giza itzultzaile bati bere lana errazteko. Behinolako burutazioa da hori, ANHITZ sistemaren helburuak zehazteari ekin genionekoa. Hastapenekoa izanagatik, ordea, burutazio horrek ondo bereizten ditu sistemaren bi alde: batetik hiztegi-ezagutza, bestetik itzulpen-prozesua; lehena tresna, bigarrena ataza.

Hastapeneko hartatik aurrerago joan nahi eta, *Nola erabiltzen da hiztegia itzulpenean?* galdera otu zitzaigun hurrena. Galdera horretan ere bi alde ageri dira —hiztegia eta itzulpena—, baina bi alde elkarlotu. Elkarrekintzak azaleratzea eskatzen du galdera horri erantzuteak. Eta erantzunean dago erdua: itzultzen denean nola usatzen den hiztegia, horixe izango da ANHITZen erabilera-eredurako oinarria.

Itzulpen-iharduneko hiztegien erabilera-ereduaren deskribapena dugu kontagai kapitulu honetan. Hariaren hasierako muturrean —§ II.1— hiztegien erabilera-azterketen errepasoa egiten da, zabal eta larri, itzulpen-iharduneko erabileretara mugatu gabe. Gure komenientziako azterketaren berri ematen da § II.2n eta modelizazioak dakartzan gorabeheri helden zaie, ezari-ezarian, kontsiderazio orokorrekin hasita —§ II.3—. Itzulpen-prozesuaren modelizazioaz mintzoko bagara ere, eduki gogoan hiztegien erabileran jarri dugula arreta eta, beraz, maila lexikalera mugatuko garela. ANHITZ ereduaren xehetasunak § II.4.3 eta § II.4.4n ematen badira ere, deskribapenari bete-betean heldu aurretik § II.4.1en eta § II.4.2n aukeratu dugun modelizazio-metodologiaz eta lengoiaiaz argibideak ematen dira.

II.1 Hiztegien erabilpena aztergai.

Hiztegien erabilpenak aparteko sekreturik ez duelako ustearen kontra, azterketa askotxo egin dira, eta egiten dira gaur egun ere, gai horren gainean eta jira-bueltan. Gehienek giza erabiltzailearen portaera eduki dute aintzat, zein hiztegi den gogokoen, zer-nolako bilaketak egiten diren, eta antzeko galderei erantzunak emateko. Pentsatzekoa denez, lan horien bultzatzaile nagusi hiztegitzaileak izan dira, neurri-neurriko produktuak egin nahi eta kontsumitzaileen joerei erreparatzea erabaki dutenak.¹³

¹³ Gaude, hala ere, ez ote diren gutxitxo egin horrelako azterketak gure inguruan. Benetan zein hiztegi du premiazkoen euskal hitzuneriak? Gauza al gara horri erantzuteko? Aztertu al da aurriritzi eta usteez harago kontu hau?

Azterketa horietako batzuen batzuk ekarri nahi izan dira hona. Horien emaitza batzuk behintzat eta, batez ere, erabilitako azterketa-metodoak nabarmenduko dira. Aukeraketan lan praktikoak, erabiltzaileek eskaintako datu enpirikoetan oinarritzen direnak, hobetsi dira.

II.1.1 Aurrekariak.

R.R.K. Hartmann-ek (1985), ikerketen zertarakoei begiratuta, lau sailetan banatzen ditu hiztegi-erabilpenari buruzko lanak eta ondo kontuan hartzekoa da sailkapen hori, alor honetan, beste hainbatetan bezalaxe, helburuak erakusten baitu bidea eta metodoa. Zer jakin nahi den hiztegi-erabilpenaz, horren arabera apailatu beharko da azterketa. Hara, bada, zein den banaketa:

- "Hiztegien tipologia": sail honetakoak dira hiztegietan sartu beharreko informazio-kategoriez egindako lanak.
- "Erabiltzaileen tipologia": erabiltzaileak taldeka aztertu eta talde bakoitzaren berezitasunak azaleratzea da tipologia honetara bildu direnen xedea.
- "Beharren tipologia": hiztegiak zein testuingurutan eta zein beharri erantzuteko erabiltzen diren jakin nahi litzateke azterketotan.
- "Trebezien tipologia": erabiltzen diren kontsulta-estrategiak jorratzen dituzten lanak multzo honetan kokatzen dira.

Lau tipologia horiek elkar eragiten dutena nabaria da. Ez dago orientazio bakarreko azterketarik, ez dago horietako tipologia bat bakarra era isolatuan landu duenik. Haatik, Hartmann-en sailkapena aski baliagarria da nork zeri eman dion garrantzia jakiteko. Erabil dezagun hortaz.

Hiztegien tipologia

Clarence Barnhart-ek (1962)n ezagutarazi zuen hiztegi elebakar komertzialei buruzko azterlana. 108 galdeketa egin zituen eskolako hiztegi elebakarrek erabiltzaileen zein motatako galderei erantzun beharko lieketen argitzeko. Ohiko sei informazio-mota hartu eta ordenatu egin zituen ikasleen erantzunetan oinarrituta. Barnhart-en arabera esanahia da hiztegietatik usuen nahi izaten dena, ondoren ortografia, ahoskera, sinonimoak, erabilpena eta, azkenik, etimologia.

Zelko Bujas-ek (1975) eginkizun jakin batean aztertu nahi izan zituen hiztegien sartu beharrekoak, gogoan zuela hiztegi guztiek ez dutela berdin balio zernahitarako. Aldizkari modernoek ulerkuntza finkatu zuen egitekotzat eta 18 ikasle jarri zituen horretara, hiztegien

hutsune eta gabeziak agerian uzteko. Are, garaian garaiko hitz erabiliak jaso eta hiztegiak eguneratu beharra azpimarratu zuen. Erabilitako hiztegian, kasu batera, beste 2200 sarrera jasotzea proposatu zuen ondorengo edizioetarako.

James Baxter-ek (1980), ingelesa estudiatzen ari ziren 342 unibertsitateko ikasle japoniarren erantzunak jaso eta ikasleek hiztegi elebidunak erosteko duten joera azpimarratu zuen. Autorearen ustez, ikasgeletako portaerak ikusita, nabarmena da hiztegi elebidunen eragina. Ondorio gisa, hala dio (1980:333):

It is clear that, in light of the preceding analysis, increased use of monolingual English dictionaries should be encouraged.

Hiztegi elebakarrak, elebidunak eta "elebidunduak"¹⁴ alderatu zituzten Batia Laufer-ek eta Linor Melamed-ek (1994). Azterketa hau burutzeko 15 maiztasun txikiko sarrera eman zitzaizkien 123 goi-mailako ikasleri. Hamabost sarrera horietatik bost definizio elebakarrekin batera eman zitzaizkien, beste bost elebidunetik hartutako baliokideekin, eta azken bostak hiztegi elebidundu batean zegokien informazioarekin. Hitz horiek ulertzeko eta esaldi-sorkuntzan erabiltzeko gaitasuna neurtu zen gero. Ondorioak hitz gutxitan laburtzea zaila bada ere, esan dezagun hiztegi elebidunduak izan zirela eraginkorrenak batezbeste.

Badira beste lan batzuk hiztegietako informazioaren egokitasunaz, baina oraindik azterketa enpiriko sakonen beharra somatzen da, izan ere, zaila baita sakonki aztertu ezean egokitasuna neurtzea. Bien bitartean zenbait galdera hortxe dago airean: ba al du etorkizunik hiztegi orokor edozertarakoak? Lehenxeago aipatu diren sei informazio-kategoriez gain, ba al da sartu beharreko besterik? Gramatika- eta fonetika-kodeei zer irizten zaie? Definizioen egiturek ulergarritasuna bermatzen al dute?

Erabiltzaileen tipologia

Randolph Quirk-ek (1973) lan serioa egin zuen erabiltzaile-tipoen irudi errealistagoa lortzeko. Autore honek behaketa zuzenetik ez baina itaunketetatik jaso zituen datuak. 220 unibertsitateko ikasle itaundu zituen —30 galdera bakoitzari—, arteetakoak batzuk, zientzietakoak besteak, erdira gizonezko eta emakumezkoak. Saio hau, metodologi gabeziak gorabehera, presente edukitzekoa da.

¹⁴ Hiztegi "elebidundu" hauek hibridoak dira, informazio elebakarra eta elebiduna batera biltzen dute. Hiztegi ingeles bertsio elebidunduak pasa den hamarkadan agertu ziren. *Oxford Student's Dictionary*-aren hebraitarrentzako bertsioa halakoxea da.

"Nolako erabiltzailea, halako hiztegia", lelo horrek ederki adierazten du ikuspuntu hau landu dutenen ustea¹⁵. Erabiltzailearen nolakotasunak dira aztergai, eta ahaztezin gainera, hiztegi berezitu eta egokituaren defentsa egingo bada behintzat (Tomaszczyk, 1979).

Hiztegi-saltzaileek, jakina, egiten dituzte merkatu-azterketak jakiteko zeinek zer hiztegi erosten duen, baina azterketa horien emaitzak nekez publikatzen dira. Horiek alde batera, lexikografo askoren susmoek bere horretan dira, susmo bai, baina berrespenik gabe: erabiltzailearen ingurune kulturalak —herria, adina, heziketa, e.a.— duen garrantziaz, esaterako, zer dakigu?

Beharren tipologia

Jerzy Tomaszczyk-ek (1979) hizkuntza arrotza ikasten ari direnen beharrak aztertzen, 450 bat galdeketa —57 galdera— egin eta hizkuntz ikaskuntzako iharduera desberdinei buruzko datuak jasoz. Garbi asko erakusten du lan honek hiztegi-erabilera iharduera-motaren eta trebetasunaren arabera dela.

Hain zuzen ere, iharduera desberdinei erreparatuz, bereziki sorkuntza aktiboari eta ulerkuntza pasiboari, saiatu zen hiztegi-beharrak aztertzen Henri Béjoint (1981). Ingelesa estudiatzen ari ziren 122 ikasle frantsesi egin zizkien 21na galdera eta ohartu zen hiztegiak sorkuntzan baino areago ulerkuntzan erabiltzeko joera zegoela, eta gainera, ez zutela behar bezalako probetxurik ateratzen hiztegi-takako informazioez sortze aktiboan ari zirela.

Hartmann-ek (1983) 67 irakasle eta 118 ikasle britainiar itaundu eta ondorio benetan interesgarriak atera zituen, besteak beste, premia nagusiek esanahiarekin nahiz gramatikarekin dutela zerikusia eta, oro har, erabiltzaileei ezer gutxi erakusten zaiela hiztegi-erabilpenaz.

Trebezien tipologia

1983an argitaratutako txosten batean Evelyn Mitchell-ek irakurketa-estrategiei buruz egindako lan baten berri ematen du. 94 eskola-ume eskoziarri 5 testu irakurri eta ulertzea jarri zien eginkizun eta, haren esanetan, ikasleek zailtasun handiak izan zituzten. Antza denez, oso prozesu konplexua da sarrera egokia aurkitzea, sarreraren egitura ulertzea, definizioaren zatirik behinenak identifikatzea, testuinguruarekin bat datorren adiera zein den erabakitzea eta, azkenik, hitza testuinguruan txertatu eta esanahia biribiltzea. Hain zaila omen da ezen, Marsha Bensoussan-ek (1983) erabakitakoa hona ekarriz, hiztegiaren erabilpenak ez duen hainbesterainoko eraginik ulerkuntzan. Hara zergatik: asko dakitenek beharrik ez dutelako eta

¹⁵ Ez dezagun ahaztu, hala ere, badela helburu orokorreko hiztegiaren aldeko agertzen denik.

gutxi dakitenek hiztegia ondo erabiltzen ez dutelako. Eta, egia esateko, lagin handi baten azterketa estatistikoan oinarritu zen hain baieztapen harrigarria egiteko.

Glyn Hatherall-ek (1984) ere harritu gaitu beste zenbait ondorioekin. Alemanieratik ingeleserako itzulpenetan hiztegia nola erabiltzen den aztertzeko 22 ikasle jarri zituen lan horretan. Datuak ateratzeko itzulpenak behatu eta jaso egin zituen aurrena, eta galderak egin zizkien ikasleei ondoren. Ondorioetatik pare bat nabarmenduko ditugu. Bata: ikasle aurreratuagoek maizago jo zuten hiztegiara¹⁶. Bestea: ia guztiek, aurreratuek barne, hitzez-hitz itzultzeko joera erakutsi zuten.

Goraipatzekoa da testuinguru honetan Josh Ard-ek (1982) egindako azterlana, bai hiztegi elebidunen erabilera mugatzeko egin zuen aportazioagatik¹⁷, bai erabilitako metodoagatik. Autore honek ikasle ingelesan idazketa-iharduera ikertu zuen, datuak jasotzeko bi ikasleren portaera filmatuz eta, horrez gain, ahozko protokoloak baliatuz.

EURALEX/AILA proiektuaren baitan (Atkins B.T. & Knowles F.E., 1990) hiztegien erabilpena eta eraginkortasunari buruzko azterlan xehe eta zabala burutu zen. Ikasle frantziar, alemaniar, italiar eta espainolen —lau ezagutza-mailatan sailkatuta— 1100 erantzun maneiatu ziren egindako testean. Datu eta portzentaia askotxo lortu ziren, eta interesgarriak alafede, baina hona ondorio orokor batzuk baizik ez ditugu ekarriko. Hala, garbi nabarmendu zen testu-ulerkuntzan eta itzulpenean hiztegi elebidunak gehiago usatzen zirela eta hitzen erabilpen-informazioa nahi zenean, oster, elebakarrak nahiago izaten zirela. Orobat azpimarragarriak dira lan horretan hitz anitzeko sarreraren bilaketa-teknikez eta informazio gramatikal nahiz metalengoaiazkoaren erabilpenari buruz idatzitakoak.

Stefania Nuccorini-k (1994) berriki plazaratutako lan batean hiztegi-erabilpen okerraren arrazoia trebezia-falta dela azpimarratu du, edo zehatzago esanda, oker gehiago sortzen dela hiztegiak gaizki erabiltzeagatik hiztegien gabeziengatik baino. Egilearen iritziz, askotan hiztegia ez erabiltzea da erroreen kausa (jakintzat hartzen diren hitzak hiztegien kontsultatu ez eta, horra errorea). Probetan unibertsitateko ikasleek hartu zuten parte; 222 itzulpen-lan —ingelesez italierara— aukeratu eta egindako erroreak sailkatu eta analizatu egin ziren. Bost errore-kategoria bereizi ziren: erabilpen desberdinak, hitz homonimoak eta polisemikoak, adierazpen idiomatikoak, hitz deribatuak eta hitz konposatuak.

Orain artean aipatu diren lanak baliagarriak eta oso gogoan hartzekoak badira ere, hain dira ugariak gaur egun hiztegi-, erabiltzaile-, behar- eta trebezia-tipoak, ezen oraindik hiztegi-

¹⁶ Agian, zioen Hatherall-ek, ikasle ez hain aurreratuek informazioa edireteko duten konfiantza falta izan daiteke horren kausa.

¹⁷ Geroxeago, II.3 atalean, ukituko dira hiztegi elebidunen erabilerak dakartzan arazoak.

erabileraren alorrean zer landu asko dagoen. Azterketa sakon, fin eta zientifikoen beharra dago. Azterketa-teknikak eta metodologia ere estandarizatu egin beharko lirateke. Are, informazioa biltegitatu eta eskuratzeko teknika berrien ekarpenek berealdiko aldaketak eragin eta eragingo dituztela-eta, etorkizuneko hiztegiak nolakoa izan behar duen definitzean erabilera-azterketek pisu handia beharko lukete. Gu behintzat horretan gaude.

II.1.2 Hiztegi elektronikoen erabilera.

Gero eta zabalduago, gero eta arruntago ditugu hiztegi elektronikoak. Patrikako hiztegietatik entziklopedia handietaraino bada produktu anitz euskarri elektronikoan merkaturatu dena. Zenbaiten ustez tresna hauek hain dira diferenteak paperezko hiztegietatik ezen produktu desberdintzat har daitezkeen. Produktu diferenteak izan ala ez, informazioaren egituraketak eta erabiltze-aukerek agintzen dute eta ez hainbeste euskarriak berak. Erabilera da, hortaz, giltzarri. Eta erabilera-azterketaren premia nabarmen ikusi bada orainokoa, paperezko hiztegiatan, ez da gutxiago nabarmentzen hiztegi elektronikoei heltzean.

(Marchionini, 1989)n trataera egokia ematen zaio paperezkoetatik entziklopedia elektronikoetara pasatzeak dakarren erabilera-egokitzapenari. Hasteko, dio Marchioninik, desberdintasun kontzeptualak eta fisikoak bereizi behar dira bi ereduaren artean. Kontzeptualki ikusita, sistema elektronikoen aukera-ugalketak konplexutasuna eta, ondorioz, beste trebezia kognitiboen beharra eragiten du. Protokoloen, elkarrizketen eta ahozko simulatzearen bidez jasotzen ditu egileak entziklopedia baten bi bertsioaren —bata paperezkoa, bestea elektronikoa— erabilpenari buruzko xehetasunak. Autoreak dienez paperezkoetan ohitutako erabiltzaileek ez dituzte bete-betean aprobetxatzen baliabide elektronikoak, besteak beste, kontsultak egiteko, nabigatzeko edota informazioa iragazteko eskaintzen zaizkien aukerak.

Orobat da interesgarria (Large and Beheshti, 1994)n plazaratutako lana. Honakoan 48 seigarren mailako ikasleren bilatze-denbora eta bilatze-teknikak konparatzen dira, *Compton's Encyclopediaren* paperezko eta CD-ROM bertsioak erabiltzen dituztela. Ikasle gazte eta eskarmenturik gabeek neke handirik gabe ikasi eta erabiltzen omen dituzte bilatze-bide desberdinak. Dena dela, paperezkoa erabili edo CD-ROMa erabili, ez da emaitzetan alde handirik antzematen.

Nahiz eta hiztegi edo entziklopedien esparrutik aldentzea ere baden, ez genduke aipatu gabe utzi nahi Paul Solomon-en azterlana (1993) Lineako Atzipen Publikoko Katalogoei buruz. Gure irudiko, esparru-aldaketa gorabehera, lan honetako problemaren formulazioak, *"the challenge is to understand why some people succeed and others fail in order to develop flexible, anticipatory mechanisms to support users as they explore information structures to*

retrieve information" (1993:246), eta erabilitako metodologiak argibideak eman ditzakete hiztegi elektronikoen erabilera aztertzerakoan ere.

Kontsultak, nabigazioak, bilaketak eta informazio-egiturak aipatzen ditugula, hiztegi eta entziklopedia elektronikoak informazio-eskurapenerako sistematzat hartu behar dira. Gisa horretan, interfaze gero eta lagungarriagoak, hipertestuak, eskariak formulatzeko laguntzak, beste aplikazio informatikoen integrazioak, multimedia baliabideak, horiek denak aztertu behar dira sistemon ahalbideak ebaluatzeko. Azken batean Marcia J. Bates-ek (1990)en aireratzen duen galdera dago: *Where should the Person Stop and the Information Search Interface Start?* Eta bistakoa da muga dinamikoa dela hori, hots, sistemek gero eta "eremu" zabalagoak hartzen dituztela beren gain.

II.2 Gure azterketaren deskribapena.

Hastapenetik, sistemaren alde funtzionala mamitzen hasi aurretik, erabiltzailearen neurriko sistema diseinatu nahiak markatu du gure egitekoa. Erabiltzailearen beharrak hartu nahi izan ditugu ipar, eta haiei eman beharreko erantzunak, ahalik eta aberatsenak, jomuga. Arestian azaldu denez, badago besteren lanik eta eskarmenturik, ondotxo baliagarriak, baina horien artean gure tamainakorik, gure asmo zehatzei doi erantzuten dionik badenentz zalantzak izan ditugu. Egia esan, nahiago izan dugu besteek egindakoa aberasgarritzat hartu eta gureari ekitea. Hiru arrazoren kariaz jokatu dugu horrela: batetik, helburu —eta jakinmin— desberdinak ditugulako; bigarrenik, azterketa-teknika gure erara atondu nahi izan dugulako, eta atzenik, ingurunea bera ere, euskara tarteko, besterena ez bezalakoa delako.

Helburuei dagokienez, esan dezagun gure interesek ukitu kualitatiboa dutela, kuantitatiboa bainoago. "Zenbatetan", "zenbatek", "ze portzentaiatan" estiloko galdegaiak gutxi-edo axola zaizkigu, beste "nola", "nondik", "zergatik", "zertarako" horiekin alderatuta. Beharren eta erabileren ikuspegi zabala eduki nahi izan dugu, hartara ANHITZek era askotako kontsultak egiteko aukera eman diezaien itzultzaileei. Oinarrizko funtzioak kontsulta-behar horien isla dira. Gainera, oinarrizko kontsultatan mamitzen ez diren portaerak ere ez ditugu alde batera utzi nahi izan. Horiek, goi-mailako betekizunetan adierazita, prozesuaren deskribapen orokorrean ageriko dira. Ez dezagun ahantz gure xedea unitate lexikalen itzulpena modelizatzea dela, bereziki hiztegiekin gertatzen den elkarrekintza nabarmenduz.

Behin helburuak argituta, azterketaren nondik norakoak finkatzea dihoakio horrelako egitasmo bati. Xehetze lan horretan badira nahitaez ukitu beharreko gaiak, eta ukitu ezezik, horien inguruan erabakiak hartzea ere beharrezkoa da azterketa taxuz burutuko bada. Ikus dezagun zein diren landu beharreko gairik behinenak.

Azterketaren ezaugarriak xedatzeko irizpideak

- *Informazioa jasotzeko teknika:*

Ohikoenak itaunketa egituratuak, behatze-teknikak eta protokoloak dira (Johnson & Johnson, 1991).

Galdeketak egokiak dira oinarrizko arau eta jokamoldeak, probabilitate txikiko gertaerak eta arrazoibideak azaleratzeko. Oso maneigarriak dira eta hasierako hurbilpenetarako egokiak. Zaila da, ordea, xehetasunetara ailegatzea eta ez dago jakiterik zein mailataraino den fidagarria jasotako informazioa. Azkeneko hau da duten akats nagusia.

Behatze-teknikak bereziki aproposak dira hipotesiak berresteko eta, eginkizun konplexuetan, zehazkizunez jabetzeko. Baina hauek ere ez dira aitzakiarik gabeak. Teknika hauetan datuen interpretazioa ez da berehalakoa eta inferentzia eskatzen dute ikusitakotik prozesuaren egitura deskribatzeko. Horrez gain, zuzeneko behatze-teknikei, behatzailea presente dagoela egiten direnei, erabiltzailearen ihardueran eragitea eta portaerak aldaraztea eragozten zaie. Zeharkakoek, bideo-kameraz egiten direnek kasu baterako, giza behatzailearen ausentzia dute aringarri.

Protokoloak ahozko txostenak dira, ari denak, egileak, eskainitakoak. Betebeharraren egilea eta hizlaria pertsona bera izaten da. Protokoloen artean gutxienez bi mota bereizi behar dira: batzuei konkurrente deritze, eta hauetan kontakizuna iharduerarekin batera gertatzen da, egin ahala kontaktzen da; besteei atzerantzkoak esaten zaie, egin ahala gogoan hartu eta egitekoari buru emandakoan, orduan eta ez lehenago, kontaktzen baita. Protokoloek, batez ere konkurrenteei, ariera oztopatzea egozten zaie. Dituzten gabeziak estaltzeko, mesedegarri izaten da behaketak eta protokoloak elkar osatuz erabiltzea.

Hiztegi-erabilpenari buruzko azterketetan, orain artekoetan, galdeketak erabili dira gehienbat. Baina ahots kritikoak ugalduz joan dira eta metodo horiek zalantzan ere jarri ditu zenbaitek. Hara zer galdetzen zuen Hatherall-ek (1984:184) juxtu hiztegi-erabilpena aztertzeke egiten ziren galdeketez ari zela:

Are subjects saying here what they do, or what they think they do, or what they think they ought to do, or indeed a mixture of all three?

Eta aurreraxeago, beste pasarte batean, horrela zioen:

I conclude that, whatever the difficulties, the only reliable method of collecting data on dictionary user behaviour is by direct observation.

Komeni da, guk uste, zuri-beltzeko baieztapen hori ñabartzea¹⁸: zuzeneko behatzea ez da akatsik gabea eta gainerako teknikak ere fidagarriak dira hein batean. Behatze-teknikak hobetsi, besterik ez gutxietsi: horra gure iritzia.

- *Zeini eta zenbati egin probak:*

Adituei erreparatu ohi zaie, hauengandik jasotakoa fidagarriago eta baliagarriagoa delakoan. Eta hala bide da. Baina hiztegi-erabilpenean zein da aditu?

Itzultzaile profesional batek azkar esango dizu hiztegira maiz jo behar duen itzultzaileak baduela ajeren bat. Itzultzaileek oso ondo ezagutzen dituzte hizkuntzetako lexikoak eta, hiztegia erabili bai baina, oso arazo konkretu eta espezifikotan jotzen dute hiztegira. Hauengandik jasotakoak munta handia du, baina ez da nahikoa.

Beste muturrean hizkuntz ezaguera handirik eduki ez eta normalean itzulpengintzan aritzen ez direnak leudeke. Hauek, bistan da, sarri joko dute hiztegira itzultze lanetan jarritz gero. Baina hauen iharduera ezin liteke erabilera-eredutzat hartu.

Bada hirugarren talde bat: itzulpenak egiten nolabaiteko ohitura —ez da ariketa arrotza— dutenak, baina hizkuntzaren bat, edo biak, hain ondo menperatzen ez dutenak. Hauek hiztegira maiz joko dute, eta intentzior jo ere eskuarki. Esan daiteke itzulpengintzan aditu ez badira ere, hiztegien erabilpenean badakitela zertan ari diren. Guretzat aparteko arreta merezi dute, itzultzaile profesionalek baino hiztegien behar handiagoa dutenez, beharrak eraginda, kontsulta-espektro zabalagoa erakuts baitezakete.

Beste kontu bat da datuak jasotzeko probak zenbat laguni egin beharko litzaiekeen erabakitzea. Erabaki hau estuki lotuta dago azterketa-teknika eta helburuekin. Nabaria da, konparazio batera, azterketa estatistikoetarako lagin handia beharko dela, nabaria den bezala behatze-teknika erabiliz gero lagina murriztu beharra dagoela. Batzuetan datuen ugaltzeak ez du ezagutzaren ugaltzea ekartzen, besteetan aldiz derrigorrezkoa da.

- *Zer datu jaso:*

Zer jakin nahi den, hark erakutsiko du zer jaso. Baina ez gaitezen sinplekeriatan eror: ez da hain nabaria datuen eta ezagutzaren arteko erlazioa. Aztertzaileak oso

¹⁸ Autore honek berak kontuan hartzen zituen zuzeneko behatzearen mugak eta akatsak, eta horregatik formula misto bat —"idatzizko protokolo" deituko genioke— erabili zuen.

kontuan edukiko du zein datu izango zaizkion baliagarri zerbaiti buruzko ezagutza areagotzeko. Baina, hala ere, maiz zeharkako datuek, ustez munta txikikoek, bistararazten dute portaeraren alderdi nabarmenen bat¹⁹. Horiei tarte bat egitea komeni da. Datu egituratuak gune gisa hartuta, soilik horiek erabili ordez, aberasgarri dateke molde horietatik kanpo gertatzen dena ere jasotzea, nahiz eta gero datuon interpretazioa askoz ere konplexuagoa gertatuko den.

Beste kontutxo bat ere bada aipagarri: zenbateraino izango dira xehe eta fin jasotako datuak? Galdera horren inguruan ere erabaki beharra suertatzen da. Ez da gauza bera, bistan denez, erabiltzaileak hiztegi-kontsulta bat egin duela esatea, edo kontsulta sarrera jakin bati buruzkoa izan dela aipatzea, edota sarrera baten n-garren adierari dagokion definizioa kontsultatu duela nabarmentzea. Datu larriagoak komeniko dira batzuetan, finagoak besteetan.

- *Denbora:*

Denborak zerikusi handia izan lezake azterketaren nondik norakoan. Hasteko, denbora-mugapena jartzearen komenientzia ikusi behar da. Autore batzuek, Hatherall-ek (1984) edo Nuccorini-k (1994) kasu baterako, denbora faktorea kontuan izan dute azterketak egitean. Hatherall-en ustez hiztegien egokiera neurtzeko kontsultek eramaten duten denborak ere badu zeresana, azken batean, bilaketetarako erraztasunak eta informazioaren eskuragarritasuna denbora-unitatetan ere neurtzen baita. Nuccorini-k denborak eragiten duen estuasuna eta presioa bazterrean utzi gabe jorratu zituen itzulpenean egiten diren erroreak²⁰.

Dena dela, ez dakigu zein mailataraino den arrunta denboraren presiopean hiztegiak kontsultatzea. Beti dago, nola ez, denbora faktore hori, bazterrezina alafede, baina zorrotasun gehiegiak azterketa desitxuratzea ere ekar lezake. Ez zaigu egoki iruditzen patxadaren kalterako izango den denbora-mugapen estua ezartzea, non eta helburuen artean denbora bera neurtzea ez dagoen.

- *Nolako lana:*

Unitate lexikalen itzulpenean hiztegi-erabilera zer-nolakoa den aztertu nahi bada, argi dago lan horretantxe jarri beharko direla azterketariak. Testuak eta hiztegiak beharko dira halabeharrez: testuak, langai; hiztegiak, lanabes.

¹⁹ Horregatik ez gara gu aldeztu aurretik finkatutako galdeketen zale, hauek egitura zurruna baitute eta zaila baita ustekabekoei bidea irekitzea. Normalean portaera estereotipatuak erakusten dira, ez ordea aurririkusi gabeak.

²⁰ Examina batean egindako itzulpen-lanak aztertu zituen autore honek.

Probetarako testuak aukeratzea ez da nola-hala egitekoa, kontu handiz baizik. Hala egin ez eta testu desegokiak hautatuko balira, testu horiek datu-bilketaren asmoak zapuztu egin ditzakete. Horregatik, zein hizkuntzatan dagoen idatzia, arloa, luzera, zailtasun-maila, hori guztia ondo neurtu beharrekoa da testua aukeratzean. Eskuarki testu zailak eta ez oso luzeak nahiago izaten dira, betiere espazio txikiago eta maneigarriagoan datuak biltzearen aldera.

Hiztegiei dagokienez ere bada zer erabaki eta finkatu. Ikuspegi zabalenetik —nork bere gogoko hiztegiak erabil ditzala— murrizteneraino —hiztegi bakar bat jartzea erabilgarri— nahi beste aukerak leudeke²¹. Gai honetan oso zaila da irizpide orokorreaz hitz egitea, azken finean, jaso nahi diren datuek zeharo baldintzatzen baitute hiztegien nolakoa eta kopurua.

Gure azterketaren nondik-norakoak

Gurera etorriz, aipa dezagun lehenik eta behin datuak jasotzeko erabili dugun teknika zuzeneko behaketan eta ahozko nahiz idatzizko protokoloetan oinarritu dela. Jakitun gara behariaren presentziak eragin dezakeen distortsioaz, baina naturaltasunetik gehiegi ez urruntzearen azterketarien lana presiorik gabe eta konfidantzaz egiten saiatu gara. Presiorik gabe, ezen hasiera-hasieratik jakin baitute probak ez direla inolako examina edo iaioitasun-erakusketa, eta konfidantzaz, behatzailea laguna edo gutxienik ezaguna izan baita²². Protokoloak, arestian esanda bezala, gogaikarriak gerta daitezke, eta ondorioz portaera-aldarazleak. Itzultzea berez nola-halakoa bada ere, lanean ihardun ahala kontatzeak —edo idazteak— are eta astunagotzen du ariera. Zama hori arintzeko behatzailearen presentzia aktiboa bultzatu dugu. Ikusitakoa, begien bistakoa, behariak jaso du besterik gabe²³. Hori bai, ikustezina dena —asmoak, helburuak, etabar— azterketarien ahotik jaso behar izan da ezinbestez.

Itzulpen-lanaren ildo nagusiak aztertzaileak xedatu ditu, hark aukeratu ditu testuak, haren esku egon da zein testu-zati itzuli behar den erabakitzea. Batzuetan parrafo osoak itzuli behar izan dira, beste testu batzuetan, aldiz, esaerak, lokuzioak edo hitzak itzultzea agindu da. Horrelakoetan aurrikusitako arazo konkreturen bat aztertu nahi izan da. Beraz, espero izatekoa zenez, aztertzailea agintari ere izan da nolabait.

²¹ Erabiltzen diren produktuak guztiz desberdinak izaterainoko aldeak egon litezke aukeren artean: batek, konparazio batera, paperezko hiztegia erabil lezake eta beste batek, berriz, elektronikoa.

²² Komeni da halakoetan irakasle-ikasle edo beste edozein erlazio hierarkiko ebitatzea, ahal bada, azterketaria eta behatzailea "maila berean" egon daitezela bermatzea.

²³ Estreinako probetan paperean jaso dira zuzenean, baina gerora, idazteak luze jotzen duela-eta, zinta magnetikotan erregistratu dira.

Haatik, komeni da jakitea aztertzaileak finkatutako lan-esparru horretan azterketariak askatasun osoa izan duela zernahi egiteko, bai edozein hiztegi aukeratu eta erabiltzeko, bai bere kasa itzultzeko. Aztertzailearen aginpidean ezarritako muga horrek —edo, beste era batera esanda, azterketariari emandako autonomia horrek— emaitza positiboak ekarri ditu. Izan ere, horri esker detektatu baitira alde aurretik ikusten ez ziren portaerak eta hiztegi-erabilerak.

Zazpi lagunek hartu dute parte itzultze-lanetan, horietatik lau itzultzen dezente arituak —ezin aditutzat hartu, ordea²⁴— eta beste hirurak itzultzaile profesionalak. Lanbidez desberdinak diren bezala hizkuntz ezagueran ere alde nabarmenak daude beraien artean. Heterogeno izate hori, egia esateko, apropos bilatu da, hiztegi-erabilera desberdinak azaleratzeko intentzioz.

Testuak ere, hala berean, heterogenoak izan dira: bat, frantsesez idatzitako testu teknikoa, bestea, frantsesezko literarioa, eta hirugarren bat, aztertzaileak idatzitako euskarazkoa. Itzulpenak, denbora estuasunik gabe, norabide bietan egin dira: frantsesetik euskarara eta euskaratik frantsesera. Testuekin batera hiztegiak ere jarri zaizkie eskueran itzultzaileei²⁵, eta ezaugarri desberdinetako hiztegiak jarri ere: Xabier Kintanaren Hiztegia 2000, Ibon Sarasolaren Hauta-Lanerako Euskal Hiztegia, Euskararako hiztegia - Adorez2, Izagirrerren Lokuzioen Hiztegia, Lhanderen "Dictionnaire Basque-Français", Le Plus Petit Larousse, Sinonimoen hiztegia - Adorez3, UZEI hiztegi teknikoak, Larousse Français Espagnol²⁶. Erraztasunak ematearren jokatu dugu horrela, ez inola ere itzultzaileen aukeramen mugatzearren.

Era horretan, itzultzaileen gorabehera, asmo eta ebazpide guztiak erregistratu ditugu. Hiztegira egindako kontsulta bakoitzeko erregistro antzeko bat osatu da ondoko informazioaz: hautatutako hiztegia, kontsultatu nahi izan den terminoa, zein hiztegi-sarrera atzitu den, kontsulta-mota, emaitza eta, komeniz gero, bestelako oharrak edo argibideak. Kontsulta-erregistro horiek sekuentzialki ordenatuta biltegitatu dira, hartara ikuspegi diakronikoa ere osatzeko. Erregistro-sekuentzia horretan dago gordea gure aztergai nagusia —unitate lexikalaren itzulpena hiztegiaren laguntzaz— modelizatzeko behar dugun informazioa. Baina, ez gaitezen aurrera, modelizatzearen berri eman aurretik egin diezazkiogun kritika batzuk burututako azterketari.

²⁴ Barka biezagute arinkeriaz botatako iritzi hau. Zehatzago esatera, profesioz itzultzaileak ez direla esan beharko genuke.

²⁵ Itzultzaile profesionalak ez diegu hiztegi-eskaintzarik egin, bestela ere ondotxo hornituta zeuden-eta.

²⁶ Asko hitz egin liteke hiztegi hauen egokieraz, baina ukaezina da arrunt samarrak direla gure artean eta euskara-frantsesa itzulpenak egiten hasita horrelatsukoak erabiltzen direla.

Kritika batzuk geure buruari

Aitor dezagun, erreparorik gabe, egin den azterketa kualitatiboa osatzeko —eta indartzeko— ez litzaigukeela gaizki etorriko datu kuantitatiboago batzuk ezagutzea. "Zenbat kontsulta egiten dira, batez beste, unitate lexikal bat itzultzeko?", edo, "zer arrakasta-tasa lortzen da kontsulta-mota bakoitzeko?". Tipo horretako galderei emandako erantzunak izan zitezkeen deduzitu ditugun erabilera-oinarrien sendogarri. Baina horrelako datuak jasoko badira beste era bateko teknikak erabili behar dira, besteak beste, lagin handiagoak eta hurbilpen estatistikoak.

Itzultzaile profesionalekin burutu den azterketak ere ez gaitu gogobete. Bagenekien alde zurretik zaila izango zela hauekiko langintza, baina emandako testuek ez dituzte behartu hiztegia erabiltzera. Eta, jakina, hiztegia erabili ezean hankamotza da guztiz gure azterketa. Ordainez, elkarriketatu egin ditugu, eta egia esateko, elkarriketa horiek zenbait gaitan sakontzeko baliagarriak izan dira. Hizpide izan ditugu, adibidez, itzulpena zuzendutako hiztegiek beharko lituzketen ezaugarriak, lexiko egokia aukeratzeak dakartzan buruhausteak, itzultzaileari laguntzeko sistema konputazionalen premia edo sistema hauek eskaini beharko luketen funtzionalitatea. Hori guztia grabatuta geratu da zinta magnetikotan, gerora ANHITZ sistemaren funtzionalitatea espezifikatzeko erabili diren zinta horietan.

Bestalde, hiztegi elektronikorik ez dugu erabili, eta hori ere bada kritikatzeko modukoa. Azterketa egin genuenean aukera gutxi izan genuen horretarako. Geroztik bai, agertu da hiztegi elektronikorik eta posible litzateke gaur egun euskara-frantsesa itzulpen-saioak tresna informatikoez egitea. Balio beza kritika honek etorkizuneko asmoen adierazgarri gisa.

II.3 Modelizaziorantz: kontsiderazio orokorrak.

Azterketariei ikusi eta entzundakoek eman digute zer pentsa eta zer hausnar. Zertzelada puntualetan oraingoz sartu gabe ere, ez zaizkigu alboratzekoak iruditu itzultzaileen arierei ikusi eta igarri dizkiegun jite orokorrak. Azken horiek erabiliko ditugu idazgaitzat atal honetan.

Eta hasi, konstatazio biribil batekin hasiko gara: hiztegia erabiltzean zailtasun handiak topatzen dira, batzuetan porrot egiterainoko zailtasunak. Uste hori aski hedatua dago hiztegi-erabilera aztergai izan duten autoreen artean eta ederki baieztatu dugu egindako azterketan. Besterik da, ordea, zailtasun horiei aurre nola egiten zaien galdetzean edo behatzean. Hor ez dago adostasunik, hor "ahal den bezala" esateko —egiteko— joera da nagusi, eta, noski, hori orokorra, orokorregia da guk nahi dugunerako.

Nolako hiztegiak

Pitin bat xehetu nahi eta azkar asko planteatzen da duda-muda betierekoa: nolako hiztegiak dira erosoena edo zein hiztegi-motak sortarazten du arazo gutxien? Eta, horiekin batera, nolakoa behar luke propio itzulpena zuzendutako hiztegiak? Badirudi, bapatean esanda, eleaniztun —gutxienez elebidun— beharko lukeela itzulpenarako hiztegiak. Baina beta pixka bat hartuz gero bestelako iritziak ere, elebakarren aldekoak, indar hartzen du. Autoreak ez datoz bat iritziak ematean. James Baxter-ek (1980) hiztegi elebakarren aldeko apustu argia egiten duela ikusi dugu. Glyn Hatherall-ek (1984) ere, gutxi edo gehiago, iritzi horri eusten dio. Baina ez dira hain garbi mintzatzen beste batzuk: Ard-ek (1982) ez du uste hiztegi elebidunak ez erabiltzeak errore lexikalen kopurua gutxitzen duenik; (Atkins & Knowles, 1990)en zertarakoen arabera juzkatzen dituzte hiztegi elebidunak —eta ez beti gaizki—; (Laufer & Melamed, 1994)en hiztegi elebakarrak, elebidunak eta "elebidundak" alderatzen dira, *which are more effective, for what and for whom?* galderari erantzuna emateko saio batean²⁷; eta Béjoint eta Moulin (1987:104) bakoitzari berea ematen ahalegindu dira hitzotan:

bilingual dictionaries are ideal for quick consultation in many cases, and monolingual dictionaries, though more difficult to use, have the extra merit of introducing the user right into the lexical system

Guk ateratako ondorioa ez da izan hiztegi elebidunak erabat baztertzekoa. Esan liteke, eta horretan ados geundeke, hiztegi elebidunek erakusten duten "zuzeneko bidean" informazio gutxiago eman ohi dela eta, beraz, errore lexikalak areagotu egiten direla. Baina, hori horrela izanda ere, elebakarren "zeharkakotik" joaten direnek ere arazoak izaten dituzte, informazio gehiago jasotzetik, informazio hori ondo ulertu eta maneiatzera koxka handia baitago. Garbi asko ikusi dugu zenbaitek hiztegi elebidunak kenduz gero ez lukeela beti itzulpenik aurkituko, eta elebidunekin berriz, nahiz eta errore arriskua areagotu, itzulpenak errazago lortzen dituela.

Itzultzaile ez-adituak eta adituak

Halaxe bada, erabiltzaile ez-adituak dira —hizkuntz ezaguera handia ez dutenak bereziki— elebidunetan makulu aproposa aurkitzen dutenak. Hauen jokabidea xedebaliokidea lehenbailehen aurkitzea da, baita jatorrizko hitzaren esanahia bete-betean harrapatu ez badute ere. Elebidunetarako joera hori, guk antzeman dugunez, hiztegi elebakarretako definizioak —eta gainerako informazioa— ulertzeko dauzkaten zailtasunek eragiten dute. Batzuetan definizioetako hitzen bat konprenitzen ez dutelako eta besteetan

²⁷ Eta hauen esanetan, hiztegi elebidunak baliagarriak izan edo ez, norentzat eta zertarako diren izaten da.

definizioaren egitura bera konplexuegia delako, kontua da zaila gertatzen zaiela sarrerako unitatearen esanahia atzematea. Eta hor ez du balio elebakarrek informazio gehiago ematen dutela esatea, hala izanagatik, ez baitira nahikoa lagungarri itzulpena burutzeko²⁸. Kuriosoa da, bestalde, itzuli ezinak eraginda, zenbaitetan hirugarren hizkuntza bat —gaztelania normalean— erabiltzera jo dela, eta hura zubi edo pibotetzat hartu dela itzultzera iristeko. Inguru elebidunak ezezik, eleaniztunak ere mesedegarri direla iradokitzen du azken puntu horrek.

Ez-adituen ihardunean hizkuntza batetik besterako jauzia baliokide-bilaketaz egiten den bezala, adituenean ulerkuntza eta sorkuntza prozesuak garatzen dira eskuarki²⁹. Hiztegi-kontsultak, orain maiztasun txikiagoz gertatzen direnak, iharduera horri lotzen zaizkio eta ondorioz hiztegi elebidunek zer eskaini gutxiago izaten dute. Elebidunen indar galtze horrekin batera ulerkuntzarako eta sorkuntzarako hiztegiak nagusitzen dira. Eta horra eztabaidatzeko gaia: hiztegi arruntak bietarako, ulertzeko nahiz sortzeko, al daude pentsatuta? Egin ditugun behaketek eta protokoloek garbi erakutsi digute ezetz, ulertze-lanetarako daudela pentsatuta, baina ez inola ere sorkuntzarako. Salbuespenak salbuespen, hiztegi-laguntza aberatsagoa jasotzen da testu-ulertzen testu-sortzen baino. Hortaz, akats hori konpontzera jo beharko du itzultzaileari lagundu nahi dion hiztegi-sistema konputazionalak.

Lokuzioak

Lokuzioen, esamolde idiomatikoen eta, oro har, hitz anitzeko terminoen itzulpenak buruhauste bat baino gehiago sortarazten du. Bi eratara egiten zaio aurre horrelakoen itzulpenari: horiek ere hitzez hitz itzultzen saiatuz edo termino osatua hiztegiren batean bilatuz. Dударik ez da bigarren hau dela jokabide zuzena, lokuzioen esanahia ez baita osagaiena konposatuz lortzen. Arazoa da, ordea, hizkuntza behar bezala ezagutzen ez duenak nekez igarriko diola lokuzioen presentziari. Gainera, hori igarri eta lokuzioa oso-osoan hartuta ere, hiztegi-kontsultan arazoak suertatzen dira sarri. Batetik, hiztegi arruntek ez dietelako leku askorik eskaintzen eta, bestetik, ez dagoelako hiztegi-sarrera horiek kokatzeko irizpide finkorik. Lexikografoak ez datoz bat kokatze-irizpideak finkatzean eta erabiltzaileek ezin asma zein estrategia jarraitu behar den lokuzio-sarrerak hiztegietan aurkitzeko. Egin ditugun itzul-saioetan Koldo Izagirreraren *Lokuzioen Hiztegia* izan dugu eskuartean eta garbiro

²⁸ Egia da lexikografoek ahaleginak eta bi egiten dituztela definizioak erraztearren, baina hala ere zailtasunak nabarmenak dira. Definizioen metalengoia eta idazkera estandarizatu ezean, hizkuntzaren gutxienezko ezaguera behar du hiztegi elebakarren erabiltzaile izan nahi duenak. Edo, bestela, ANHITZ bezalako hiztegi-sistemaz baliatu beharko luke hiztegi elebakar arrotzetan eroso antza mugitzeko. Baina horretaz § II.4.3n luze eta zabal ariko gara.

²⁹ Ez pentsa ordena sekuentzial hertsia jarraitzen denik, itzulpen automatikoan ohikoa den bezala. Bi prozesu kognitibo horien bereizketa, giza itzulpenaz ari garela, askozaz ere lausoago eta nahasiagoa da.

ikusitu dugu zaila dela hiztegi hori behar bezala erabiltzea³⁰. Adibidez, gure azterketetan erabilitako testu batean "*eltzetik babak atera*" zegoen eta *Lokuzioen Hiztegia* "*babak eltzetik atera*" da sarrera. Nahiz-eta bilaketa ondo zuzendu "*atera*" hori ardatz hartuta, kasualitatez aurkitu zuen zenbaitek sarrera hori, ordena-aldaketa dela eta.

Testu-hiztegi dikotomia

Atzipen-arazoez ari garela, esan dezagun bidenabar ez dela beharrezkoa lokuzioetara jotzea arazo horiekin topo egiteko. Evelyn Mitchell-ek, adibidez, ederki nabarmendu zituen 1983. urtean plazaratutako lan batean hiztegi-sarrerak eta adierak kausitzeko ikasleek dituzten zailtasunak. Guk ere ondorio beretsuak atera ditugu, gure esanetara aritu diren itzultzaileek ere behaztopa bihurriak aurkitu baitituzte sarrerak aurkitzean eta adierak aukeratzean. Bi zailtasun-mota gertatzen dira gure ustetan: batzuek zerikusitu zuzena dute morfologiarekin eta besteei zailtasun semantiko-pragmatiko dei genietzaieke. Nolanahi ere, denek dute ezaugarri komuna: testutik hiztegiara edota hiztegitik testura pasatzean gertatzen dira. Testu-hiztegi dikotomia ageri zaigu zailtasun askoren iturburu, beraz.

Morfologia

Atzipen-arazo morfologikoak testuko unitatea benetan forma lexikala delako gertatzen dira. Ez dago zuzeneko korrespondentziarik formen eta sarreraren artean, sarrerek ez baitute flexiorik, ez deklinabidezkorik, ez aditz-jokorik. Testuko unitate lexikalen eta hiztegiko sarreraren arteko zubi hau ez da iraganerraza morfologiaz ondo jantzita ez dagoen erabiltzailearentzat. Gure itzul-saioetan zenbait itzultzailek zailtasun handiak izan zituen "*sauraient*" edo "*attendis*" bezalako unitateak itzultzean. Eta oztopo horiek larriagotu egiten dira euskara bezalako hizkuntza eranskaria badarabilgu. Orduantxe bai behar duela gutxienezko ezagutza-maila hiztegi-erabiltzaileak, bestela jai.

Semantika-pragmatika

Behaztopa semantiko-pragmatikoak aipatu ditugu oraindik orain, morfologikoekin batera, erabiltzaileari kezka bat baino gehiago sortarazten diotelakoan. Semantiko-pragmatiko deitze hori unitate lexikalek testutik jasotzen dituzten ezaugarriei zor zaie. Ez da aherentzia behar testuinguruean dagoen edozein unitatek baduela berezko esanahiaz aparte testuak eranstean diona³¹. Esanahi erantsi horrek berebiziko garrantzia du itzulpenean, eta

³⁰ Ados gaude (Ibarzabal & Navarro, 85)ean hiztegi horren erabilgarritasunaz jaulkitzen diren iruzkinekin.

³¹ Denotazio eta konnotazio kontzeptuak erabili ohi dira antzeko bereizketak egiteko, baina zentzu zabalagoan hartu nahi dugu guk testuinguruaren eragina.

hiztegiek zeharka baino ez dituzte ukitzen sarreren testuingurupeko erabilerak. Erabiltzailearen esku geratzen da testuko unitate lexikaletik, esanahiez ondotxo hornituta dagoen horretatik, sarreren eta azpisarreren artean aukeratzea, eta gero, sorkuntza-lanetan murgiltzean, hautatu den hiztegiko unitatetik xede-forma lortzea jatorrizko esanahiei eutsiz. Nahiko lukete itzultzaileek —nahi zuten gurekin ihardun zutenek— horretan laguntza gehiago jasotzea hiztegietatik. Baina, nola ordea? Gure iritzirako, hiztegietakoa informazioa bestelako tresnez osatu behar da. Testu-corpusak, adibide-bankuak, hipertestuak, horiek denak zeresan handia izango dute testuinguruak eta hiztegi-sarrerak elkarrenganatzerakoan. Krista Varantola ere uste berekoa dugu (1994:606)n:

text corpora can be used to bridge the gap between the general-purpose, context-free dictionary information and the highly context-sensitive needs of the user.

Aurrekoa gutxi balitz, testuaren ingurutik harago joanda, esan gabe doa baldintzapen kulturalak, erabilera estereotipatuak etab.ak ahaztezinak direla testuetatik hiztegietara iragatean. Eta pasabide hori bai dela zaila.

Hizkuntzen hurkotasuna

Hizkuntzen arteko hurbiltasunak duen eragina ere ez zaigu oharkabea pasa. Hurbiltasun linguistikoak —alfabeto bera erabiltzea, egitura morfosintaktiko antzekoak edukitzea, familia berekoak izatea— erraztu egin ditzake inguru elebidunetako hiztegi-kontsultak: hala begitandu zaigu, errazagoa —erraxkeri puntu bat ere balegoke— dela frantses-gaztelania hiztegiak maneiatzea euskara tartean denean baino³². Baina ezaugarri linguistikoez haratago, bada bestelako hurbiltasuna, kultural dei genezakeena, eta horrek ere erruz eragiten dio itzulpenari eta hiztegi-erabilerari. Josh Ard-ek horrela zioen ingelesa ikasten ari ziren estudiante japoniar eta espainolen hiztegi-erabilera aztertu zuenean (1982:2)n:

the success rate depends on the native language background of the students. Students from languages 'close' to English ... are more likely to be successful.

Geroztik, Gillian Moss (1992) edo Hilari Nesi (1994) bezalako autoreek hurbiltasun hori dela eta gertatzen diren erabilera okerrak aipatu dituzte. Autore hauen arabera, badira idazkera oso hurbileko terminoak esanahi desberdinak dituztenak, "lagun gaiztoak" omen, eta kontuz ibili ezean "lagun" horiek erabiltzera jotzen omen du itzultzaile fidakorrak. Ez da hau, guk uste, fenomeno oso zabaldua eta kezkatzekoa, baina ez legoke gaizki sasibaliokidetasun horien berri ematea hiztegietan. G. Moss-ek espainiera eta ingelesaren arteko adibide batzuk jartzen ditu: *exit-éxito, success-suceso, remove-remover*, etab. Baina badira, hurbilago,

³² Hizkuntz ezagutzan leudekeen aldeak aparte utzita.

euskara-espainiera adibideak, *okasio-ocasión, sesio-sesión* edo *arte-arte*, kasu baterako; eta euskara-frantsesa adibideak ere: *bide-vide*.

Kontsiderazio horiek gogoan ekin diogu itzulpen-prozesuaren modelizazioari, horren ondorioz ateratako ereduak ere gogoeta horiek jaso ditzan.

Modelizaziorantz: hurbilpen informala

Ez da samurra maila lexikaleko itzulpenaren nondik norakoak eredu batean jasotzea. Nork bere bidea jarraitzen du, hori kasurik onenean. Normalagoa da ezein eredurik ez jarraitzea, soilik ahal dena edo bururatzen dena egitea: itzulinguruka itzultzea. Hori, ikusi eta hainbestez, hala da.

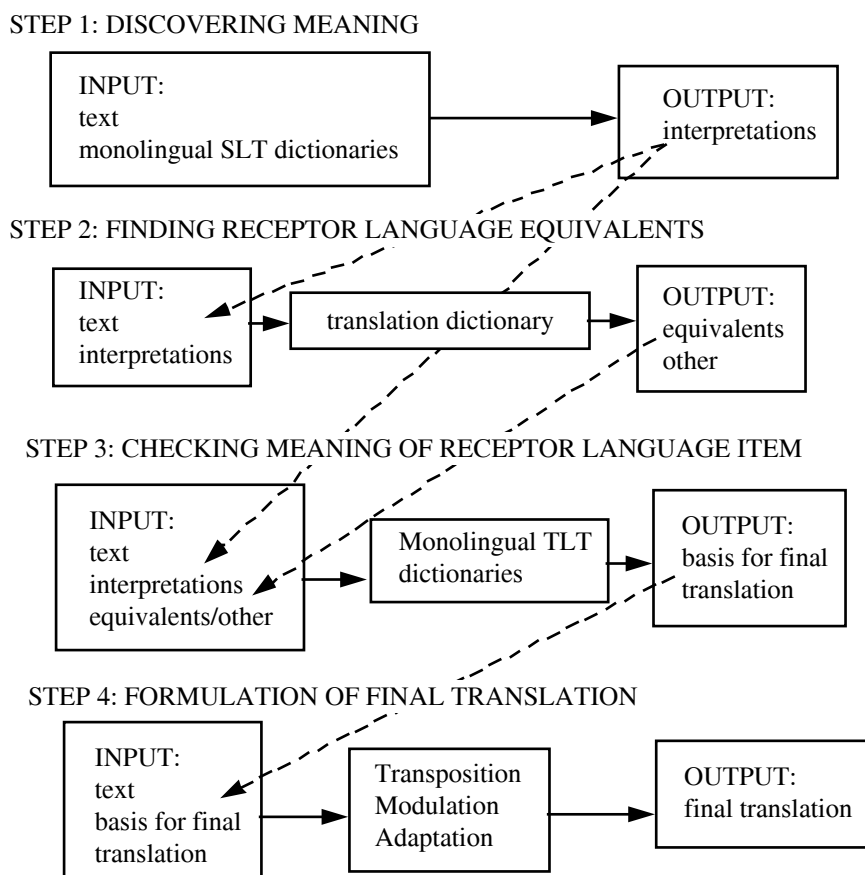
Baina, metodorik gabeko iharduera itxuraz anarkiko horrek ere baditu ezaugarri bereak. Sandiak ez du hiztegi elebidunik erabiliko, Urliak nahiago izango du informazio gramatikala beste ezer baino eta Berendiak erabilera-erregistroen premia izango du sarri; hala ere, batzuek eta besteek prozesu kognitibo berdintsuak darabiltzate. Abigail Neubach eta Andrew D. Cohen-en (1988:2) esana da honakoa:

when learners ... consider turning to the dictionary, they go through what amounts to a series of complex cognitive processes, where at every stage they have to make important decisions.

Hots, prozesu batean zehar ikusten da hiztegi-erabilpena, ez inola ere ekintza puntual isolatu gisa. Prozesu hori unitate lexikalen itzulpena denean, badira ataza eta azpiataza saihestezinak, nola edo hala aurrera eraman beharrekoak, ez bada behintzat narraskeriatan erori nahi. Batzuk erraz aritzen dira, itzulpena berehala asmatuko dute, ikusi batean irudi lezake unitate lexikala itzultzea horixe bera dela, unitatea, urrats bakarreko betebeharra, banaezina eta sinplea. Baina beste batzuk astiroago doaz, zailtasunak topatu, hiztegira jo, hiztegian ere zailtasunak aurkitu, horrela, behin eta berriz, urrats askoren buruan agian helburua lortzeko. Batzuek buruz egiten dutena, besteek agerian egiten dute, horixe da alderik handiena. Aurreraxeago deskribatuko dugun ataza-eredu edo egiturak ondo jaso nahi ditu iharduera horiek denak.

Izan dira gu baino lehenagotik itzulpen-prozesua modelizatzeko saioak egin dituztenak. Motibazioetan eta metodologian gurearekin antza duelako, Maastricht-eko *School of Translation and Interpreting*-en garatutako ikerkuntz proiektu baten berri emango dugu hemen. Bi helburu ditu aipatu proiektuak: lehen, itzulpengintza ikasten ari direnei hiztegia egokiro erabiltzen irakastea, eta bigarrena, itzulpenean lagungarriak izateko hiztegiak nola

antolatu beharko liratekeen argitzea. Itzulpen-saioetan 34 ikaslek hartu zuten parte, eta II.1 irudian ikus daitekeen itzulpen-ereduari atxeki zitzaizkion beren lanean. Hara, zehatzago, zer egin behar zuten ikasleek urratsez urrats:



II.1 irudia.- Itzulpen-prozesuaren errepresentazioa (Starren P. & Thelen M., 1988:449)

1. urratsa (step 1): Sarrera gisa testu bat eta hiru ingelesezko hiztegi elebakar eman zitzaizkien³³, horiek emandakotik eta testuingurutik nederlanderazko interpretazioak formulatzeko.
- 2. urratsa (step 2): Aurreko interpretazioa eta testua abiapuntutzat hartuta ekin zioten urrats honi. Honakoan hiztegi elebiduna erabili behar zuten³⁴, aurreko intepretazioa eta hiztegiak emandako baliokideak bat zetozen ala ez ikusteko. Bat etorri ala ez, elebidunaren baliokide zerrenda hartzen zen urrats honetako irteeratzat.

³³ Erabilitako hiztegiak: *Oxford Advanced Learner's Dictionary of Current English*, *Longman Dictionary of Contemporary English* eta *Collins Dictionary of the English Language*.

³⁴ Hauxe: *Van Dale Groot Woordenboek Engels-Nederlands*.

- 3. urratsa (step 3): Aurreko bi urratsetako emaitzak nederlanderazko hiztegi elebakarrean³⁵ txekeatu eta azken itzulpenerako oinarria finkatzea zuen xede urrats honek.
- 4. urratsa (step 4): Transposizioa, modulazioa eta adaptazioa bezalako itzulpen-prozedurak aplikatzen ziren honetan, azken itzulpenak xede-hizkuntzaren arau gramatikalak eta konbentzio kulturalak jaso zitzaizkion.

Bada desberdintasun garrantzizko bat lan honen eta gurearen artean: ikasleak soka motzean eduki eta urrats horiek ematera behartu dira honakoan, aldiz gurean, laxoan utzi dira. Baina, azkenean, hauek oinarritzat hartutako ereduak eta guk deduzitutakoak antzekotasun handia dute elkarren artean, geroxeago § II.4.3n ikusiko dugun bezala.

II.4 Itzulpen-prozesuaren modelizazioa.

Itzulpen-ereduaren berri eman aurretik modelizaziorako erabili ditugun metodologiaz eta lengoaiar arituko gara, II.4.3ko deskribapena ulergarria(go) izan dadin. Ondorengo bi ataletan, hortaz, ezagutzaren modelizaziora egingo dugu jauzi, begietaratu —eta begitandu— zaizkigun prozesu eta egitateak nola formalizatuko ditugun argitzeko.

II.4.1 Modelizazio-metodologiaz.

Hiztegiz lagundutako hitzen itzulpenaz eta prozesu horretan antzeman ditugun azpiprozesuei buruz ikasi duguna eredu formal batean jaso nahi dugu, horra helburua. Eginkizun bati buruzko ezagutzaren deskribapen egituratuari *eredu kontzeptual* deitu ohi zaio ezagutzaren injinerutzan. (Fensel and van Harmelen, 1994:105)en *eredu kontzeptualaren* definizio argigarria ematen da:

conceptual model: an abstract, implementation independent description of the relevant problem solving expertise. A conceptual model should describe the task which is solved by the system and the knowledge which is required by it.

Software-injinerutzako "espezifikazio funtzionala"ren senide bikia da ezagutzaren injinerutzan erabiltzen den "eredu kontzeptuala". Egun, software-injinerutzan, ez dira zalantzan jartzen espezifikazioa eta sistemen diseinua nahiz inplementazioa bereiz tratatzeak ekarri dituen onurak. Halatsu gertatzen ari da azken aldion ezagutzaren munduan ere, eta horregatik Ezagutzaren Oinarritutako Sistemen garapenerako metodologia askok bere egin dute

³⁵ Van Dale Groot Woordenboek Der Nederlandste Taal hiztegia.

espezifikazio/inplementazio bereizketa —prototipatze azkarraren kaltetan—. Horien arteko batzuen berri —aipamen soila baino amiñi bat gehiago— emango dugu segidan.

Prozesuen modelizaziorako hurbilpenak

Problemen ebazpenerako prozesuen modelizazio-ereduak atazen³⁶ deskribapenean eta ataza horiek ebazteko behar den ezagutzaren zehaztapenean oinarritzen dira. Horiek guztiak antzeko objektu-motak maneiatzen dituzte: atazak, azpiatazak, helburu eta azpihelburuak, ebazpen-estrategiak, atazen arteko erlazioak eta planak, egoerak, murriztapenak, e.a. Baina, izan, badira aldeak batzuetatik besteetara.

(Johnson and Johnson, 1991)n KAT, *Knowledge Analysis of Tasks*, metodologia deskribatzen da, zeina, egileen esanetan, "is concerned with analysing and modelling the knowledge people possess and utilize in carrying out tasks". Hurbilpen horretan atazen egitura taxonomikoak eratzen dira, atazen arteko ordenazioak planetan adierazten dira eta ataza bakoitzari Ezagutza-Egitura bat egokitzen zaio, non, besteak beste, estrategiak —jakintza prozedurala—, objektuak eta oinarritzko ekintzak biltzen diren.

(Yen and Lee, 1993) eta (Lee and Yen, 1993)n azaldutako *Task-Based Specification* metodologian beste modu batera ulertzen dituzte atazen inguruko kontzeptuak. Atazak berak, kontzeptu abstraktu gisa, esanahi desberdina hartzen du, horra: "*we treat a task as a state transition and specify the states before, during, and after the operation using precondition, protection and postcondition which are partial descriptions of the state*". Egoeratan oinarritutako eredua darabilte bai domeinuaren espezifikazioko atal batean, bai espezifikazio funtzionalean, baita portaeraren espezifikazioan ere.

(Karbach, Linster, Voß, 1990)en lau hurbilpen aurkeztu eta aztertzen dira. Ondorengo hauexek dira lan horretan jasotakoak:

Generic Task deitutakoak "elkarrekintza problema" hartzen du gunetzat, hau da, ezagutza ezin daitekeela erabilera kontuan izan gabe errepresentatu eta, hortaz, problemen ebazpenerako egituratan, ataza generikotan, kodetu behar dela. Ataza, zentzu zabalean hartuta, problemen ebazpenerako metodoaren aplikazioa —domeinutik independentea— da. Arrazonatze-estrategiak, bloketan egituratu eta generikotzat hartuz gero, ataza konplexuen osagai gisa berrerabil daitezke (Chandrasekaran & Johnson, 93).

Method-to-Task hurbilpena Ezagutzan Oinarritutako Sistemen eraikuntza erraztu beharrak eraginda mamitu zen, zehatzago esanda, adituaren eta ezagutza-injineruaren

³⁶ Testuinguru honetan atazak esanahi hau hartzen du: problema-ebazpenerako metodo baten aplikazioa, non ebazkizun dagoen problema domeinu jakin batean kokatzen den.

ingurumariak bereizteko komenientziak motibatuta. Bereizketa hori problema-ebazpenerako ezagutzen eta domeinu-ezagutzen zertzen da, eta bi ezagutza-mota ondo bereizi horiek, era berean, izaera desberdineko eredu kontzeptualetan³⁷ sailkatzen dira. Honakoan ere, izaera beretsuko atazak, nahiz eta domeinu desberdinetan aplikatu, eredu generikotan bildu nahi dira, gero egokiera denean instantziatzeko. Hartara, problema-ebazpenerako atala berrerabili egin daiteke domeinu parekideetan.

Role-Limiting-Method izeneko hurbilpenean berriz, ekintza-sekuentzia baten identifikazioan, hautaketan eta inplementazioan datza problemen ebazpena. Horretarako metodo ahulak —ataza-mota askori aplikatzeko modukoak— eta rol-mugatzaileak —aurrekoen espezializazioak— erabiltzen dira. Rol-mugatzaile diren metodoek, jakintzaren eskuratzea eta kodetzea gidatzeaz gain, atazari lotutako kontrol-ezagutza aurredefinitzen dute.

Hiru aurrekoekin batera, KADS hurbilpenari egiten zaio tartea azterketa konparatibo horretan. KADSen eredu kontzeptuala, § Ien azaldu den bezala, lau geruza desberdinetan adierazten da: domeinu-geruza, inferentzi geruza, ataza-geruza eta estrategi geruza. Eredu kontzeptualarekin batera, diseinu-eredua —eredu kontzeptual baten inplementaziorako transformazioa— har daiteke KADS hurbilpenaren bereizgarri nagusitzat.

Gure hautaketa

KADS metodologian aurkitu dugu hurbilpenik egokiena. Horrixe ikusi dizkiogu modelizaziorako oinarri argienak, bai osagaiak bereiztean, bai ezagutza sailkatzean, eta malguenak, nork bere neurriko eredu kontzeptualak deskribatzeko aski zabalak. Oinarri horiek gida eta molde izan dira guretzat, terminologia-iturri eta, kontzeptuen nahas-mahasean, argigarri.

KADSen eraginpean, eredu kontzeptuala oso-osorik ez baina, ataza-geruza eta, zeharka, geruza horrek inferentzi nahiz estrategi geruzekin dituen harremanak espezifikatuko dira gauden kapitulu honetan. ANHITZ eredu kontzeptualaren baitan, propio itzulpen-prozesuarekin eta hiztegi-erabilpenarekin zerikusia daukan ezagutza jasotzen da geruza horietan. Hala bada, segi dezagun nola —zein pausotan eta zein lengoaiatan, esan nahi baita— deskribatuko dugun azaltzen.

³⁷ Arestian eman dugun Fensel eta van Harmelen-en eredu kontzeptualaren definizioak baino esanahi murriztagoa du Method-to-Task hurbilpenean hartu den horrek.

Modelizazio-urratsak

Modelizazioan erabili den teknika (Akkermans *et al.*, 1993b)n jasotzen denaren antzekoa da. Horrela izenda eta defini daitezke emandako hiru urratsak:

- Kontzeptualizazioa: ataza nagusiaren ezaugarri eta nondik norakoak ezagututa, kontzeptualizazioak agerian uzten ditu zein diren eta nola izendatzen diren azpiatazak, objektuak, helburuak eta egoerak. Nolabait, atazaren espazioa mugatu eta definitzen da.
- Espezifikazioa: hots, maneiatzen diren kontzeptuen zehaztapen formala. Bereziki atazak, egoerak eta beraien arteko erlazioak dira espezifikazioan lantzen direnak.
- Operazionalizazioa: atazen ebazpenerako erabiltzen diren inferentzi eta kontrol-egiturak jorratzen dira honakoan. Deskribatzen diren egiturak exekutagarriak izaten dira hein batean.

§ II.4.3n, ANHITZ eredu kontzeptualaren deskribapenean, azalpen-bide horri lotuko gatzazkio.

II.4.2 Modelizazio-lengoaiaz.

Atazak, egoerak, erlazioak, kontrol-egiturak nola adierazi, galdegai horrek, hots, eredu kontzeptualaren adierazpideak, garrantzi handia du arestiko modelizazio-metodologietan. Are, zalantza handirik gabe esan daiteke adierazpide hauek, lengoaia hauek, ezaugarritzen dituztela hein handi batean modelizaziorako hurbilpenak.

Adierazpideari buruzko hasierako galdera zabal eta orokorra laster fintzen da, eta ugaldur, lengoaien dohain eta mugak aztergai hartuz gero. Ondoko galderak, konparazio batera, berehalakoan etorriko zaizkio burura zeregin horretan murgiltzen denari:

- Zein dira maneiatzen diren kontzeptuak eta beren esanahi zehatza?
- Nola adierazten dira ataza-egiturak eta atazen deskonposizioak?
- Nola zehazten dira prozesuan zehar gertatzen diren egoerak?
- Zein dira kontrol ezagutza errepresentatzeko erabiltzen diren primitiboak?
- Zein eratako baldintzazko adierazpenak erabil daitezke?
- Atazei buruzko ezagutzaren formulazioa domeinutik independentea al da?
- Adieraz al liteke ez-determinismoa?

Irizpide horiek gogoan gainbegiratu ditugu KADS hurbilpenean erabili izan diren lengoaia batzuk, gure komenientziei ondo egokituko zaiena zein, edo nolakoa, izango den argitzeko.

KADS ereduako modelizazio-lengoaiak

Esana dugu KADS ereduak malguta dena, eta malgutasun horren kariaz, hain zuzen, ez da batere harritzekoa formalizazio-lengoaia bat baino gehiago erabili izana eredu horretan. Horietako batzuen berri ematen da (Fensel & van Harmelen, 1994)³⁸, eta batez ere hortik jasotako informazioaz hornitu dugu ondoko zerrenda:

- *Operational Models of Problem-Solving* (OMOS) (Linster, 1993): "Role-Limiting-Method" eta "Method-to-Task" hurbilpenak KADS ingurunean egokitu nahiak eragin zuen operazionalizazio-lengoaia honen sorrera. OMOSen inferentzi mailako ekintzak ataza mailako ekintza primitiboak dira, eta ekintza primitibo hauek konposatzeko sekuentzia-, baldintza- edo iterazio-egiturak erabil daitezke.
- MODEL-K (Voß & Karbach, 93; Karbach *et al.*, 94): bi azpilengoaia bereizi behar dira: espezifikazio-lengoaia, atazak eta beren finte hierarkikoa deskribatzeko, eta operazionalizazio-lengoaia, kontrol-ezagutza adieraziz ataza bakoitzari gorputza eranstea.
- MoMo: aurreko bien ondorengotzat har daiteke. Batetik, OMOSen adierazgarritasun-mugak zabaltu, eta bestetik, MODEL-Kren ahalmen konputazionala mugatu nahi du. Inferentzi geruzan onartzen den ez-determinismoari kontrol osagarria ezartzen zaio ataza-geruzan, eragiketa primitiboak —inferentzi mailako ekintzak— sekuentzia-, baldintza- eta iterazio-eragiketez konbinatuz.
- FORKADS (Wetter, 1990): KADS ereduaren oinarri formalak finkatzen ditu, batez ere lehen mailako logika-lengoaia baliatuz. Bereizketa argia egiten da ataza-buru eta ataza-gorputzaren artean. Atazen eta azpiatazen gorputzak beste atazei edo inferentziei —oinarrizko atazak— egindako deien sekuentziak dira eta, aurrekoetan bezala, iterazioak edo baldintzazko sententziak ere erabiltzen dira. Horrez gain, badira beste bi sententzia axioma-aldaketak eragiteko.
- *Knowledge Acquisition and Representation Language* (KARL) (Fensel, 1991): lengoaia honek atazari buruzko formalizazioa eta operazionalizazioa biltzen ditu: prozesuaren deskribapen formalari deskribapen operazionala egokitzen zaio automatikoki.
- (ML)² (Akkermans *et al.*, 1993a; van Harmelen & Balder, 1993): nahiz eta lengoaiaren azpimultzo bat operazionaliza daitekeen, batez ere formalizazio-lengoaiatzat hartzen da. Inferentzi ekintzak inplikazioen bidez deskribatzen dira, non sarrera/irteera erlazioak xehetzen diren. Ataza-geruzan lau eragiketa-mota dira erabilgarriak inferentzi ekintza bakoitzeko: Sarrera/Irteera bikote bat (S/I bikotea) konputatuta dagoen txeketa, konputatu gabeko S/I bikoterik oraindik badagoen txeketa, S/I bikote bat konputatzea (historia-aldagai batean gordez) eta S/I erlazio

³⁸ Benetan gomendatzen diogu artikulua honen irakurketa gai honetan gehiago sakondu nahi duenari.

jakin bat betetzen den txekateza. Primitibo hauek sekuentziazioa, hautaketa ez-determinista eta iterazio ez-deterministaren bidez konbina daitezke.

- QIL eta K_{BS}SF lengoaiak, aurrekoak ez bezala, jakintzaren errepresentazio orokorrerako pentsatuta daude, bereziki KADS ereduak errepresentatzeko baino. Hala ere, biak erabili izan dira KADS ereduak deskribatzeko.

QILEk ez du ataza-deskonposizio finkorik erabiltzen, horren ordez, azpiataza sekuentziak sortzen ditu dinamikoki. Azpiataza-sekuentzien planifikazioa estrategi geruzan lantzen da KADSen.

K_{BS}SF VITAL proiektuan erabilitako modelizazio-lengoaia da —VITAL lengoaia deritzo testuinguru horretan— (Jonker & Spee, 92). Lengoaia honek ez du ataza- eta inferentzi geruzen arteko bereizketarik egiten. Ataza guztiak, primitiboak ala konposatuak izan, Bmoduluen ("Behaviour module") bidez modelizatzen dira. Bmodulu horiek, batetik, atazaren sarrera/irteera espezifikatzen dute eta, bestetik, atazaren portaera deskribatzen dute lengoaia prozedural batez.

Conceptual Modelling Language: gure aukera

Conceptual Modelling Language (CML) CommonKADS ereduen espezifikaziorako definitu den lengoaia egituratua eta erdi-formala³⁹ da. (Schreiber *et al.*, 1994)n deskribatzen den CMLk domeinu-ezagutza, inferentzi ezagutza, ataza-ezagutza eta problemen ebazpen-metodoak tratatzen ditu. Artikulu horretan CMLren eraikitzaileak aurkezten dira ezagutzamota bakoitza espezifikatzeko eta eraikitzaile horiek BNF notazioaz xehetzen dira eranskinean.

Arestiko kontuak errepikatuz, gauden kapitulu honetan ataza-ezagutza landuko dugu batez ere. Hortaz, murriz dezagun gure azalpena horretara, gero ere izango da-eta beste ezagutzei buruzkoak jorratzeko abagunerik.

CMLn atazei buruzkoa bi ezagutza-motatan egituratzen da: ataza-ezagutza eta problema-ebazpenerako ezagutza. Ataza-ezagutza kontrol-informazioa, atazaren burubideei dagokiena, finkoa da, aldaezina, zenbaitetan zurrunegia. Hain zuzen, hori ebitatu eta kontrol-forma malguagoei lekua egiteko jasotzen da CMLn problema-ebazpenerako ezagutza, zeinek bi alde dituen orobat: ebazpenerako metodoak eta ezagutza estrategikoa. Metodoek atazaren helburua nola lortu deskribatzen dute, eta ezagutza estrategikoak metodoak nola aukeratu eta zein ordenatan aplikatu behar diren zehazten du, horretara ataza-eredua dinamikoki sortuz.

³⁹ Erdi-formal izate horrek ez du esan nahi errepresentazio formalak onartzen ez dituenik, baizik eta ez duela errepresentazio-formalismo jakinik eskatzen. CMLk notazio testual egituratua eta notazio diagramatikoa biltzen ditu.

Ataza- eta ebazpen-ezagutza batera bildu ditugu gurean. Ez diogu iritzi beharrezko bereizketa hori mantentzeko gure eremuan, batetik, deskribapenaren argitasun-galerarik ez dagoelako⁴⁰ eta, bestetik, kontrol-egituraren malgutasuna atazari itsatsitako kontrolean lor daitekeelako. CMLren egokitze hau azalekoa da funtsezkoa baino, esan nahi baita, ez diogula metodoen berri emateari utziko, aiantziko ez dugun bezala ezagutza estrategikoaren deskribapena. Aurrerantzean, ataza-ezagutza deituko diogu bat egitearen ondorioz eratu dugun ezagutza-motari.

ANHITZ ereduko atazen deskribapenerako lengoaiak, CML egokituak, ondoko klausula-egitura darabil:

```
task-knowledge41
```

```
  task:
```

```
    task-specification
```

```
      task-definition
```

```
        goal:
```

```
        input:
```

```
        output:
```

```
      task-body
```

```
        type:
```

```
        parents:
```

```
        sub-tasks:
```

```
        additional-roles:
```

```
        acceptance-conditions:
```

```
        competence-conditions:
```

```
        control-structure:
```

```
    task-structure
```

```
      decomposition-structure:
```

```
      data-flow:
```

Klausula horiei buruzko argibideak emango ditugu segidan banan-banan:

- task-knowledge (ataza-ezagutza): Helburu jakin bat atazaren bidez nola irits daitekeen deskribatzen du.
- task (ataza): Atazaren izena.
- task-specification (ataza-espezifikazioa): Atazaren espezifikazioa nabarmentzen da atal honetan.

⁴⁰ Alderantziz baizik, erredundantziak —ataza-ezagutzako hainbat ezaugarri problemen ebazpen-metodoetan errepikatzeak— deskribapena ilundu egin lezake.

⁴¹ Tabuladore bidez adierazi dira egitura honetan klausula-azpiklausula erlazioak. Hala, goal azpi-klausula task-definition orokorragoari dagokio, eta sub-tasks, berriz, task-bodyri.

- **task-structure** (ataza-egitura): Klausula honetan ez da espezifikazioan esaten ez denik azalduko, baina notazio grafikoa erabiliko da atazaren deskribapena, gaingiroki bada ere, errazago ikus dadin. Klausula lagungarria da.
- **task-definition** (ataza-definizioa): Zer lortu behar den adierazten da, espezifikazio deklaratio baten bitartez.
- **task-body** (ataza-gorputza): Helburua nola lortzen den xehetzen da.
- **goal** (helburua): Atazaren aplikazioaren bidez lor daitekeen helburuaren deskribapen testuala.
- **input/output** (sarrera/irteera): Atazak datu gisa hartu eta emaitza gisa lortzen dituen rolen⁴² definizioa. Definizioak bi alde ditu: batetik izena, eta bestetik deskribapen testuala.
- **type** (mota): Hiru ataza-mota bereiziko ditugu:
 - **Ataza konposatuak (composite):** Azpiatazatan deskonposatzen diren atazak.
 - **Ataza primitiboak (primitive):** Inferentziekin (inferentzi geruzarekin) zuzenki lotuta daudenak.
 - **Transferentzi atazak (transfer):** Erabiltzailearen eta sistemaren arteko elkarrekintzako atazak. Ataza hauek, berez, ez dira ezagutza-ereduko parte, baina ANHITZ ereduari sartu egingo ditugu, zenbaitetan argigarri dira-eta.
- **parents/sub-tasks** (gurasoak/azpiatazak): Ataza-deskonposizioaren berri ematen da. Ataza gurasoa azpiataza umetan deskonposatzen da.
- **additional-roles** (rol gehigarriak): Atazaren egitean parte hartzen duten rolak zerrendatzen dira. Sarrera/Irteerako rolak eta rol gehigarri hauek izango dira atazak maneiatuko dituen objektuak.
- **acceptance-conditions** (onartze-baldintzak): Ataza abian jarri dadin gertatu behar duten baldintzak zehazten dira. Baldintza hauek predikatu logikotan zertzen dira, rolen gaineko predikatutan nagusiki. Itxuraz deskribatzaileak, operazionalizazioarako behar-beharrezkoak dira eta zeregin funtsezkoa dute atazen planifikazioan. Onartze-baldintzetan oinarrituta erabakitzen da zein diren ataza abiaezinak eta zein den abiagarrien artean egokiena. CML lengoia onartze-baldintzei buruzko klausula problemen ebazpen-metodoetan azaltzen da⁴³.

⁴² CommonKADS metodologian ezagutza-rolak objektuei emandako izen orokor eta funtzionalak dira, objektu horiek daukaten erabilpenari atxekitakoa. Ez du zertan gertatu behar rolak eta domeinuko objektuen motak bat etortzea, gerta bailiteke objektu-mota batek baino gehiagok rol bera edukitzea atazaren batean.

⁴³ Klausula honetako baldintzen sintaxia eta erabilera § IV.ean azalduko da, harako utzi baititugu operazionalizazioarekin zerikusi zuzena duten ezaugarriak. Kapitulu honetan baldintzen deskribapen abstraktua ematera mugatuko gara soilik.

- **competence-conditions** (konpetentzi baldintzak): Ataza burututakoan geratzen den egoera deskribatzen da. Egoera horrek atazaren helburuarekin bat etorri behar du⁴⁴.
- **control-structure** (kontrol-egitura): Atazaren helburua lortzeko azpiatazak nola antolatu behar diren finkatzen da. Azpiatazen gaineko kontrola lengoaia prozedural batez deskribatzen da⁴⁵, sekuentziazioa, hautaketa determinista eta iterazio determinista bezalako kontrol egiturak baliatuz. Esan gabe doa klausula honek operazionalizazioa duela xede, horregatik ANHITZek ulertu eta exekutatu dituen baldintzak eta aginduak erabili behar dira, hots, zehazki definitutako lengoaia exekutagarria⁴⁶.
- **decomposition-structure** (deskonposizio-egitura): Ataza-hierarkiaren berri ematen da, era grafikoan.
- **data-flow** (datu-fluxua): Atazatik atazarako datuen fluxu-diagrama jasotzen da klausula honetan.

II.4.3 ANHITZ ereduaren deskribapena.

ANHITZ ereduaren ezaugarriak deskribatzeko erabiliko dugun bidea ereduaren sorrerarako erabili dugunaren antzekoa da. Beraz, hiru estadio bereizi —kontzeptualizazioa, espezifikazioa, operazionalizazioa— eta ereduaren ezaugarriak horien arabera paratuko ditugu.

II.4.3.1 Kontzeptualizazioa.

Esku artean dabilkigun problemaren espazioa mugatzea du helburu kontzeptualizazioak. ANHITZ espazioko entitateak hiru multzotan sailkatuko ditugu: objektuak/rolak, azpiatazak eta egoerak.

⁴⁴ ANHITZ ereduaren ez bezala, CML lengoian bi maila bereizten dira: helburua atazari dihoakio, konpetentziak ebazpen-metodoei. Hori horrela, ataza aurrera eramateko ebazpen-metodo egokiena aukeratzea, konpetentziak eta helburua parekatzean datza, beti ere prozesuaren egoera eta onartze-baldintzak bat datozela.

⁴⁵ ANHITZ ereduaren lengoaia prozedurala aukeratu badugu ere, CMLk ezagutza-injineruaren esku uzten du formalismo egokiaren aukeraketa.

⁴⁶ Lengoaia horri buruzko argibide gehiago IV. kapituluaren emango da.

ANHITZ Objektuak/Rolak

Unitate lexikalen itzultze-lanetan alde edo moldez maneiatzaren diren objektuak eta objektu horiek jokatzaren dituzten rolak xedatu nahi izan dira. ANHITZ espazioaren baitan bost rol-mota bereizi ditugu:

- Ingurune-rola: rol hau jokatzaren duten objektuek itzulpenaren nondik norakoak deskribatzaren dituzte.
- Erabilera-rola: unitate lexikala testuinguru jakin batean nola erabiltzaren den argitzekoak.
- Rol semantikoa: esanahia lantzeko erabiltzaren diren objektuek rol hau betetzaren dute.
- Rol morfosintaktikoa: portaera morfologiko nahiz sintaktikoa jorratzaren dituztenak mota honetakoak dira.
- Rol lexikala: lexikoaren mailako ezaugarriak lantzeraz zuzenduak.

Finduz gero, objektuen izaerari estuago atxekiz, bost mota nagusi horien pean beste hainbat rol ezagut daitezke. Rol xeheago horiek baliatuz, segidan ANHITZ objektuen eta rolen⁴⁷ zerrenda aurkeztuko dugu⁴⁸:

objektua

rola

testu_eredua [testu_mota] [testuingurua]	testuingurua:<er> sintagma:<ms>
itzultzaile_eredua [hizkuntza_ezaguera] [iaiotasun_maila]	itzultzailea:<in>
konnotazioak	konnotazioak:<er>
hiztegia	sorburu_hiztegia:<in> xede_hiztegia:<in> hiztegi_elebiduna:<in> hiztegi_hobetsia:<in>
hizkuntza	sorburu_hizkuntza:<in> xede_hizkuntza:<in> gogoko_hizkuntza:<in>
arloa	erabileremua:<er>
unitate_lexikala ⁴⁹	sorburu_unitatea:<le>

⁴⁷ Kontuan izan objektu batek rol bat baino gehiago joka dezakeela eta, bestalde, objektu batek baino gehiagok rol bera joka dezaketela.

⁴⁸ Aurkezpen-taulako rol guztiei etiketa bana jarri zaie zein motatakoak diren adierazteko. Etiketa hauek erabili dira: ingurune-rola: <in>, erabilera-rola: <er>, semantikoa: <se>, morfosintaktikoa: <ms> eta lexikala: <le>. Rol bakoitzari mota-etiketa bakarra itsatsi diogu, horrela ANHITZ espazioan rol horrek duen esanahia nabarmenago gera dadin.

	xede_unitatea:<le>
hiztegi_sarrera	sorburu_sarrera:<le> xede_sarrera:<le> sarrera_lagungarria:<le>
morfologia	sorburu_morfologia:<ms> xede_morfologia:<ms>
kontzeptua	sorburu_kontzeptua:<le> xede_kontzeptua:<le> kontzeptu_lagungarria:<le> atributua:<se> atributuaren_balioa:<se>
definizioa	sorburu_definizioa:<se> xede_definizioa:<se> erabiltzaile_definizioa:<se> tesauruserako_ideia:<se>
murritzapen_espresioa	tesauruserako_ideia:<se>
diferentzi_espresioa	kontzeptu_arteke_aldea:<se>
elkarbide_espresioa	kontzeptu_arteke_harremana:<se>
adibidea	erabilera_egiaztagarria:<er>
patroi_sintaktikoa	patroi_sintaktikoa:<ms>
lokailu_sintaktikoa	lokailua:<ms>
atributu_primitiboa	atributua:<se>
erlazio_sintaktikoa	erlazioa:<ms>
azpikategorizazioa	portaera_sintaktikoa:<ms>
funtzio_lexikala ⁵⁰	funtzio_lexikala:<ms> erlazioa:<ms>

ANHITZ atazak

Itzultzaileekin egindako azterketetan, hala zertan aritzen diren nola zein eginkizunetan murgiltzen diren, horrelakoak antzemangai izan ditugu. Kontzeptualizazioan eginkizun horiek identifikatu eta ataza-azpiataza gisa kokatu ditugu ANHITZ eremuan. Hiru motatan sailkatu ditugu atazak: batzuk konposatuak, besteak primitiboak eta hirugarren batzuk transferentzi atazak. Halaxe bada, atazak zerrendatzeko sailkapen horri —CML lengoaiak erakutsi bezala— lotuko gatzazkio:

Ataza konposatuak:

- Itzulpen_unitatea_itzultzea
- Sorburu_unitatearen_ulerkuntza
- Esanahia_eridetea

⁴⁹ Lan honetan, unitate lexikala azaleko gauzatze simple —hitz bakarra— edo konplexua —hitz bat baino gehiago— duen forma lexikal flexionatua da.

⁵⁰ Mel'cûk-ek (1982) definitutako funtzio lexikalak dira objektu horiek.

- Unitate_lexikaletik_hiztegi_sarrerara
- Sarrera_diskriminatzea
- Hiztegi_sarreratik_definizioa_lortzea
- Hiztegi_sarreratik_kontzeptura
- Adiera_diskriminatzea
- Kontzeptuaren_esanahia_lortzea
- Esanahia_diskriminatzea
- Esanahia_egiaztatzea
- Baliokidea_bilatzea
- Baliokidetza_hipotesiak_lortzea
- Hiztegi_sarreratik_baliokidea_lortzea
- Adieraren_baliokidea_lortzea
- Baliokidea_diskriminatzea
- Baliokidetza_egiaztapena
- Xede_unitatearen_sorkuntza
- Sorkuntza_hipotesiak_lortzea
- Sorkuntza_hipotesiak_diskriminatzea
- Sorkuntza_hipotesia_egiaztatzea
- Hiztegi_sarreratik_unitate_lexikalera

Ataza primitiboak:

- Analisi_morfologikoa: anali
- Hiztegi_sarrera_diskriminatzea: dis-sar
- Adieren-lista-sortzea: konlist
- Kontzeptua_diskriminatzea: dis-kon
- Definizio_galdea: ddef
- Definizioa_diskriminatzea: dis-def
- Definizio_birformulaketa: rdef
- Definizio_egiaztapena: vdef
- Propietate_galdea: dpro
- Propietate_egiaztapena: vpro
- Elkarbide_bilaketa: drap
- Elkarbide_egiaztapena: vrap
- Diferentzien_galdea: ddif
- Kontzeptu_bilaketa_tesaurikoa: rths
- Baliokide_galdea: equiv
- Baliokidea_diskriminatzea: dis-bal
- Baliokidearen_egiaztapena: vequiv
- Adibide_galdea: adib

- Lokailu_sintaktikoen_galdea: pat-sint
- Bateragarritasun_semantikoaren_egiaztapena: konp-sem
- Kokakide_lexikalen_galdea: kokak
- Sorkuntza_diskriminatzea: dis-sor
- Aditz_erregimenaren_galdea: adi-reg
- Sorkuntza_morfologikoa: prod
- Itzulpen_lexikala: itzul-lex

Transferentzi atazak:

- Testuingurua_eskuratzea
- Itzultzaile_eredua_eskuratzea
- Gogoko_hizkuntza_aukeratzea
- Hiztegi_hobetsia_aukeratzea
- Konnotazio_finkatzea

ANHITZ egoerak

Nekeza da egoera guzti-guztiak aurrikustea kontzeptualizazioan. Beti ere, bistan da, zein mailatarainoko fintasunarekin adierazi nahi diren, hark erakutsiko du zenbat egoera bereizi, izendatu eta definitu behar diren. Egin dezagun kontu hiru hauek direla hasieran igarritako egoerak:

- ulertua(sorburu_definizioa)
- aurkitua(xede_kontzeptua)
- sortua(xede_unitatea)

Hiru egoera horiek ataza-helburuen eta egoeren arteko korrespondentziaz ohartarazten gaituzte⁵¹. Korrespondentzia horren eremua, gainera, hedagarria da zeinnahi ataza eta egoeratar, izan ere, ataza ororen burutzapenak berez baitakar egoera-aldaketa. Hortaz, zertan ibili egoerak definitzen atazekin nahikoa ondo balia bagintezke?

Aurreregi joan baino lehen aberats dezagun orain arteko hurbilpena. Atazen eta egoeren arteko korrespondentziarena onartuta ere, ezin dugu ahantzi itzulpena prozesua dela eta prozesu horretako edozein egoerak aurretik egindako —eta, agian, egiteke dauden— atazen eraginak jaso behar dituela. Horrek itzulpen-egoeren konbinazioa iradokitzen du egoera konplexuak adierazteko. Eta hori kontuan mamitzen da gure proposamena: and, or eta not eragile logikoak erabiltzea egoerak idazteko. Adibidez:

- (hiztegi_sarreratik_definizioa_lortua(sorburu_definizioa)) and

⁵¹ sorburu_unitatearen_ulerkuntza -> ulertua(sorburu_definizioa),
baliokidea_bilatzea -> aurkitua(xede_kontzeptua),
xede_unitatearen_sorkuntza -> sortua(xede_unitatea)

```
(not (esanahia_diskriminatua(sorburu_definizioa)))
- ((ulertua(sorburu_definizioa)) or
  (aurkitua(xede_kontzeptua)))
  and
  (not (sortua(xede_unitatea)))
```

Horrela sor daitekeen amaigabeko zerrendak osatzen du ANHITZ egoeren unibertsoa.

II.4.3.2 Espezifikazioa.

Oraindik orain esana dugu ereduko kontzeptuen —kontzeptualizazioan identifikatutako horiexen— zehaztapen formala lantzen dela espezifikazioan. Zehatzago genion gainera: kontzeptuak eta beraien arteko harremanak dira zehaztu beharrekoak, horra.

Zehaztapenerako aukeratu dugun manierak ataza hartzen du gune. Atazek aldatzen dituzte egoerak, atazek erabiltzen dituzte objektuak. Atazatan lotzen eta jartzen dira harremanetan batzuk besteekin. Aproposa da, hortaz, atazatan oinarritzea espezifikaziorako. Arestion eman dugu ANHITZ atazen deskribapenerako klausula-egitura, CML lengoaiaren irudikoa, eta hori baliatuko dugu, hain zuzen, atazak espezifikatzeko. Hala ere, control-structure klausulak alderdi operazionala nabarmentzen duela-eta, horri operazionalizazioan emango diogu tarte. Acceptance-conditions eta competence-conditions klausulek, bestalde, funtzio bikoitza betetzen dute gure hurbilpenean: espezifikatzea, batetik, atazaren sarrera/irteera baldintzak zehaztuz, eta planifikatzea, bestetik, uneko ataza egokiena zein den erabakitzean lagunduz. Gauden honetan maila oso abstraktuan baizik ez ditugu landuko klausulok eta ez diogu planifikazioari arretarik eskainiko, § IVn egokiera izango dugu-eta.

Ulertzekoa denez, ataza guztien espezifikazioen berri eman beharrean, bakan batzuk, gure ustetan adierazgarriak-edo, deskribatuko ditugu segidan. Ataza primitiboak § II.4.4rako utzi ditugu, horiek lotura estua baitute inferentzi geruzarekin eta, gainera, erabilera apartekoa dutenez, egoki iritzi baitiogu bereiz azaltzeari.

Horrela bada, luzamendutan ibili gabe, adieraz ditzagun jarraian adibide gisako hiru atazaren espezifikazioak.

Itzulpen-unitatea_itzultzea

Goi-ataza, edo ia hobeto, ataza gorena da hau gure ereduan. Testuinguru jakin batean dagoen unitate lexikala xede-hizkuntzara bihurtzea du xede. Zeregin horretan hiztegiek berealdiko garrantzia dute, horregatik, azken buruan, ANHITZ ereduan hiztegi-kontsultekin zerikusia duten azpiatazatan adierazten dira problema honen ebazpideak.

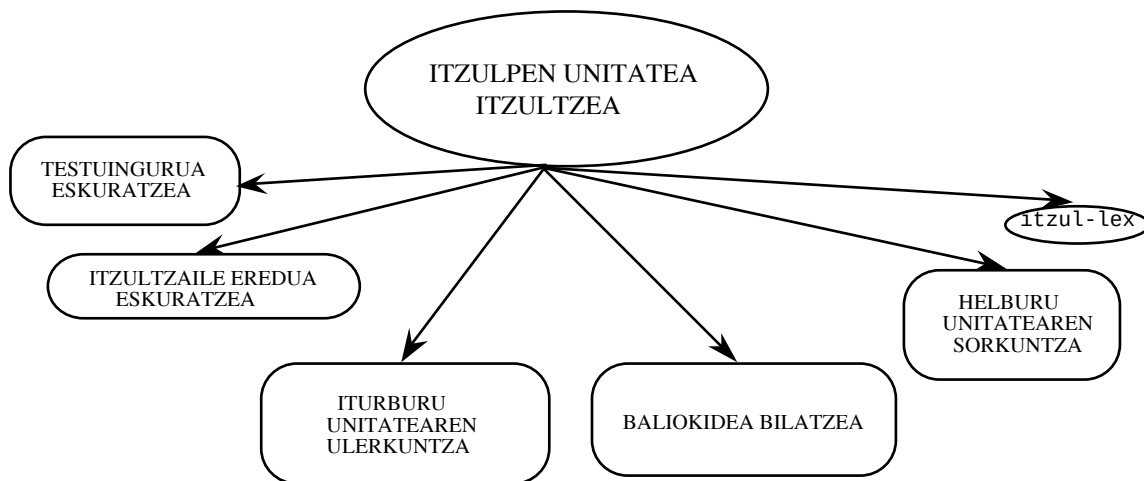
Espezifikazioa:


```

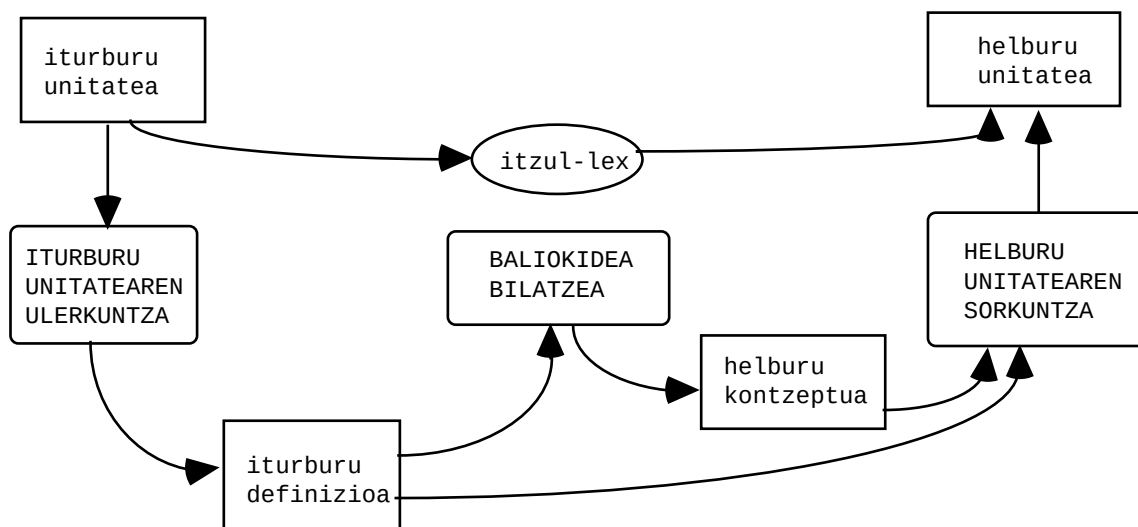
task-knowledge
  task#1: sorburu_unitatea_itzultzea
    task-specification
      task-definition
        goal: unitate lexikal baliokidea aurkitzea xede-hizkuntzan
        input: sorburu_unitatea: unitate lexikala
        output: xede_unitatea: unitate lexikala
      task-body
        type: konposatua
        parents: nil
        sub-tasks:
          itzultzaile_eredua_eskuratzea,
          testuingurua_eskuratzea,
          sorburu_unitatearen_ulerkuntza,
          baliokidea_bilatzea,
          xede_unitatearen_sorkuntza
        additional-roles:
          testuingurua: testu_eredua,
          itzultzailea: itzultzaile_eredua,
          hikuntzak: hizkuntza*,
          hiztegiak: hiztegia*,
          sorburu_sarrera, xede_sarrera: hiztegi_sarrera,
          sorburu_kontzeptua, xede_kontzeptua: kontzeptua,
          sorburu_morfologia, xede_morfologia: morfologia,
          sorburu_definizioa, xede_definizioa: definizioa,
        acceptance-conditions:
          (eskura(sorburu_unitatea))
        competence-conditions:
          (sortua(xede_unitatea))

```

task-structure
decomposition-structure:



data-flow:



Hiztegi-sarreratik definizioa lortzea

Behin hiztegi-sarrera bat daukagula, haren definizio ulergarria eta egokia lortzera jotzen du ataza honek. Egiteko zailik baldin bada ataza honetan, egiteko hori adiera egokia aukeratzea da. Adiera diskriminatzea zaila da, bai, baina zenbaitetan definizio ulergarriak ematea ere ez da makala. Itzultzailearen mailak, gogoko hizkuntzak, hobetsitako hiztegiak, horiek denek lagundu beharko dute definizio aproposena lortzen. Konparazio batera, sorburu-hizkuntzaren nola-halako ezaguera duen itzultzaileari definizio elebidundua —gogoko hizkuntzan— eskainiko zaio hobetsitako hiztegitik aterea⁵². Aitzitik, ordea, sorburu-hizkuntzaren ezaguera handiko itzultzaileari definizio elebakarra ematea izango da onena. Joera horiek, besteak beste, (Ard, 1982)n edo (Laufer & Melamed, 1994)n azaltzen dira, eta gure eskarmentuak ere halatsu erakutsi digu.

Espezifikazioa:

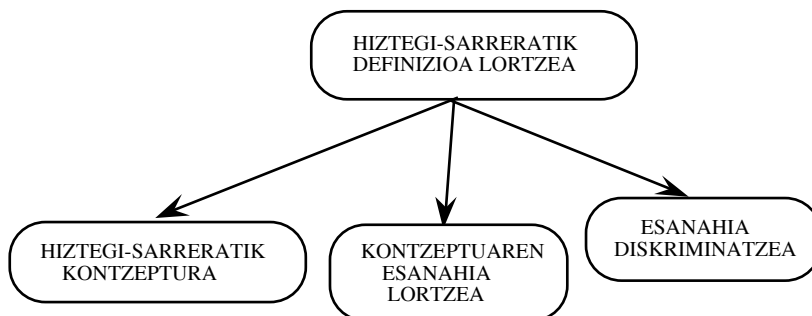
⁵² Arazoa, hortaz, definizio elebakarretik elebidundua lortzean datza. Hori definizio_galdea ataza primitiboan landuko da.

```

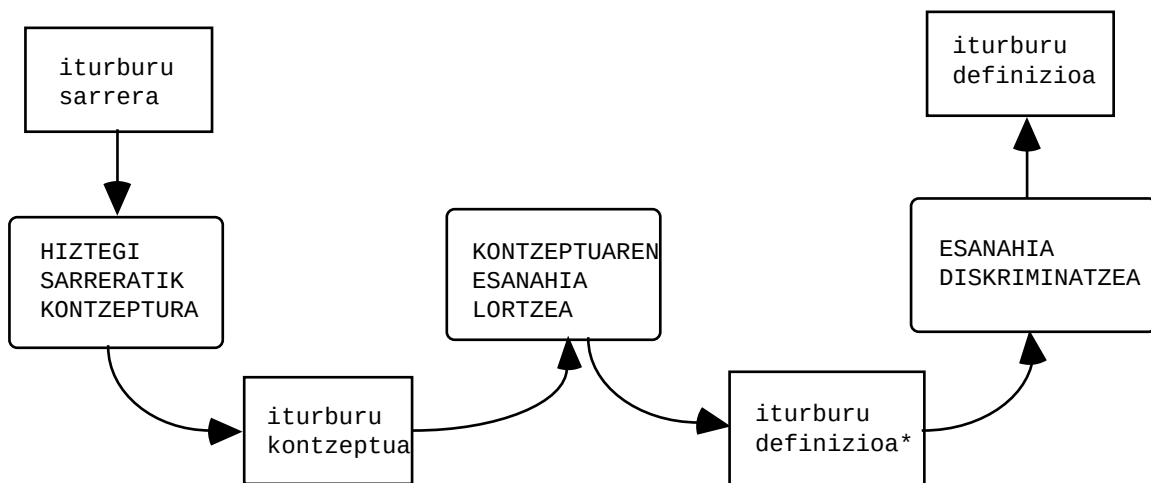
task-knowledge
task#8: hiztegi_sarreratik_definizioa_lortzea
task-specification
task-definition
goal: hiztegi-sarreraren definizio egoki eta ulergarria lortzea
input: sorburu_sarrera: hiztegi_sarrera
output: sorburu_definizioa: definizioa
task-body
type: konposatua
parents: esanahia_eridetea
sub-tasks:
hiztegi_sarreratik_kontzeptura,
kontzeptuaren_esanahia_lortzea,
esanahia_diskriminatzea
additional-roles:
testuingurua: testu_eredua,
itzultzailea: itzultzaile_eredua,
hizkuntzak: hizkuntza*,
sorburu_kontzeptua: kontzeptua,
kontzeptu_lagungarrien_lista: kontzeptua*,
definizio_lista: definizioa*,
definizio_lista_lagungarria*: (definizioa)*
acceptance-conditions:
(eskura(sorburu_sarrera))
competence-conditions:
(sortua(sorburu_definizioa))

```

task-structure
decomposition-structure:



data-flow:



Xede-unitatearen sorkuntza

Josh Ard-ek (1982:7) horrela zioen:

Successful written communication often depends on the use of a common, expected word, not just a possible one.

Itzulpen egokia lortzeko giltzarria horretan dago, testuinguruan era naturalez txertatuko den unitatearen sorkuntzan. Starren eta Thelen-ek (1988) itzulpenaren formulazioaz ari direla hiru mekanismo aipatzen dituzte: transposizioa, modulazioa eta adaptazioa; dena dela, autore hauek ez dute hiztegiren aipurik egiten, seguru asko jabetzen zirelako hiztegien antolaketa ohikoak ez duela laguntza gehiegirik ematen. Guk uste, ordea, eta guk bezala beste askok, sortze- edo kodetze-beharrei erantzungo dieten hiztegiak premiazkoak direla. Eta, hala, xede-unitatearen sorkuntzan hiztegia nola erabiltzen den —erabili beharko litzatekeen— jorratzen da ataza honetan.

Espezifikazioa:

task-knowledge

task#36: xede_unitatearen_sorkuntza

task-specification

task-definition

goal: testuinguruari ondo egokitzen zaion unitatea sortzea

input: xede_kontzeptua: kontzeptua,
xede_definizioa: definizioa

output: xede_unitatea: unitate_lexikala

task-body

type: konposatua

parents: itzulpen_unitatea_itzultzea

sub-tasks:

sorkuntza_hipotesiak_lortzea,
sorkuntza_hipotesiak_diskriminatzea,
sorkuntza_hipotesia_egiaztatzea,
hiztegi_sarreratik_unitate_lexikalera

additional-roles:

testuingurua: testu_eredua,
itzultzailea: itzultzaile_eredua,
hizkuntzak: hizkuntza*,
sorkuntza_lista: kontzeptua*,
?egiaztatua: boolean,
xede_sarrera: hiztegi_sarrera,
xede_morfologia: morfologia

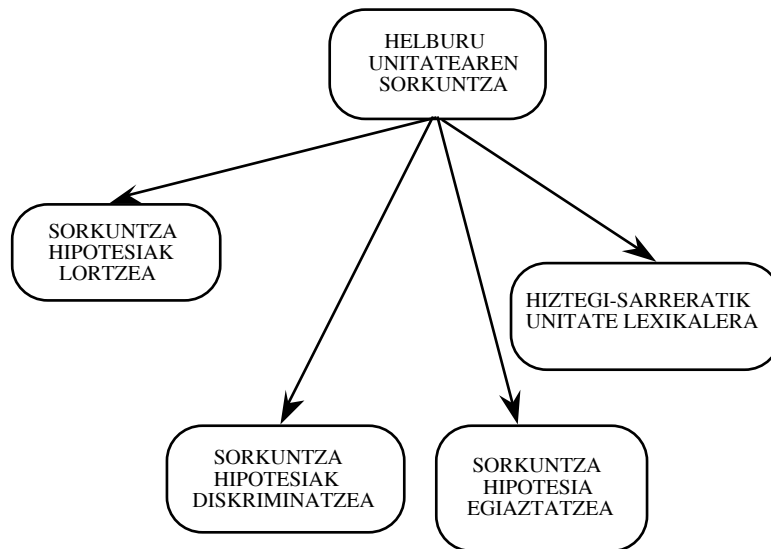
acceptance-conditions:

(eskura(xede_kontzeptua)) or
(eskura(xede_definizioa))

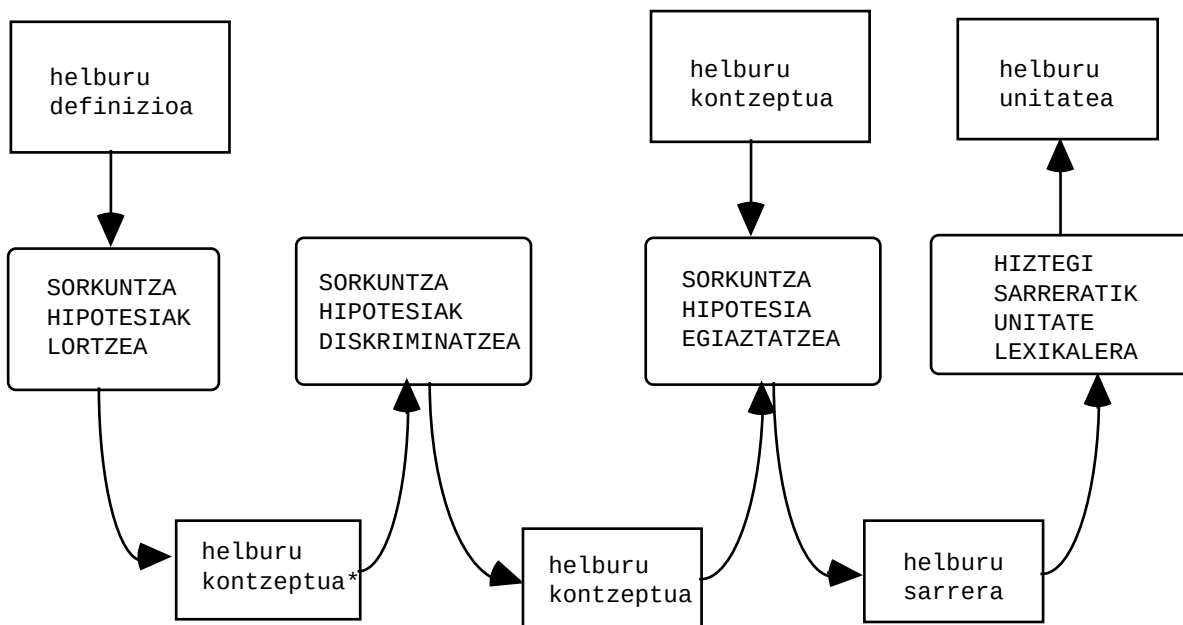
competence-conditions:

(sortua(xede_unitatea))

task-structure
decomposition-structure:



data-flow:



Orain arteko espezifikazioak mostratarako adina ozta-ozta badira ere, eten beharrean gaude, horien luzerak ez du-eta segitzea agintzen. Ordainez, argigarria datekeelakoan, II.2 irudian ataza guztien deskonposizioa bildu dugu, gorenetik hasi eta primitiboetaraino, espezifikazio atalaren buruan ikuspegi orokorra, nahiz-eta larria, emateko asmoz.

II.4.3.3 Operazionalizazioa.

Hirugarren formalizazio-urrats honetan atazen kontrol-egiturei heltzen zaie, atazek zer egiten dutenari baino nola egiten dutenari jaramon eginez. "Nola" galdegai horrek kode algoritmikoa eskatzen du erantzun gisa eta horregatik kontrola adierazteko, control-structure klausulan, lengoaia algoritmiko agintzailea erabili dugu gurean.

Hara nolakoa den, informalki deskribatuta, lengoaia hori⁵³:

- Kontrol-mekanismoak:
 - Sekuentziazioa: ";" ikurraz adierazten da.
 - Baldintzazko konposaketa: "case B_1 : S_1 ; ...; B_n : S_n ; end-case" eredu jarraitzen da, non B_i -ak baldintzak diren eta S_i -ak sententziak. Betetzen den lehenengo baldintzari doakion sententzia da exekutatzen dena.
 - Iterazioa: "loop S_1 ; exit when B ; S_2 ; end-loop" moldearen arabera. S_1 eta S_2 sententzi multzoak egikaritzen dira harik eta B irteera-baldintza bete arte.
 - Aukeran: "[S]", S sententzia exekutatzea barne-kontrolaren menpe ez dagoenean. Ohikoa da makoen erabilera transferentzi atazetan, horiek exekutatu edo ez erabiltzailearen esku baitago.
 - Noiznahi: " $\bullet \rightarrow S$ ", S sententzia noiz exekutatzen den zehatz-mehatz jakiterik ez dagoenean. Normalean sententzia lagungarriak edo osagarriak dira, eta esaterik badago atazaren kontrol-egituran gertatzen direla, baina zaila da oso une jakin batean kokatzea.
- Sententziak:
 - Sententzia hutsa: null;
 - Ataza konposatu parametrizatuak sententziak dira.
 - Ataza primitibo parametrizatuak sententziak dira.
 - Transferentzi ataza parametrizatuak sententziak dira.
- Parametrizazioa:
 - Datu parametroak: "in p ", non p parametroa den.
 - Eraitza parametroak: "out p "
 - Datu-Eraitza parametroak: "in out p "
- Baldintzak:
 - Baldintzak egoerak bezala adierazten dira.

Honakoan ere hiru adibidez baliatuko gara operazionalizazioan zernolako erantzak lortu ditugun erakusteko:

⁵³ Apropos saihestu dira deskribapen honetan inferentzia- eta estrategia-geruzak ukitzen dituzten lengoaiaren atalak. Horiek § IV.ean agertuko dira.

Itzulpen-unitatea itzultzea

Itzultzeak, eta zer esanik ez hiztegien laguntzaz itzultzeak, berezkoa du saioak egin eta, gogoko emaitzak lortu ezean, atzera jotzea beste saialdiei ekiteko. R.R.K. Hartmann (1994:293) iritzi berekoa da: "*the coverage of lexical equivalents in the bilingual dictionary is a hit-and-miss, trial-and-error task...*". Gure ustetan, hiztegi elebidunei buruzkoa hedagarria da beste atazetara ere. Ezaugarri horrek nabarmen markatzen du kontrol-egitura.

control-structure:

```
begin
  [testuingurua_eskuratzea (out testuingurua, out hizkuntzak);]
  [itzultzaile_eredua_eskuratzea (out itzultzailea);]
  loop
    case
      not ?(ulertua(sorburu_definizioa)):
        sorburu_unitatearen_ulerkuntza (in sorburu_unitatea,
                                         out sorburu_definizioa);
      not ?(aurkitua(xede_kontzeptua)):
        baliokidea_bilatzea (in sorburu_unitatea,
                              out xede_kontzeptua);
      ?ulertua(sorburu_definizioa):
        xede_unitatearen_sorkuntza (in xede_definizioa,
                                     out xede_unitatea);
      ?aurkitua(xede_kontzeptua):
        xede_unitatearen_sorkuntza (in xede_kontzeptua,
                                     out xede_unitatea);
    end-case;
    exit when ?sortua(xede_unitatea);
  end-loop;
  •-> itzul_lex (in sorburu_unitatea, out xede_unitatea);
end
```

Hiztegi-sarreratik definizioa lortzea

Sarreratik definiziorainoko bidean bidegurutze franko kausituko du itzultzaileak: adiera bat aukeratzea dela, definizio ulergarria osatzea dela, e.a. Gainera, maiz, kontzeptutik kontzepturako, definiziotik definiziorako ibileran murgilduko da itzultzailea, gehienetan definizio bateko kontzeptu ezezagunaren definiziora salto eginez. Zenbaitetan, ordea, irudi du noraezean doala itzultzailea ideien asoziazioz-edo han eta hemen kontsultak eginez. Ibilgo horretan ageri diren kontzeptu nahiz definizioei lagungarrien rola egokitu diegu, eta eredu horretako definizio-bilaketa "noiznahiko" kontrolaz ($\sqrt{->}$) adierazi dugu.

control-structure:

```
begin
  loop
    hiztegi_sarreratik_kontzeptura (in sorburu_sarrera, out
                                    sorburu_kontzeptua);
  loop
    case
      not ?(eskura(definizio_lista)):
        kontzeptuaren_esanahia_lortzea (in sorburu_kontzeptua,
                                         out definizio_lista);
        esanahia_diskriminatzea (in definizio_lista,
                                  out sorburu_definizioa);
      ?(eskura(definizio_lista)):
        esanahia_diskriminatzea (in definizio_lista,
                                  in out sorburu_definizioa);
    end-case;
  end-loop;
end
```



```

        exit when egokia(sorburu_definizioa);
    end-loop;
    exit when sortua(sorburu_definizioa);
end-loop;
•-> kontzeptuaren_esanahia_lortzea (in kontzeptu_lagungarria,
                                out definizio_lista_lagungarria);
end

```

Xede-unitatearen sorkuntza

"How the system works in practice to take the dictionary user from (pre-lexical) meaning to an appropriate word or phrase to express this meaning", (Rundell & Ham, 1994:173)tik erauzi ditugun hitzok ederki adierazten dute ataza honen operazionalizazioan zeri heldu diogun.

control-structure:

```

begin
  loop
    case
      not ?sortua(xede_kontzeptua):
        sorkuntza_hipotesiak_lortzea (in xede_definizioa,
                                       in testuingurua,
                                       out sorkuntza_lista);
    end-case;
    sorkuntza_hipotesiak_diskriminatzea (in sorkuntza_lista,
                                         in out xede_kontzeptua);
    sorkuntza_hipotesia_egiaztatzea (in out xede_kontzeptua,
                                     out ?egiaztatua);
    exit when ?egiaztatua;
  end-loop;
  hiztegi-sarreratik-unitate-lexikalera (in xede_sarrera,
                                           in xede_morfologia,
                                           out xede_unitatea);
end

```

II.4.4 Ataza primitiboak.

Ataza primitiboak, behe-mailakoak, zuzenean inferentzi geruzarekin lotura dutenak direla esana dugu aurrez, eta hala da, izan ere, deskonposatzen hasita, inferentziatan deskonposatzen baitira, ez ataza sinpleagotan. Ataza bakunak dira, tarteko egoerarik gabe hasi eta buka egiten direnak.

Erabiltzailearen ikuspegitik ataza hauek oinarri-oinarrizko kontsultekin identifikatzen dira. Identifikazio hori dela kausa, ANHITZen eta erabiltzailearen arteko elkarrekintza ataza primitiboen bidez adierazten da gehienetan. Horren adierazgarriak izaki, komenigarria iruditu zaigu guztiak, diskriminatzeakoak salbuetsita, zerrendatzea eta espezifikatzea banan-banan.

ANHITZ inguru elebidunaren ataza primitiboak HIZTSUA sistema elebakarraren (Arregi *et al.*, 91; Artola, 93) funtzionalitatearen egokitzapen eta aberasketa gisa ikus daitezke. HIZTSUA sistemarako deskribatu ziren funtzio haiek itzulpenera zuzendutako ingurumarian paratu dira, aurretik argitaratutako zenbait lanetan (Agirre *et al.*, 93; Agirre *et al.*, 94b) adierazi den bezala.

II.4.4.1 Ataza diskriminatzaileak.

Aukera bat baino gehiago dagoenean, edozein dela ere itzulpen-egoera, funtzio diskriminatzaileek lehentasun-ordena ezartzen dute aukera horien artean. Nola unitate-lexikal bati hiztegi-sarrera anitz edo hiztegi-sarrera bakarrari adiera bat baino gehiago egokitu lekizkiekeen, hala gertatzen da adieren definizioekin, baliokideekin edota sorkuntza-hipotesiekin. Horietan denetan aniztasuna —itzultzaileari hamaika zalantza sortaraziko dizkion aniztasun gogaikarria!— da nagusi.

Gauzak horrela, ANHITZ sistemari begira galdera da ea hiztegiek zein neurritaraino erraztu lezaketen itzultzailearen etengabeko aukeratze-lan hori. Edo, zein teknika erabil daitekeen hiztegi-tako informazioan oinarri hartu eta testuinguruaren arabera aukerak lehenesteko. Egia da ikerlan askoren gaia dela testuingurupeko desanbiguazioarena (Agirre & Rigau, 95), eta ez inola ere lerrootan laburbiltzekoa. Gure xedea teknika horietakoren bat baliatu eta ANHITZ sisteman txertatzea da, hartara itzultzaileari laguntza aberats eta osotuagoa emateko diskriminatze-zereginetan.

Ohargarria da ANHITZeko ataza diskriminatzaile hauetan helburua ez dela aukera bakarra hartzea eta besteak alboratzea. Horrela balitz atzera jotzeko ahalbidea galdu egingo litzateke, aukera bakarrak ez baitu besterik, hori edo deus ez. Zuhurtzia handiagoz aritzeak aukerarik ez baztertzea eskatzen du, lehentasunak markatzea bai, baina azken helburua iritsi bitartean lehentasun txikieneko aukerak ere presente edukitzea.

Kontzeptua `_diskriminatzea` (dis-kon) hartuko dugu adibidetzat ataza diskriminatzaileen espezifikazioak erakusteko, gainontzeko guztiak parekoak direnez, bakarra erakustearekin nahikoa baita:

```
task-knowledge
  task#12:      dis-kon
    task-specification
      task-definition
        goal:      unitate lexikal bati egokitu zaion adiera-lista
                    lehentasunez ordenatzea itzulpen-testuinguru baten
                    arabera.
        input:      adiera_lista: adiera*
        output:     adiera_lista: adiera* (ordenatua)
      task-body
        type:       primitiboa
        parents:    adiera_diskriminatzea
        sub-tasks:  nil
        additional-roles:
          testuingurua: testu_eredua
```

II.4.4.2 Ulerkuntzarakoak.

Azpiatal honetara bildu ditugun ataza primitiboak sorburu-unitatearen ulerkuntzan erabiltzen dira gehienbat. Sailkatzeko nahiak eraginda daude hemen, ez baitago esaterik beste atazetan erabiltzen ez direnik. Maiztasun-irizpideak erabili ditugu ataza primitiboak han eta hemen sailka banatzeko.

Definizio-galdea (ddef)

Usuen erabiltzen den ataza da ulerkuntzan. Arrakastatsua izan edo ez, zer esan handia du itzulpen-inguruneaz ANHITZek duen ezagutzak. Esate baterako, ez da gauza bera itzultzailearen gustuko hiztegia eta hizkuntza jakinda edota horiek jakin gabe definizio-galdea burutzea.

Hala bada, definizio-galdea erarik malguenean mamitu nahi izan da ANHITZen. Hiztegiak eta emaitza-hizkuntza aukeran daude eta ez da inolako konbinaziorik debekatzen. Posible da, konparazio batera, frantsesezko hiztegi elebakarraren definizioak euskaraz eskatzea. Definizio elebakarrak elebiduntzeko gaitasun hori benetan lagungarria suertatzen zaie hizkuntz ezaguera aberatsegia ez duten itzultzaileei. Oro har, gainera, definizio-galdearekin egin den bezalaxe, ANHITZeko funtzio primitiboek badute hizkuntz bihurketarako aukera.

```
task-knowledge
  task#15:      ddef
    task-specification
      task-definition
        goal:      kontzeptu baten definizio ulergarria eskuratzea
        input: kontzeptua: kontzeptua -- rol anitz joka ditzake
              azalpen_maila: (testuala, lokala, heredatua, deduzitua)
              hiztegia: hiztegia,
              emaitza_hizkuntza: hizkuntza
        output:    definizioa: definizioa
      task-body
        type:      primitiboa
        parents:    kontzeptuaren_esanahia_lortzea,
                  baliokidetza_egiaztapena
        sub-tasks:  nil
```

Definizioa objektuak bi erako gauzatzeak izan ditzake. Izan daiteke testua, hiztegi arruntetan ohikoa den bezala, edo izan daiteke erlazio-multzoa. Azken hurbilpen honetan, kontzeptu baten definizioa *kontzeptu horrek beste kontzeptu batzuekin dituen erlazioen multzo* bezala hartzen da. Ikuspegi erlazional horrek jarraitzaile asko ditu semantika lexikalaren munduan (Evens, 88). Gurean, "definizio erlazionala" erlazioen konjuntzio da, erlazio-multzo baino areago.

Azalpen_maila parametroak lau maila bereizteko aukera ematen du definizioa aurkeztean:

- Testuala ("textuel"): Adierari dagokion definizio-testua itzultzen da, besterik gabe.
- Lokala ("original"): Definizio erlazionala bueltatzen da. Erlazioak, kontzeptuaren definiziotik ezezik, kontzeptu horrek besteren definiziotan dituen agerpenetatik ere jasotzen dira.
- Heredatua ("herité"): Lokalean bezala, honakoan ere definizio erlazionala itzultzen du DDEFek, baina kontzeptuen arteko erlazio taxonomikoak ere ustiatuz.
- Deduzitua ("deduit"): Inferentzi geruzako mekanismoak baliatuta deduzitzen diren erlazio guztiak ematen dira maila honetan.

Adibideak:

Itzul.: (ddef '|maison I 1| 'textuel 'th-lppl 'frantsesa)

ANHITZ: "e'difice , logement"

Itzul.: (ddef '|maison I 1| 'original 'th-lppl 'frantsesa)

ANHITZ: ((and (|maison I 1| SYNONYMES |e'difice I 1|)
(|maison I 1| SYNONYMES |logement I 1|)
(|maison I 1| POSSESSEUR+INV |bloc I 1#1|)))
non |bloc I 1#1|= "bloc de maisons"

Itzul.: (ddef '|maison I 1| 'deduit 'th-lppl 'frantsesa)

ANHITZ: ((and (|maison I 1| SYNONYMES |logement I 1|)
(|maison I 1| SYNONYMES |e'difice I 1|)
(|maison I 1| SYNONYMES |logis I 1|)
(|maison I 1| SYNONYMES |demeure I 1|)
(|maison I 1| SYNONYMES |habitation I ?|)
(|maison I 1| POSSESSEUR+INV |bloc I 1#1|)
(|maison I 1| HYPERONYME |ba^timent I 1|)
(|maison I 1| HYPONYME |mairie I 1|)
(|maison I 1| HYPONYME |institution I 2|)
(|maison I 1| HYPONYME |cha^teau I 2|)
(|maison I 1| HYPONYME |cha^teau I 1|)
(|maison I 1| HYPONYME |baraque I 2|))),

Hiztegiko definizioaz gain, deduzitutako erlazioak ere ikus daitezke, azalpen-maila "deduit" baita.

Itzul.: (ddef '|maison I 1| 'herite 'th-hleh 'euskara)

ANHITZ: ((and (|etxe I 1| HYPERONYME |eraikuntza I 1|
(|etxe I 1| PREDICATION |presta(tu) I 1#1|)
(|etxe I 1| MEMBRE-DE |hiri I 1|)
non |presta(tu) I 1#1|= "gizona bizitzeko prestatu"

Adibide honetan hizkuntz bihurteta egiten da definizio-galdearekin batera.

Definizio-birformulaketa (rdef)

ANHITZek emandako definizioa ulertu ez eta orduan eskatzen du itzultzaileak definizioa birformulatzea. Beti ere DDEF baten ondoren doa, haren emaitza definizioa aldatu eta esanahia ulergarriagoa egiteko.

```
task-knowledge
  task#16:      rdef
    task-specification
      task-definition
        goal:      kontzeptu baten aurretiko definizioa birformulatzea
        input: definizioa: definizioa,
              azalpen_maila: (testuala,      lokala,      heredatua,
                             deduzitua)
              hiztegia: hiztegia,
              emaitza_hizkuntza: hizkuntza
        output:    definizio_birformulatua: definizioa
      task-body
        type:      primitiboa
        parents:    kontzeptuaren_esanahia_lortzea
        sub-tasks:  nil
        additional-roles:
          kontzeptua: kontzeptua
```

Adibideak:

```
Itzul.: (rdef '((and (|ikastetxe I 1| HYPERONYME |etxe I 1|)
                     (|ikastetxe I 1| DEFINI-PAR |etxe I 1#7|)
                     (|ikastetxe I 1| ANCETRE |gauza I 1|)
                     (|ikastetxe I 1| ANCETRE |eraikuntza I 1|)))
        'deduit 'th-hleh 'euskara)

ANHITZ: ((and (|ikastetxe I 1| HYPERONYME |eraikuntza I 1|)
              (|ikastetxe I 1| PREDICATION |presta(tu) I 1#1|)
              (|ikastetxe I 1| MEMBRE-DE |hiri I 1|)
              (|ikastetxe I 1| DEFINI-PAR |etxe I 1#7|)
              (|ikastetxe I 1| ANCETRE |gauza I 1|)
              (|ikastetxe I 1| PREDICATION |ari_izan I 1#1|)),
  non    |etxe I 1#7|= "irakaskuntzan ari den etxea", eta
        |ari_izan I 1#1|= "irakaskuntzan ari izan"

|etxe I 1| hiperonimoaren definizioa txertatu da |ikastetxe I 1|-en definizioan.
```

Definizio-egiaztapena (vdef)

Egiaztapen-funtzioak, oro har, garrantzi handikoak dira hiztegien erabilpenean, izan ere, aurretiko usteak berresteko asmotan jotzen baitute maiz hiztegi itzultzaileak, batez ere aurreratuek. Joera horren berri ematen dute hiztegien erabilera-azterketan ihardundako beste autore batzuek: (Neubach & Cohen, 1988:7)n, kasu baterako, *"One of these students said that she used the monolingual dictionary a lot out of intellectual curiosity in order to reinforce her expectations,..."* jakinarazten da. Eta gure azterlanetan ere jaso ditugu antzeko usteak.

Definizio-egiaztapena horietarik bat da, itzultzailearen ustezko definizioa egiaztatzeko aukera ematen duena:

```
task-knowledge
  task#22:      vdef
    task-specification
      task-definition
        goal:      kontzeptu baten definizioa egiaztatzea
        input: kontzeptua: kontzeptua,
                  erabiltzaile_definizioa:
                    definizioa, --definizio erlazionala
        output:      hiztegia: hiztegia,
                    emaitza_hizkuntza: hizkuntza
                    ?definizio_egiaztatua: boolean,
                    definizioa: definizioa      --hiztegitik ateratako
                                                definizioa
      task-body
        type:      primitiboa
        parents:    esanahia_egiaztatzea,
                    baliokidetza_egiaztapena
        sub-tasks:  nil
```

Adibideak:

```
Itzul.: (vdef  '|palais I 1|
              '(and (|palais I 1| hyperonyme |construction I 1|)
                    (|palais I 1| caracteristique |grand I 1|))
              'th-hleh 'euskara)
```

```
ANHITZ: True,
(Nire definizioan ere:
  (|jauregi I 1| hyperonyme |eraikuntza I 1|)
  (|jauregi I 1| caracteristique |handi I 1|))
```

Propietate-galdea (dpro)

Galde-funtzio honen bitartez kontzeptuaren atributu batek zein balio dituen jakin daiteke. Normalean kontzeptuaren esanahia bete-betean harrapatzetik gertu dagoela egiten du DPRO galdea itzultzaileak, artean ilunska dagoen alderdiren bat argitzeko.

Badira DPROren bidez adierazten diren garrantzi eta maiztasun handiko galdeak. Horien artean behinena sinonimo-galdea da, batez ere sorkuntzan erabilia. Hori dela-eta kokatu dugu DPRO sorkuntza_hipotesiak_lortzearen azpiataza gisa ere.

```

task-knowledge
  task#17:      dpro
    task-specification
      task-definition
        goal:      kontzeptu baten atributu baten balioak ematea
        input: kontzeptua: kontzeptua,
                atributua: kontzeptua edo atributu_primitiboa,
                hiztegia: hiztegia,
                emaitza_hizkuntza: hizkuntza
        output:     atributu_balio_lista:
                    (atributu_primitiboa, atributuaren_balioa)*
      task-body
        type:      primitiboa
        parents:    kontzeptuaren_esanahia_lortzea,
                    baliokidetza_egiaztapena,
                    sorkuntza_hipotesiak_lortzea
        sub-tasks:  nil

```

Adibideak:

Itzul.: (dpro '|buru I 5| 'hyperonyme 'th-lppl 'euskara)

ANHITZ: ((HYPERONYME |pertsona I 1|))

Itzul.: (dpro '|kale I 2| 'ensemble-de 'th-hleh 'euskara)

ANHITZ: ((ENSEMBLE-DE |etxe I 1#2|)
 non |etxe I 1#2|= "elizaren inguruan biltzen diren etxeak"

Elkarbide-bilaketa (drap)

Bi kontzepturen arteko erlazioa bilatzen da ataza honetan. Guk elkarbide deitu diogu sare kontzeptualean bi kontzeptu lotzen dituen bideari, DRAPeK emaitza gisa itzuliko duenari, hain juxtu.

```

task-knowledge
  task#20:      drap
    task-specification
      task-definition
        goal:      bi kontzepturen arteko elkarbidea lortzea
        input: kontzeptua1, kontzeptua2: kontzeptua,
                hiztegia: hiztegia,
                emaitza_hizkuntza: hizkuntza
        output:     kontzeptu_arte_harremana: elkarbide_espresioa
      task-body
        type:      primitiboa
        parents:    esanahia_diskriminatzea,
                    baliokidea_diskriminatzea
        sub-tasks:  nil

```

Adibideak:

Itzul.: (drap '|attache I 1| '|korapilo I 1| 'th-hleh 'euskara)

ANHITZ: ((and (|lokarri I 1| THEME+INV |korapilatu I 1|)
(|korapilatu I 1| INSTRUMENTAL |korapilo I 1|)))

Bi hizkuntzako kontzeptuen arteko elkarbidea erakusten da, euskarazko hiztegian egina eta euskaraz aurkeztua.

Elkarbide-egiaztapena (vrap)

Egiaztapen-funtzio honetan emandako elkarbide batek bi kontzeptu lotzen dituenentz esaten da. Erabiltzailearen esku dago ustezko elkarbidea adieraztea. Ez da sarritan erabiltzen den ataza.

```
task-knowledge
task#21:      vrap
  task-specification
    task-definition
      goal:      bi kontzeptuen arteko elkarbidea egiaztzea
      input: kontzeptua1, kontzeptua2: kontzeptua,
              kontzeptu_arte_ustezko_harremana:
                                elkarbide_espresioa
      output:      hiztegia: hiztegia,
                    emaitza_hizkuntza: hizkuntza
                    ?elkarbide_egiaztatua: boolean,
                    kontzeptu_arte_ustezko_harremana: elkarbide_espresioa
    task-body
      type:      primitiboa
      parents:    esanahia_diskriminatzea,
                  baliokidea_diskriminatzea
      sub-tasks:  nil
```

Adibideak:

Itzul.: (vrap '|matazatu I 1| '|fil I 1|
'(and (and (|matazatu I 1| theme |fil I 1|)))
'th-hleh 'euskara))

ANHITZ: True,
(and (|matazatu I 1| THEME |hari I 1|))

Adieren-lista-sortzea (konlist)

Hiztegi-sarrera bati dagozkion adierak bueltatzen ditu ataza honek.


```

task-knowledge
  task#10:      konlist
    task-specification
      task-definition
        goal:      Hiztegi-sarreraren adierak lortzea
        input: hiztegi-sarrera: hiztegi-sarrera,
                hiztegia: hiztegia
        output:     kontzeptu_lista: kontzeptua*
      task-body
        type:      primitiboa
        parents:   hiztegi_sarreratik_kontzeptura
        sub-tasks: nil

```

II.4.4.3 Baliokidea lortzekoak.

Hizkuntza desberdinetako hitzen baliokidetasuna hitzetik hortzera erabili dute lexikografia eleaniztunaz eta itzulpenaz ardura dutenek, eta ia denak bat datoz bene-benetako baliokidetasuna ezinezkoa dela esatean. Al-Kasimik (1977:63) horrela dio:

Absolute equivalents which have exactly the same semantic and grammatical function in both languages are rare.

Eta Albrecht Neubart oraindik kategorikoagoa da (1992:29):

the fictitious claims are the lexical equivalences. The factual aims are the alphabetical entries enabling the user to build hypotheses about the relations between words in two languages.

Argibideak ziren horiek, ez besterik; eta gogoan hartuko baditugu ere ez gara hasiko orain eztabaida teorikoetan. Egia ukaezina da hiztegiek ematen dituztela nola-halako baliokideak, kontzeptutik kontzepturakoak, eta horiek erabiltzen dituela itzultzaileak besterik ezean. Egia horri helduko diogu bada ANHITZen, kontuan hartuz, hori bai, baliokidera hurbiltzeko testuinguruari buruzko xehetasunak, erabilera-erregistroak, konnotazioak eta erabilera metaforikoak detektatzea beharrezkoa dela.

Baliokidea lortzeko ahaleginetan antzeman eta definitu ditugun atazen artean bat nabarmentzekotan EQUIV, baliokide_galdea, nabarmenduko genuke. Baina badira horrez gain beste primitiboak, normalean lagungarritzat hartzen badira ere, aski baliagarriak izaten direnak baliokide-hipotesiak sortzeko, diskriminatzeke edota egiaztatzeke. Horietatik batzuk azpiatal honetan deskribatuko ditugu, bestetzuk dagoeneko espezifikatu dira eta badira ondorenean deskribatuko direnak ere. Zein non erabiltzen den zehatzago jakiteko komeni da II.2 irudiko atazen deskonposaketa begien bistatik ez galtzea.

Baliokide-galdea (equiv)

Bereziki landu da ataza honetan "*hutsune*" *lexikalen* arazoa. Hutsune lexikaltzat, gurean, bi erako fenomenoak hartu dira:

- Sorburu-kontzeptuaren hitz bakarreko baliokide sinplerik ez dagoenean xede-hizkuntzan: horrelakoei aurre egiteko xede-hizkuntzako kontzeptu konplexuagoak erabiliko ditugu, sintagmatan edo esalditan adierazten diren kontzeptuak. Horiei, gure domeinuko hizpidera joz, *kontzeptu sintagmatiko* deritze (ikus § III.4.2.1.1).
- Sorburu-kontzeptuak sarrerarik ez duenean hiztegi elebidunetan: Sarrerarik ukan ez arren, hiztegi elebidun alderantzizkoan xede-kontzeptu gisa azalduko balitz, EQUIVek baliokidea aurkituko lioke itxura bateko hutsune horri. Bestela, hiztegi elebidunetan inolako erreferentziarik ez dagoenean, baliokide erabatekoaren ordeztu baliokide ahalik-eta hurbilenaren bila jotzen du EQUIVek, gehienetan kontzeptu orokorrak itzuliko ditu baliokide partzial gisa eta bakanago kontzeptu espezifikoagoak.⁵⁴

task-knowledge

```
task#33:      equiv
  task-specification
    task-definition
      goal:      kontzeptu      baten      baliokideak      lortzea      beste
                hizkuntzan
      input: sorburu_kontzeptua: kontzeptua,
                erabileremua: arloa,
                konnotazioak: konnotazioak,
                sorburu_hiztegia,                  xede_hiztegia,
                hiztegi_elebiduna: hiztegia
      output:    baliokide_lista:
                (xede_kontzeptua,baliokidetza_maila)*
  task-body
    type:      primitiboa
    parents:    adieraren_baliokidea_lortzea
    sub-tasks:  nil
```

Adibideak:

Adibideotan ageri ez diren parametroek besterik ezeko balioak hartzen dituzte.

Itzul.: (equiv '|txabola I 1|)

ANHITZ: ((|cabane I 1| 0.5))

Itzul.: (equiv '|baraque I 2|)

ANHITZ: ((|etxe I 1| 0)),

⁵⁴ Baliokidetza-mailak zenbakiez adierazten dira: erabatekoa (0.5), orokorrak (1) eta espezifikoak (0).

ez dago baliokide erabatekorik hiztegi elebidunean, baina ANHITZek |etxe I 1| kontzeptua itzultzen du, nahiz eta |baraque I 2| espezifikagoa den.

Itzul.: (equiv '|abere I 1|)

ANHITZ: ((|mammife`re I 1#1| 0.5)
(|be'tail I 1| 0.5)
(|be^te I 1| 0.5)
(|mammife`re I 1| 0)),
non |mammife`re I 1#1|= "mammife`re domestique",
|mammife`re I 1#1| kontzeptu sintagmatikoa eman da, besteren artean, |abere I 1| kontzeptuaren baliokide gisa.

Baliokidearen-egiaztapena (vequiv)

Egiaztapen-ataza honek EQUIVen antzeko langintza du, baina VEQUIVi ustezko baliokidea ematen zaio, horren egiazkotasunaz erabaki dezan.

task-knowledge

```
task#38:      vequiv
task-specification
task-definition
goal:         bi kontzeptu elkarren baliokide diren egiaztatzea
input: sorburu_kontzeptua, xede_kontzeptua: kontzeptua,
erabileremua: arloa,
konnotazioak: konnotazioak,
sorburu_hiztegia, xede_hiztegia,
hiztegi_elebiduna: hiztegia
output:       ?baliokideak: boolean,
baliokidetza_maila: (=,<,>)
task-body
type:         primitiboa
parents:      baliokidetza_egiaztapena
sub-tasks:    nil
```

Diferentzien-galdea (ddif)

Bi kontzepturen arteko aldeak edo diferentziak ageriko egitea da ataza honen xedea⁵⁵. Zaila bada kontzeptuen elkarbideak detektatzea, are zailagoa da diferentziak zehaztea. DDIFen erabilera zentzuzkoa izan dadin sare kontzeptualeko taxonomia berean egotea eskatzen zaie kontzeptuei. Alabaina, hizkuntza bat baino gehiagorekin ari garela, posible da gurean hizkuntza desberdinetako kontzeptuen aldeak lantzea.

⁵⁵ Teoria konponentzial baten arabera, hitzaren esanahia beste hitzen esanahietatik bereizten duten ezaugarrien multzo bezala ikus daiteke.

```

task-knowledge
  task#19:      ddif
    task-specification
      task-definition
        goal:      bi kontzeptuen arteko aldeak lortzea
        input: kontzeptua1, kontzeptua2: kontzeptua,
              hiztegia: hiztegia,
              emaitza_hizkuntza: hizkuntza
        output:     kontzeptu_arte_ aldeak: diferentzi_espresioa
      task-body
        type:      primitiboa
        parents:    esanahia_diskriminatzea,
                  baliokidea_diskriminatzea
        sub-tasks:  nil

```

Adibideak:

Itzul.: (ddif '|gizaki I 2| '|piztia I 1| 'th-hleh 'frantsesa),
 |gizaki I 2| eta |piztia I 1| kontzeptuen arteko aldea jakin nahi da, baina frantsesez.

ANHITZ:
 ((and (and (and (|e^tre II 1#5| caracteristique |humain I 1|))
 (and (|animal I 2#1| caracteristique |sauvage I 1|)))))

Itzul.: (ddif '|iltzatu I 1| '|korapilatu I 1| 'th-hleh 'euskara)

ANHITZ:
 ((and (and (and (|iltzatu I 1| instrumental |iltze I 1|))
 (and (|korapilatu I 1| instrumental |korapilo I 1|)))
 (and (and (|iltzatu I 1| theme |gauza I 1#6|))
 (and (|korapilatu I 1| theme |lokarri I 1|)))))
 non |gauza I 1#6|="sakabanaturik zeuden gauzak"

Propietate-egiaztapena (vpro)

Itzultzaileak kontzeptu bati ustez ezagutzen dizkion propietateak egiaztatzen ditu ataza honek.

```

task-knowledge
  task#23:      vpro
    task-specification
      task-definition
        goal:      kontzeptu baten propietateak egiaztatzea
        input: kontzeptua: kontzeptua,
              atributu_balio_lista:
                (atributu_primitiboa, atributuaren_balioa)*,
              hiztegia: hiztegia
        output:     ?propietate_egiazkoak: boolean
      task-body
        type:      primitiboa
        parents:    esanahia_egiaztatzea,
                  baliokidetza_egiaztapena
        sub-tasks:  nil

```

Adibideak:

Itzul.: (vpro '|abere I 1| ')((caracteristique |ornodun I 1|)) 'th-hleh)

ANHITZ: True

Itzul.: (vpro '|gizadi I 2| ')((ensemble-de |gizabanako I 1|)) 'th-hleh)

ANHITZ: True

Itzul.: (vpro '|gizaki I 2| ')((ensemble-de |gizabanako I 1|)) 'th-hleh))

ANHITZ: NIL

II.4.4.4 Sorkuntzarakoak.

Sorkuntza da, ezbairik gabe, itzultzailearen zeregin zailenetakoa, are zailagoa gainera ez dagoelako neurriko tresna lagungaririk. Laguntza eman zezaketenak, hiztegiak kasu baterako, ulerkuntzarako daude pentsatuta tradizioz. Al-Kasimik ez du uste, ordea, hala behar duenik (1977:25):

It is time for them [lexicographers] to make dictionaries which help the user in the production as well as the comprehension of foreign speech and writing.

Ez dugu lexikografoen asmamena zalantzan jarri nahi, baina paperezko hiztegiei eragozpenak ikusten dizkiogu sorkuntzarako hiztegi egokiak izan daitezen⁵⁶. Gure ustetan teknologia berrien ekarpenek jauzi kualitatiboa eragingo dute hiztegien sorkuntzarako gaitasunean. ANHITZ sistemako ataza primitibo batzuk, behintzat, hutsune hori betetzeko asmatu dira⁵⁷.

Kontzeptu-bilaketa_tesaurikoa (rths)

Idea bat eduki eta ideia hori adierazteko kontzeptu egokia bururatu ezinik aritzen dira zenbaitetan itzultzaileak. Hiztegi klasikoek, ordea, sarreratik definiorako bidea — ulerkuntza— erakusten dute. Hiztegien eskaintzaren eta itzultzaileen beharren arteko alde hori gainditzeko pentsatu da RTHS.

Ataza honen helburua murriztapen sail bat betetzen duten kontzeptuak lortzea da. Murriztapen sail horren bidez ideiak, edo definizioak nahi bada, adieraz daitezke.

⁵⁶ Batez ere, informazioaren egituraketa eta atzipenarekin zerikusia duten arazoak.

⁵⁷ Sail honetan primitibo morfosintaktiko gehienak sar genitzakeen, izaera gramatikala baitute askotan sorkuntzarako laguntza-eskariak. Azkenean, baina, sail bereizi batean biltzeari egokiago iritzi diogu.

```

task-knowledge
  task#24:      rths
    task-specification
      task-definition
        goal:      murriztapen sail bat betetzen duten kontzeptuak
                   lortzea
        input: tesauruserako_ideia: murriztapen_espresioa
                                     edo definizioa,
                                     hiztegia: hiztegia,
                                     emaitza_hizkuntza: hizkuntza
        output:    kontzeptu_lista: kontzeptua*
      task-body
        type:      primitiboa
        parents:    esanahia_egiaztatzea,
                   sorkuntza_hipotesiak_lortzea
        sub-tasks:  nil
        additional-roles:
                   emaitza_kontzeptuen_aldaia: aldaia

```

Adibideak:

```

Itzul.: (rths '(and (or (|personne I 1| descendant ?X)
                        (|personne I 1| hyponyme ?X)))
         '?X 'th-lppl 'frantsesa)

```

Pertsonak diren kontzeptuak jakin nahi dira frantsesez.

```

ANHITZ: (|vivant I 3| |voisin I 2| |supérieur I 2| |président I 1|
         |patriote I 1| |maître I 4| |maître I 2| |maître I 1|
         |habitant I 1| |femme I 1| |chef I 1|)

```

```

Itzul.: (rths '(and (and (|izaki I 1| descendant ?X)
                        (?X caractéristique |basati I 1|)))
         '?X 'th-hleh 'euskara)

```

Izakia izan eta gainera basatiak diren euskarazko kontzeptuen galdea.

```

ANHITZ: (|piztia I 1|)

```

Adibide-galdea (adib)

Hara zer esaten den (Sinclair, 87:137)n hiztegietaiko adibideei buruz, Gwyneth Fox-en eskutik:

This should help to reinforce the meaning —not by acting as a reformulation of the definition, but showing how the word is actually used in an appropriate context, a typical grammatical structure, and with words that are normally associated with it.

Garbiro erakutsi digu gure azterketak berebiziko garrantzia dutela adibideek —ondo aukeratutakoek!— hiztegien erabileran, batez ere egiaztapenetan. Ondoko zeregin hauetarako, behintzat, baliagarri gertatzen dira:

- Adiera bat testuinguru kultural zehatz batean erakustea, hau da, egiaz existitzen dela frogatzea.
- Definitutako sarreraren esanahia ilustratzea.

- Gramatika- edo estilo-portaerari buruzko argibideak ematea.
- Kokakidetasun-kasu tipikoak erakustea.
- Adiera-desanbiguazioan laguntzea.

Hala ere, Ludmila Minaevak ez du hain erabateko oniritzirik agertzen *Dictionary examples: friends or foes?* deitutako artikuluan. Horrela dio (1992:78):

True, complete sentences used as examples, together with definitions and labels can give a fuller picture of the word's meaning and connotation ... If, however, they are overloaded with irrelevant extralinguistic information, they are hardly illustrating ... Most of the time, sentences containing this kind of information can only baffle the user.

task-knowledge

task#25: adib

task-specification

task-definition

goal: kontzeptu baten erabilera-adibideak erakustea.

input: kontzeptua: kontzeptua,

hiztegia: hiztegia,

emaitza_hizkuntza: hizkuntza

output: erabilera_egiaztagarria*: adibidea*

task-body

type: primitiboa

parents: esanahia_egiaztatzea,

baliokidetza_egiaztapena,

sorkuntza_hipotesia_egiaztatzea

sub-tasks:

nil

Kokakide_lexikalen_galdea (kokak)

Testuingurupeko erabilerak, hitzen kokakidetasunak, konbinatoria lexikalak, eta horiek jorratzeko premiazkoa den corpus-azterketak, astindua eragin diote lexikografiaren munduari azken urteotan. Eta, nola lexikografoek informazio horiek corpus-etatik eskuratu eta hiztegitatzeko, hala nabaritzen dute premia itzultzaileek laguntza hori hiztegietatik jasotzeko. Batzuk eta besteak bat datoz horretan, eta hainbat idazlan izan dira horren lekuko, adibidez (Smadja & McKeown, 91; Heylen *et al.*, 93).

Gure azterketak erakutsi digunez, nonbait izatekotan, sorkuntzan da bereziki nabarmena premia hori. Itzultzaileak eskertzen du gogotik helburu-testuari zuzentasuna eta "jatortasuna" emateko eskaintzen zaizkion laguntzak. Hiztegi orokorretako definizioek edo adibideek, ikusi dugunez, balio dute zenbaitetan. Baina ez beti. Erabilgarriagoa dateke kokakide lexikalei buruzko informazioa era esplizituan edukitzea hiztegietan. Badira, izan, esamolde idiomatikoetan eta kokakideetan berezitutako hiztegiak, non sailkapen- eta atzipen-problemek tratamendu landuago eta zehatzagoa jasotzen duten; baina hauek ere nekez asetzen dute erabiltzailearen beharra.

Hiztegi berezitu horien artean Mel'cuk-ek plazaratutako *Dictionnaire explicatif et combinatoire du français contemporain*-aren erakeran bada eredu interesgarria. Beste gainontzeko hiztegiengan, hala orokorretan nola berezitueta, irizpide sintaktikoak erabiltzen dira kokakidetzari buruzko informazioak antolatzeko. Ez ordea Mel'cuk-enean, non irizpide semantikoak ere maneiatzen diren. Hala, hiztegi horretan, oinarri-sarrerari aplikatu dakizkiokeen *funtzio lexikalen* bidez atzitzen dira kokakideak.

Aurrera baino lehen, finka dezagun nola erabiliko dugun gurean kokakidetzaren lexikalaren kontzeptua. Horretarako Ulrich Heid-en esanetara jo dugu, hark emandako argibideei egoki iritzi diegu eta (Heid, 94:226):

collocations are combinations of exactly two lexemes (of category noun, verb, adjective or adverb), realizing two concepts, where the choice of one of them depends on (or: is restricted by) the other.

Kokakidetzak, horrela, erabilera *tipikoak* islatzen ditu, Mel'cuk-en funtzio lexikalen (Mel'cuk 82; Mel'cuk & Zholkovsky, 88) bidez adieraz daitezkeenak. Hautapen-murriztapenetan ez bezala, kokakidetzaren horiek ezin litezke aurriki irizpide semantiko edo sintaktikoen arabera.

Horra hor bada, ondoko espezifikazioan, *kokakide_lexikalen_galdea* balizko funtzioak zer betekizun lukeen gure ustetan. Balizko esan dugu, izan ere, funtzioaren inplementazioa eginkizun baitago oraindik ANHITZen.

```
task-knowledge
task#43:      kokak
task-specification
task-definition
goal:         kontzeptu baten kokakideak ematea.
input: kontzeptua: kontzeptua,
      funtzio_lexikala: funtzio_lexikala,
      emaitza_hizkuntza: hizkuntza
output:      kontzeptu_lista: kontzeptua*
task-body
type:        primitiboa
parents:     sorkuntza_hipotesiak_lortzea
sub-tasks:   nil
```

Adibideak (eskuz jarriak):

Itzul.: (kokak |maluta I ?| Figur Euskara)
Figur funtzio lexikalak metafora estandarra adierazten du.
ANHITZ: (|luma I ?|)

Itzul.: (kokak |udaletxe I 1| Cap Euskara)
Cap funtzio lexikalak 'buru' esan nahi du.
ANHITZ: (|alkate I 1|)

Itzul.: (kokak |fakultate I ?| Cap Euskara)
ANHITZ: (|dekano I ?|)

Itzul.: (kokak |unibertsitate I ?| Cap Euskara)
ANHITZ: (|errektore I ?|)

Bateragarritasun semantikoaren egiaztapena (konp-sem)

Aurreko atazaren ildo beretik, bi kontzepturen ezaugarriak semantikoki bateragarriak diren edo ez erabaki nahi izaten dute noiz edo noiz itzultzaileek. Helburu horretara zuzendutako ataza da konp - sem deitu dugun hau.

task-knowledge

task#39: konp-sem

task-specification

task-definition

goal: bi kontzeptu semantikoki bateragarriak direnentz
erabakitzea.

input: kontzeptua1, kontzeptua2: kontzeptua,
erlazioa: funtzio_lexikala edo erlazio_sintaktikoa

output: ?bateragarri: boolean

task-body

type: primitiboa

parents: baliokidetza_egiaztapena,
sorkuntza_hipotesia_egiaztatzea

sub-tasks:

nil

II.4.4.5 Morfosintaktikoak.

Portaera morfosintaktikoa lantzea ere itzultzaileen halabeharrezko iharduna dugu. Ez dirudi baina, horretarako ere, itzultzaileak hiztegieta jotzen du maiz, haietan zer aurkituko. Eta hiztegieta informazio morfosintaktikoa izaten da, baina ez beti nahi beste. Gure behaketetako ondorioak Sinclair-en (Sinclair, 87:104) hitzotan daude adierazita:

They [dictionaries] also contain a considerable amount of grammatical information, which unfortunately is usually neither helpful nor well presented.

Ez dugu uste hiztegia gramatikatzat hartu behar denik, horretan bat gatoz Al-Kasimirekin (77:49): "...the dictionary will not be able to provide the desired integrated linguistic description of the lexicon of the language". Baina horrek ez du esan nahi teknika berrien laguntzaz hiztegieta informazio gramatikaren ustiapena hobetu ezin daitekeenik.

Ondoren zerrendatzen diren ataza primitiboak itzultzaileen ihardunetik erauzi ditugunak dira. Oraindik inplementatu edo integratu gabe daude ANHITZen.

Analisi morfologikoa (anali) eta sorkuntza morfologikoa (prod)

Zenbait hizkuntzatan, euskaran urrutira gabe, unitate lexikalen portaera morfologikoak zailtasunak ekar ditzake. Horiek arintzeko hiztegi-sistemarekin batera analizatzaile eta sortzaile automatikoak integratzea dateke egokiena.

task-knowledge

```
task#5:      anali
  task-specification
    task-definition
      goal:      unitate lexikalaren analisi morfologikoak lortzea.
      input: sorburu_unitatea: unitate_lexikala
      output:     analisi_morfologiko_lista:
                                (hiztegi_sarrera,morfologia)*
    task-body
      type:      primitiboa
      parents:    unitate_lexikaletik_hiztegi_sarrerara
      sub-tasks: nil
```

task-knowledge

```
task#49:     prod
  task-specification
    task-definition
      goal:      analisi morfologikotik unitate lexikala lortzea.
      input: analisi_morfologiko: (hiztegi_sarrera,morfologia)
      output:     xede_unitatea: unitate_lexikala
    task-body
      type:      primitiboa
      parents:    hiztegi_sarreratik_unitate_lexikalera
      sub-tasks: nil
```

Lokailu sintaktikoen galdea (pat-sint)

Diskurtso-mailako ataza gisa kontsidera daiteke, perpausak lotzean erabiltzen dena.

task-knowledge

```
task#32:     pat-sint
  task-specification
    task-definition
      goal:      patroi sintaktiko bati egokitzen zaizkion lokailuak
                  lortzea.
      input: patroi_sintaktikoa: patroi_sintaktikoa,
                  emaitza_hizkuntza: hizkuntza
      output:     lokailu_lista: lokailu_sintaktikoa*
    task-body
      type:      primitiboa
      parents:    sorkuntza_hipotesiak_lortzea
      sub-tasks: nil
```

Adibideak (eskuz jarriak):

Itzul.: (pat-sint kontzesiboak Euskara)

ANHITZ: ('nahiz' 'nahiz eta' 'ba... ere' 'arren')

Aditz-erregimenaren_galdea (adi-reg)

Aditzek azpikategorizatzen dituzten kasuak lantzen dira **adi-reg** ataza primitiboan. Itzultzaile batek baino gehiagok nabaritu zuen informazio-mota horren beharra xede-unitatea sortzean.

```
task-knowledge
task#49:      adi-reg
task-specification
task-definition
goal:         aditz      baten      azpikategorizazioari      buruzko
              informazioa ematea.
input: kontzeptua: kontzeptua -- aditza behar du izan
output: portaera_sintaktikoa: azpikategorizazioa
task-body
type:         primitiboa
parents:      hiztegi_sarreratik_unitate_lexikalera
sub-tasks:    nil
```

Itzulpen_lexikala (itzul-lex)

Analizatzaile eta sortzaile morfologikoaren konposizioaz eratortzen den ataza primitiboa da itzul-lex. Ataza honek bere gain hartzen du ANHITZ ereduko ataza gorenaren helburua, hots, unitate lexikalaren itzulpena.

```
task-knowledge
task#50:      itzul-lex
task-specification
task-definition
goal:         sorburu-unitatetik      xede-unitate      baliokidea
              sortzea.
input: sorburu_unitatea: unitate_lexikala
output: xede_unitatea: unitate_lexikala
task-body
type:         primitiboa
parents:      itzulpen_unitatea_itzultzea
sub-tasks:    nil
```

III. Hiztegi-ezagutzaren errepresentazioa.

III. Hiztegi-ezagutzaren errepresentazioa.

ANHITZ sisteman lehen mailako helburutzat hartu da ezagutza-motak bereiztea eta izaeraren arabera modulu desberdinetan biltegitratzea. Hiztegietatik erauzitako ezagutzak modulu horietako bat osatzen du, aurrerantzean ANHITZ Hiztegi-Ezagutza Base, ANHITZ HEB, deituko duguna. KADS metodologiaren zantzuak nabari dira joera honetan, dakigun bezala, KADSek ezagutza lau geruzatan banatzen baitu. Lau geruza horietan lehena domeinu-geruza da, eta ANHITZ sisteman domeinuari buruzkoa HEBan gorpuzten da. Gatozen bada Hiztegi-Ezagutza nola antolatu den deskribatzera.

KADS metodologiari beha jarri gara hasieran —ikus § III.1—, metodologia horren eraginpean erabilitako domeinu-errepresentaziorako eredueta zer ikusi. Hurbilpen orokor eta zabal horietatik xehetzera, gureru egokitzea, jo dugu ondorenean, domeinuaren —ezagutza lexikalaren— ezaugarri teoriko eta praktikoa aztertu eta agerian uzteko —§ III.2—. Horiek gogoan, gure errepresentazio-ereduaren printzipioak xedatu ditugu —ikus § III.3—, ANHITZ HEBaren deskribapenean bete-betean sartu aurretik. Deskribapena, § III.4n ikusgai, ezaugarriak metatuz burutu dugu: hasi, modulu elebakarretatik hasi gara, zeintzuk HIZTSUA sistemaren parekoak diren; gero hiztegiak batera jartzeak eta, bereziki, hizkuntza desberdinekin lan egiteak dakartzan gorabehera heldu diegu; hurrena, itzulpena zuzendutako informazio gehigarria eransteko proposamena hartu dugu idazgai, eta azkenik, inferentzi geruzari heldu diogu, KADS metodologian domeinutik bereiz tratatzen dena, baina gurean hiztegi-ezagutzaren partetzat daukaguna. Deskribapen horren buruan ANHITZ HEBaren egungo egoerari tarte txiki bat eskaini eta datu batzuk eman ditugu. Kapituluaren bukaerarako, § III.5erako, utzi ditugu gordean geneuzkan erreflexio, kezka eta oharpen batzuk.

III.1 Domeinu-ezagutzaren errepresentazioa KADS metodologian.

Atazen geruzan gertatzen den bezalatsu, domeinu-ezagutzaren trataeran ere aldeak daude KADSera bildutako lengoaietan (Fensel & van Harmelen, 94). Aldeak alde, lengoai horien arteko gehienek formalismo deklaratiboatara —frametan edo logikan oinarritutakoetara— jo dute, eta inferentzi gaitasun oinarritzkoari —herentzia edota dedukzio logikoari— eutsi egin nahi izan diote. Horiek dirudite ezaugarri nabarmenenak domeinuaren errepresentazioari dagokionean.

Edonola ere, KADS metodologiaren alderdi ahulenetakoa kontsideratu izan da domeinu-ezagutzaren modelizazioa. Itxuraz, modelizaziorako euskarri egokirik ez

proposatzeak sortarazi eta hedarazi du iritzi hori. G. Schreiber-ek domeinuaren deskribapenerako formalismo batzuk aztertzen ditu (93:71-92)n, KADS metodologian egoki txerta daitekeen domeinuaren modelizaziorako euskarria xedatzeko asmotan. Hiru dira autore horrek ukitzen dituen hurbilpenak, klasikoak hirurak:

- *Entitate-erlazio eredua*: ezaguna da formalismo honek datu-baseetan izan duen zabalkundea, baina era berean da ezaguna ez duela erlazio konplexuak —erregelak edota sailkapen-hierarkiak, kasu— modelizatzeko eraikitzaile aproposik.
- *Entitate-erlazio hedatua*: erlazio-motak ez dira denak izaera berekoak honakoan. Beste batzuen artean, badaude, konparazio batera, *is-a* estiloko erlazioak⁵⁸, zeintzuk propietateen herentziarako aukera ematen duten.

Formalismo hau egokia da, oro har, ezagutza lexikala adierazteko. Hurbiletik ukitzen gaituenez, esan dezagun Euskararen Datu-Base Lexikala —EDBL— formalismo honetan dagoela deskribatua (Agirre *et al.*, 94a)n. Hala ere, aitzakiak bilatzen hasita, ez da erraza eredu honetan beste erlazio batzuk —kausa-ondorioa, denbora, ...— modelizatzea.

- *KL-ONE*: honetan ere *is-a* erlazioak eta gainerakoak bereizi egiten dira. Entitatetik entitaterako erlazioak, edota entitatetik atributurakoak, rolen bidez adierazten dira⁵⁹. Rol horiei murriztapenak ezar lekizkieke eta beraien arteko erlazioak ere deskriba litezke. Gainera, instantzien eta kontzeptuen arteko bereizketa ere egiten da.

Azterketa horren buruan, Schreiber-ek KADS metodologian domeinu-errepresentaziorako lengoaiek eduki beharko lituzketen ezaugarriak aipatzen ditu, eta horiek guztiak *Domeinu-Deskribapenerako Lengoaia* —"Domain Description Language", DDL— biltzen ditu. DDL, azken batean, Entitate-erlazio eta KL-ONE formalismoen bertsio aberastua da, eraikitzaile eta erlazio-mota anitzez hornitua eta deskribapen grafikorako aukera zabala eskaintzen duena. Hori horrela, ahalmen deskriptibo handiko lengoaia dugu DDL.

Schreiber-ek, gainera, DDL deskribapen-lengoaia eta Ezagutzaren Errepresentaziorako inguruneak konparatuz borobiltzen du bere ekarpena. Besteak beste, KEE ingurunearekin alderatzen du DDL, eta oro har dio, KEE-k euskarri egokia eskaintzen duela zenbait DDL domeinu-deskribapen tratatzeko⁶⁰.

⁵⁸ Zenbaitetan bi motako *is-a* erlazioak desberdintzen dira: espezializazioa eta generalizazioa.

⁵⁹ Erlazioak zeharka, rolen bidez, adierazte hori da, hain zuzen ere, formalismo honen mugapenetako bat.

⁶⁰ Horrek ez du esan nahi, ordea, KEEen domeinu guztiak eroso inplementatzen direnik. Schreiber-ek KEEren akatsak ere aipatzen ditu, adibidez (93:90): "A cumbersome aspect of KEE is that it does not make an explicit distinction between concepts and instances: both map onto KEE *units*."

Azken aipamen horren harian, aurrera dezagun KEE ingurunea hobetsi dela ANHITZen Hiztegi-Ezagutza Basearen inplementaziorako. Baina, aurrena, gatozen domeinuaren — hiztegi-ezagutzaren— ezaugarriak aztertzeraz, horietara moldatuko baita KEE gurean.

III.2 Hiztegi-ezagutzaren errepresentazio-ereduak.

Hiztegiak, hitzen eta definizioen gordailu diren neurrian, ezagutza lexikalaren adierazpidetzat hartu ohi dira, ez bakarria baina bai nagusienetakoa gaurdaino. Hiztegian metatzen denak, hiztegi-ezagutzak, errealtatearen partizio sinboliko bat eratzen du, non unitate lexikal-kontzeptualak diren oinarriko. "Unibertso lexikal-kontzeptual" horren egituraketa hamaika azterlanen gaia izan da azken aldion, batez ere lengoaia naturalaren prozesamendurako sistemak lexikoaren garrantziaz jabetu direnetik.

Teoria semantiko klasikoetan hitzak eta hauek adierazitako kontzeptuak dira alde edo moldez lantzen direnak. Hitzei informazio lexikala egokitzen zaie, portaera morfosintaktikoaren berri emateko. Kontzeptuak, bestalde, zentzurik zabalenean hartuta, munduari buruzko ezagutza-unitateak dira, eta ondorioz ezagutza adierazteko egituretan antolatu ohi dira, konparazio batera, hierarkietan, grafoetan, logikan oinarritutakoetan... Eta, azkenik, eredu osatzeko, lexikoek hitzen eta kontzeptuen arteko loturak adierazten dituzte. Horrelatsu moldatzen dira gehienetan "unibertso lexikal-kontzeptualak".

Ereduei beha

Eskema klasiko horrek forma asko hartu ditu noski, eta aldaera horiek aztertzerakoan komeni da errepresentazio-maila desberdinak bereiztea. Hala, (Rappaport *et al.*, 94)n adibidez, bereizketa nabaria egiten dute beren ikuspuntua azaltzean: batetik sintaktikoki nabarmena den errepresentazioa, "Predicate-Argument Structure"en —PAS— bidez adierazitakoa, eta bestetik semantikoki nabarmena dena, "Lexical Conceptual Structure"ak⁶¹ —LCS— baliatuta lantzen dena. Autoreek aurkeztutako teoria lexikalaren motibazioa lexikotik syntaxira dagoen jauzia gainditzeko komenientziatik dator⁶², eta horregatik, maila sintaktikoa eta semantikoa bereizita daude, baina ez isolatuta. Proposatutakoa dotorea da: predikatu-argumentu eredu erabiltzea bi mailetan, bai semantika lexikala, bai syntaxi lexikala errepresentatzeko. Predikatuaren eta atributuen arteko erlazioak semantikoak dira batzuetan eta sintaktikoak besteetan.

⁶¹ Jackendoff-en egitura kontzeptualak daude hauen jatorrian.

⁶² "Government and Binding" (GB) teoritik abiatzen dira eta, horren arabera, erlazio tematikoak —maila semantikoa— leudeke erlazio sintaktikoen sorkuntzaren oinarrian.

Rapaport eta Levin-en lanak aditzen esparrua jorratzen badu ere, Pustejovsky-ren teoria (Pustejovsky, 91) areago doa. Areago, batetik, *"the lexicon is not just verbs"* printzipiotik abiatzen delako, eta bestetik, bi adierazpen-maila baino gehiago bereizten dituelako. Pustejovsky-ren ustetan semantika lexikalak trataera sendo eta formalagoa jaso behar luke, eta ikuspegi hori garatzeko ez bi, baizik eta lau errepresentazio-maila bereizten ditu. Hona hemen segidan mailok:

- Argumentu-Egitura ("Argument Structure"): (Rapaport et al., 94)n adierazitakoarekin bat dator bete-betean. Predikatu-argumentu egitura da, unitate lexikala adierazpen sintaktikotan nola itsasten den jasotzen duena.
- Gertaera-Egitura ("Event Structure"): *egoera*, *prozesua* edo *trantsizioa* bezalako gertaera-motak identifikatzeko erabiltzen da.
- Qualia-Egitura ("Qualia Structure"): item lexikala ezaugarritzen duten atributu nagusiak finkatzen dituena. Qualia-egiturak esanahiaren lau aspektu nabarmentzen ditu:
 - Osagai-rola ("Constitutive Role"): hitzaren eta bere osagaien arteko erlazioa. Adibidez, gaia, pisua, atalak eta osagaiak rol honetan kokatzen diren atributuak dira.
 - Rol Formala ("Formal Role"): domeinu zabal batean objektua bereizgarri egiten duten atributuak, hala nola, orientazioa, magnitudea, forma, neurriak, kolorea eta posizioa.
 - "Telic Role": helburuari eta egitekoari dagokiona.
 - Agente-Rola ("Agentive Role"): jatorriarekin edo eragitearekin zerikusia dutenak. Rol honetan biltzen diren atributuak: sortzailea, tresna, "natural kind" eta kausazko katea.
- Herentzi Egitura ("Inheritance Structure"): hitza lexikoiko beste kontzeptuekin nola erlazionatzen den deskribatzen du.

Pustejovsky-ren semantika lexikalaren teoriak, lexikoaren kontzepzio orokor eta aberatsa erakustez gain, badu zeresana hiztegi-ezagutzaren errepresentazioan. Horixe da (Pustejovsky and Boguraev, 93; 94)n defendatzen dena, hots, teoria horren itzalpean hiztegi-ezagutzaren ezaugarritze egokiagoa egin daitekeela. Azken batean, egitura lexikal aberatsagoak eraikitzeke ildo nagusiak proposatzen dira, predikatu-argumentu egitura, deskonposaketa lexikala⁶³ eta antolakuntza kontzeptualaren nozioak ustiatuz, batik-bat.

Antolakuntza kontzeptualaren nozioa aipatu dugunez, J. F. Sowa-k testuinguru honetan egindako ekarpenei egingo diegu tarte. Sowa, aurretik egindako hainbat hurbilpenen

⁶³ Deskonposaketaren oinarrian Pustejovsky-ren maila lexikaleko sorkuntza-eragileak ("generative devices") daude, zeintzuk primitibo-multzo finkoen ikuspegi klasikoari kontrajartzen zaizkion.

ezaugarriak bilduz, grafo kontzeptualen teoria paratu zuen (Sowa, 84). Grafo kontzeptualek, funtsean, logika-sistema bat osatzen dute, frame-etan bezalako herentzia eta sare-egiturak bere gain hartuz, eta kasu-erlazioak edo rol tematikoak esplizituki adierazteko ahalbideaz. Sowa dio (Sowa, 88; 93), bai egitura lexikalak, bai egitura kontzeptualak, biak adieraz daitezkeela egokiro grafo kontzeptualen bidez⁶⁴. Formalismo hau, Sowa berak ezezik, beste autore batzuek ere erabili izan dute hiztegien ezagutza semantikoa errepresentatzeko, (Pazienza & Velardi, 87; Velardi *et al.*, 88). Autore hauen errepresentazio kontzeptualerako eredia kontzeptuen eta erlazio kontzeptualen grafo zuzendua da. Darabilten lexikoi semantikoak ez du hizkuntzatik independente izan nahi, oso bestalde baizik, adiera bakoitzari azaleko patroia semantikoa ("surface semantic pattern") egokitzen zaio, hartara murriztapen semantikoak eta hitzaren erabilera-informazioa ere adierazteko.

III.2.1 Lexiko eleanitzen egituraketa.

Ezagutza lexikalaren ezaugarriak aztertzen hasi eta zenbait teoria lexikal-semantikotarantz lerratu gara. Egiaz, teoria horien baitako definizio, kontzeptualizazio eta formalizazioak oinarri sendoak gertatu dira lexikoen egituraketari aurre egitean. Ez da harritzekoa, hortaz, LNPrako sistema askoren lexikoek edukitzea, konparazio batera, Jackendoffen "Lexical Conceptual Structures"en zantzuak, edo Pustejovskyren "Generative Lexicon"etik jasotako nozioak, edo Sowaen "Conceptual Graph"en antza.

Lexiko horien egituraketei erreparatuko diegu bada. Horretara goazela, zilegi bekigu lexiko hiztegibakar eta elebakarrez haraindi, lexiko hiztegiak edo eleanitzetara eramatea gure arreta. Izan ere, hastapenetik baitabilkigu gogoan giza itzulpenean laguntzeko sistema hiztegiak eta eleanitza sortzeko ideia.

(Byrd *et al.*, 87)n lexiko konputazional eleanitzen errepresentazioan topatzen diren oinarritzko arazoak planteatzen dira. Hona hemen batzuk:

- Mapping-problema ("The Mapping Problem"): hiztegi bat baino gehiago elkarlotu nahi direnean, hitz berari hiztegi batean eta bestean adiera desberdinak egokitzen zaizkionez, problematikoa gertatzen da adieren arteko mapping erlazioak finkatzea. Ingeleseko Webster's Seventh, Collins Ingelesa-Italiera eta italierazko hiztegi elebakarra erabili zituzten artikuluek mapping fitxategia automatikoki sortzeko, ezaugarri semantikoen komunztaduran oinarritutako estrategia batez. Hala ere, atzenean, *"for highly polysemous and common words... the mappings will have to be done by hand"* aitortzen dute.

⁶⁴ Gainera grafoei ohar sintaktikoak ere gehitzen dizkie Sowa, eta benetan formalismo interesgarria definitzen du lengoia naturalaren prozesamendurako.

Iritzi hori indartzera dator Judith L. Klavans-en esaldi hau (90:270):

We have examined mapping between LDOCE, W7, NEW COLLINS THESAURUS, and ROGET'S THESAURUS, and have found it extremely difficult (if not impossible) to map a given sense in one dictionary into a single sense in another.

- Addenda problema ("The Addenda Problem"): sistema lexikal hiztegianitzetan hiztegi baten informazioari bestee emandakoa eransten zaio. Oinarri gisako hiztegi bat hartuta, bi bide leudeke informazio-osatze hori egiteko: oinarri-hiztegiari berari informazioa gehituz, edota aparteko "addenda" fitxategi batean informazio erantsia gordez —oinarri-fitxategian aldaketarik egin gabe—.
- Hiztegi elebidunetan zeharreko informazio lexikalaren transferentzia: transferentzia zein mailatan gertatzen den jakiteko hiru saio aurkezten dira (Byrd *et al.*, 87)n: bata [+human] marka semantikoarekin, bestea sinonimia erlazioarekin eta hirugarrena [+thatcomp] ezaugarri sintaktikoarekin. Badirudi, emaitzen arabera, marka semantikoa propagatzen dela —salbuespenak albuespen—, baina beste bi kasuetan askoz ere lausoagoak dira emaitzak.
- Erreferentzia gurutzatuak eta adiera-desanbiguatzea hiztegi elebidunetan: adiera-desanbiguazioaren arazoa latza bada hiztegi elebakarretan, are eta latzagoa da hiztegi elebidunetan, azken hauetan informazio gutxiago egon ohi delako desanbiguaziorako.

Bestalde, erreferentzia gurutzatuak direla-eta, bi zentzutako hiztegi elebidunen erabilerak asimetria sortaraz dezake. Hiztegi elebidun "simetrikotuen" komenientziaz, ordea, zalantzan daude autoreak.

Transferentzi saio hauek ikuspegi kritikoz aztertu eta hiztegi elebidunen simetrikotasunaz duen iritzia azaltzen du Bernard Al-ek (1990)en. Hitz bitan esanda, ez omen dira simetrikotzat hartu behar hiztegi elebidunak, eta horretan sinestuta, ondoko mezu hau helarazten digu informatikarior:

la solution idéale de l'informaticien —solution qui consisterait à n'avoir pour chaque paire de langues que deux dictionnaires, dont l'un aurait été construit en inversant l'autre— est loin d'être parfaite. En réalité ce qu'il nous faut, c'est quatre dictionnaires pour chaque paire de langue,...

Zailtasun horiei aurre eginez garatu zuen Pisako taldeak Database Lexikal Elebidunaren diseinu-proiektua (Calzolari & Picci, 86). Azken sistema lortzeko database elebakarrak hiztegi elebidunen bitartez lotzea du helburu proiektu honek. Integraziorako erabili diren hiztegiak aski ezagunak dira: Zingarelli italierazko hiztegia, Collins Italiara-Ingelesa elebiduna, Longman Dictionary of Contemporary English —LDOCE— eta Oxford Advanced Learner's Dictionary of Current English. Pisako proiektu honetan aintzat hartzen dira Bernard Al-en gomendioak (Picci *et al.*, 90): hiztegi elebidun bakoitza bi datu-multzoz osatuta dago, denetara lau hiztegi-bertsio osatuz. Mapping problema dela-eta, adierak

egokitzeo nabariak omen dira hiztegi elebidunen informazio-gabeziak. Lehen emandako iritziak berretsiz, eskuzko langintza beharrezkoa ikusten dute autore hauek ere lotura elebidunak sortzeo prozesuan.

ACQUILEX proiektuan ezagutza-base lexikalaren —LKB— diseinu eta garapena landu da (Copestake, 92). Erabilitako errepresentazio-formalismoaren oinarrian Carpenter-en (1990) ereduo mota-hierarkiaren nozioa dago. Mota bakoitzak hierarkian barrena heredatuko den ezaugarri-egitura bat dauka⁶⁵. Sarrera lexikalei mota bana egokitzen zaie, eta motaren ezaugarriek sarrerari buruzko ezagutza egituratzeko balio dute. Ezaugarrien balioak osatzeko hiztegi-definizioaren analisitik ateratako edo, besterik ezean, heredatutako informazioa baliatzen da. Azken herentzia lehenetsi hau —ez motena, baizik eta balioena—, beste hierarkia batean zehar gertatzen da, hots, hiztegitik eratorritako taxonomiak errepresentatzeko hierarkian, zeina "psort"-hierarkia deitzen den. Gainera, eratorpen morfologikoa edo zentzu-hedapena bezalako fenomenoak kodetzeko erregela lexikalak erabiltzen dira.

LKBren errepresentazio-lengoaia —LRL— informazio eleanitza adierazteko ere erabili da ACQUILEX proiektuan (Copestake & Sanfilippo, 93; Samiotou *et al.*, 93; Copestake *et al.*, 94). Hurbilpen honetan, dimentsio eleanitzeko informazioa kodetzeko sarrera lexikal elebakarren arteko loturak baliatzen dira⁶⁶. Itzulpen-lotura hauen ezaugarriak behinenak hauexek dira:

- Itzulpen-baliokidetza lexikalak adierazten dituzte, eta "translation link" direnez "tlink" deitzen zaie.
- tlink mekanismoak ez du errepresentazio-lengoiaren hedapenik erakartzen: tlink-ak tlink motako ezaugarri-egiturak dira, eta bi erregela lexikalen arteko erlazioa adierazten dute.
- Mekanismo hau baterakuntzan oinarritutakoa da, deklaratioa eta adierazkortasun handikoa.
- Itzulpen automatikoko transferentzi erregela ohikoak baino zabalagoak direnez, baliagarriak gerta daitezke giza erabiltzenerako.
- tlink mota desberdinak defini daitezke baliokidetzarekin zerikusia duten fenomeno desberdinak —generalizazioak, murriztapen edo espezializazioak, "hutsune" lexikalak— egoki adierazteko (Martí & Soler, 94).

⁶⁵ Egitura horiek, "Relativised Qualia Structure" deituak, Pustejovskyren "Qualia Structure"etan daude oinarrituta.

⁶⁶ Itzulpen-loturak erdiautomatikoki sortzeo ingurua —TGE— ere atondu da proiektu honetan (Ageno *et al.*, 94).

Kasurik sinpleenak, baliokidetza erabatekoak, "simple tlink"en bidez adierazten dira. Bestelakoetan, sorburu- eta xede-sarrera lexikalen artean bat etortze erabatekorik ez dagoenean, "partial tlink"ak edo "phrasal tlink"ak sortzen dira. Hara hemen tlink-en sailkapen xeheagoa:

- simple-tlink: zuzeneko korrespondentzia dagoenean sorburu- eta xede-sarreraren artean. Hauen artean nabarmentzekoak dira "compound tlink"ak: hauek ere sinpleak, baina sorburu-hitzari hitz bat baino gehiagoko baliokidea —LKBko sarrera sinplea, hala ere— egokitzen zaionean sortzen direnak.
- phrasal-tlink: baliokidea esaldia edo sintagma denean —ez LKBko sarrera sinplea— eta osoki analizatu denean.
- partial-tlink: "hutsune" lexikalak daudenean —hiztegi elebidunean ez dago sorburu-sarreraren baliokiderik— edota itzulpen-esaldia erabat analizatu ez denean. Hutsune lexikalen kasuan bi eratako tlink-ak sor daitezke: "parent-tlink"ak —taxonomiako gurasoak badu korrespondentzia— edo "grandparent-tlink"ak —gurasoak ez du baliokiderik baina gurasoaren gurasoak bai—.

Urrutiagotik, Japoniatik, "Japan Electronic Dictionary Research Institute, Ltd. (EDR)"k zuzentzen duen proiektu handi baten oihartzunak datozkigu (EDR, 93). Proiektu honen xede nagusia EDR hiztegi elektronikoa garatzea da, hainbat LNPrako aplikazioen euskarri lexikal izateko bokazioa duen hiztegi erraldoia —oraingoz Japoniera eta Ingelesa hartuz—.

Hiru printzipio nagusik gidatu dute EDR hiztegiaren diseinua:

- Azaleko mailako informazioa —morfologikoa eta sintaktikoa— daukaten hiztegiek eta sakoneko mailakoa —semantikoa— daukatenean elkarrengandik banatuta egon behar dute. Azaleko maila, hizkuntzaren menpekoea, Hitz-Hiztegian⁶⁷ biltegitratzen da; informazio semantikoa, ordea, Kontzeptu-Hiztegian. Horrela jokatzeko sakoneko mailan dagoen informazioaren berrerabilgarritasuna bermatu nahi da: hizkuntza desberdinetako hiztegiek izan dezatela informazio semantikoa erabiltzeko aukera. Azken batean, interlingua kontzeptzioa dago horren atzean.
- Sarrera lexikalen eta kontzeptuen arteko korrespondentzia Hitz-Hiztegian finkatzen da. Hitz-Hiztegiek zubi-lana egiten dute hartara, azaleko eta sakoneko hegi ondo bereizien artean.
- Formalismo gramatikal edo algoritmo zehatzei buruzko informaziorik ez da biltegitratzen EDRn. Hiztegi-sistemaren unibertsaltasuna babesten da era horretan.

⁶⁷ Erredundantea ere irudi lezake "Word Dictionary" euskaratzeko aukeratu dugun termino horrek, baina ...

Printzipio horiei eutsiz egituratzen da EDR hiztegia, osagai gisa lau hiztegi mota desberdin dauzkalarik: Hitz-Hiztegia, Kontzeptu-Hiztegia, Kokakidetasun-Hiztegia eta Hiztegi Elebiduna.

Gatozen hiztegi hauen zer-nolakoei begiratu bat ematera:

- Hitz-Hiztegia: lehen ere esana dugu azaleko mailako informazioa eta kontzeptuekiko loturak hiztegi hauetan gordetzen direla. Hala bada, sarreraren portaera gramatikalari buruzko informazioa, egokitutako kontzeptuen identifikadoreak eta erabilerari buruzko informazio gehigarria aurkitzen dira nagusiki hiztegi hauetan, nola japonierazkoan, hala ingelesezkoan (Koizumi *et al.*, 94).
- Kontzeptu-Hiztegia: kontzeptu bakoitzak identifikadore bat dauka, eta identifikadorearekin batera kontzeptu-ilustrazioa, non esanahiaren berri ematen den lengoaia naturalez, kontzeptu-sailkapena —erlazio hierarkikoak gordetzeko— eta kontzeptuaren deskribapen erlazonala —beste kontzeptuekiko erlazio ez-hierarkikoak errepresentatzeko—. Kontzeptu-sailkapenak eta deskribapen erlazonalak bereiz, hiztegi banatan, biltegitratzen dira.
- Kokakidetasun-Hiztegia: hiztegi mota honetako sarrera-hitz bakoitzari hiru informazio-eremu egokitzen zaizkio: kokakidetasun-erlazioaren etiketa, hitz kokakideari buruzko informazioa —gramatikala eta semantikoa— eta, azkenik, ziurtasun-faktorea —zenbakizko balioa probabilitatearen berri emateko—. Nabarmentzekoa da hiztegi honek betetzen duen rola esaldien sorkuntzan: sintaxi-sorkuntza ezaugarri gramatikalak erabiliz egiten den bezala, kontzeptu jakin bat adierazteko hitz egokiaren aukeraketan Kokakidetasun-Hiztegia erabiltzen da.
- Hiztegi Elebiduna: mota honetako hiztegiengan sorburu-hizkuntzako sarrerarekin batera, horri dihoakion xede-hizkuntzako hitzaren informazioa doa. Xede-hitzari buruzko informazioa hiru eremutan egituratzen da, hara nola:
 - Korrespondentzi hitza: hitzak edo string-ak ager daitezke eremu honetako balio gisa.
 - Hitz-korrespondentziaren erlazioa: erlazio hauek sorburu- eta xede-hitzaren arteko korrespondentzia deskribatzen dute. Bost erlazio klase daude definituta:
 - Baliokidetza-erlazioa: sorburu-hitzak eta xede-hitzak kontzeptu bera errepresentatzen dute⁶⁸.

⁶⁸ Izatez, erlazio hau kontzeptutik kontzepturakoa da, eta beraz nahasgarria gerta daiteke hitzen mailan deskribatuta egotea, izan ere, sorburu- eta xede-hitzek errepresentatzen dituzten kontzeptu guztiak ez baitira elkarren artean baliokide.

- Sub-erlazioa: xede-hitzak errepresentatzen duen kontzeptua sorburu-hitzak errepresentatzen duena baino "estuagoa" da.
 - Super-erlazioa: xede-hitzak errepresentatzen duen kontzeptua sorburu-hitzak errepresentatzen duena baino "zabalagoa" da.
 - Sinonimo-erlazioa: sorburu- eta xede-hitzak errepresentatutako kontzeptuak amiñi bat diferenteak dira.
 - Iruzkina ("remark"): korrespondentzi hitza sorburu-hitzaren transliterazioa da⁶⁹. Xede-hitzak zein kontzeptu errepresentatzen duen jakiteko azalpen lagungarriak behar izaten dira.
- Azalpen lagungarriak: baliokidetza erabatekoa ez denean azalpen lagungarriek erabilera egokiari buruzko xehetasunak ematen dituzte.

Multilex ESPRIT proiektuko lanen jarraipena dugu GETA laborategietan diseinatutako NADIA, datu-base lexikal eleanitzetarako kudeatze-sistema (Sérasset, 94). Proposatzen den antolaketan bi hiztegi mota desberdintzen dira: batetik, hizkuntzetako hiztegi elebakarrak, eta bestetik, interlingua hiztegiaren rola betetzen duen adiera-hiztegia. Hiztegi elebakarrak sarreren bidez dira atzigarriak eta adiera elebakarrak dituzte oinarrizko item. Adiera-hiztegian, interlingua adierak eta adiera horien arteko berdintasun erlazioak —baliokidetasun semantikoa adierazteko— biltegitratzen dira. Baliokidetza guztizkoa ez denean erlazio murritzaileak erabiltzen dira, bidenabar azpiadierei sarrera emanez. Bestalde, interlingua adierak eta elebakarrak 1:1 korrespondentziaz lotzen dira, eta korrespondentzia hori baliokidetasun erlazioarekin konposatuz zertzen dira itzulpen erlazioak⁷⁰.

III.2.2 Itzulpenerako lexikoak.

... selecting appropriate translations of source lexical items has been one of the hardest problems in MT.

(Tsujii & Fujita, 89:275)tik ekarritako esaldi horri kasu egiten badiogu, pentsatzekoa da itzulpen automatikorako sistemen lexikoek izango dutela itzulpen lexikalera zuzendutako informazio espezifikoa, agian, Datu- edo Ezagutza-Base Lexikal helburu orokorrekoek ez dutena. Hori horrela balitz, eta esan dezagun ja horrela dena, komenigarria litzateke ikustea sistema horien lexikotik zerk egon beharko lukeen ANHITZ sistemaren HEBan, makinari ez

⁶⁹ Gogoan izan Japoniera-Ingelesa edota Ingelesa-Japoniera hiztegiez ari garela.

⁷⁰ Antolaketa horrekin esperimentatzeko 5 hizkuntzako —ingelesa, frantsesa, alemanera, errusiera, txinera— datu-base lexikala inplementatu da, hizkuntza bakoitzeko 500 adiera inguru eta interlingua hiztegian ia 600 adiera errepresentatuz.

baina gizakiari itzultzen laguntzeko. Asmo horrek eraginda doaz ondoko printzak, handik eta hemendik jasotako ideiak baizik ez direnak.

Zertzeladetan sartu gabe, ikus dezagun gaingiroki itzulpen automatikorako unitate lexikaletan zer informazio espezifikatzen den. Baliagarria iruditu zaigu horretarako (Bläser *et al.*, 92)n egiten den deskribapena⁷¹, nahiz eta sistema batetik bestera aldeak dauden noski. Ondoko hauek dira unitate lexikalen espezifikazioan aurkitzen diren informazio motak:

- kategoria gramatikala
- hitz-adierak
- ezaugarri morfologikoak
- komunztadura ezaugarriak
- balentzia, hau da, konplemento derrigorrezko/aukerazkoen framea
- konplemento bakoitzaren betegarrien ("filler") espezifikazioa
- bateragarritasun semantikoari buruzko murriztapenak eta kokakidetasunak
- hitz anitzeko terminoen ezaugarritzea
- erabileremua
- itzulpen-erlazioak
- transformazio lexikalak

Bonnie Jean Dorr-ek bere "Machine Translation: A View from the Lexicon" liburu mardulean (1993) itzulpen automatikorako hurbilpen baten berri ematen digu —UNITRAN⁷² sisteman inplementatutakoa— eta hurbilpen horren harian *dibergentzia*⁷³ fenomeno sailkapen interesgarria burutzen du. Autoreak dio: "*divergence arises when the natural translation of one language into another results in a very different form than that of the original*". UNITRANek dibergentzia sintaktikoak eta lexikal-semantikoak lantzen ditu, interesgarriak denak, baina guk hemen lexikal-semantikoei bakarrik egingo diegu tarte, horiek baitira ANHITZek hurbilgotik ukitu beharko dituenak. Ikus dezagun zertan den dibergentzia lexikal-semantikoen sailkapena —ikus baita ere (Dorr, 94)—:

- "Conflational": itzulpenak beste osagai lexikalak —argumentuak izan litezke— gehitzea, edota, alderantziz, kentzea, eskatzen duenean. Adb.: *I stabbed John* ≈ *Yo le di puñaladas a Juan*⁷⁴.

⁷¹ LMT itzulpen automatikorako sisteman erabiltzen den LOLA datu-base lexikala deskribatzen da artikuluko horretan.

⁷² Sistema hau Lexikoan Oinarritutako Itzulpen Automatikotik ("Lexical-Based MT") hurbil legoke autorearen esanetan.

⁷³ Termino hau beste hurbilpen batzuetan erabilitako "complex lexical transfer" horren antzekoa da.

⁷⁴ Adibideak ez dira gure golkotik atera, autorearenetik baizik.

- "Structural": aditzaren objektua izen-sintagma denean hizkuntza batean eta sintagma preposizionala bestean.
- "Thematic": Argumentuen birkokapena gertatzen denean, adibidez, tema aditzaren objektu izatetik subjektu izatera pasatzen denean, edo alderantziz.
- "Categorical": kategoria-aldaketa dagoenean.
- "Demotional" edo "Promotional": bi mota hauek oso antzekoak dira, batean nahiz bestean, esaldiko gune rolaren trukaketa gertatzen da osagai lexikalen artean. Bi kasuetan aditzari egokitzen zaion itzulpena egitura adberbiala da. Adb.: John *usually* goes home $\approx$ Juan *suele* ir a casa.
- "Lexical": unitate lexikal batentzat itzulpen literala ez baina desberdin bat erabili behar denean. Beste dibergentziekin batera gertatu ohi dira, haiek eraginda, eta horregatik albo-ondorio gisa ere hartzen dira.

Dibergentzia hauek egoki tratatzeko UNITRANen erabiltzen den sistema lexikal semantikoa Jackendoff-en Lexical Conceptual Structure-ren hedadura da, parametrizazio-aukeraz hornitua (Dorr, 91)⁷⁵. Errepresentazio-eredu horretan oinarri hartuta bideratzen da *aukeraketa lexikala* ("lexical selection"), ANHITZ sistemaren eredu kontzeptualeko goi-atazarekin zerikusi nabaria duena.

Galdera da, puntu honetara ailegatuta, ea giza itzulpen lexikalera zuzendutako sistema batek, ANHITZek kasu baterako, ez lituzkeen kontuan hartu beharko gisa horretako dibergentziak; eta baietz iritziz gero, ea nola adieraz daitezkeen horiek hiztegi-ezagutzan, itzulpen automatikorako sistemetan adinako joritasunik gabe noski, baina probetxua ateratzeko moduan bestalde. Airean utziko dugu galdera hau momentuz, baina duen garrantziaz ohartuta.

Ezagutzan Oinarritutako itzulpen automatikoan —ikusi adibidez (Gates *et al.*, 89) KBMT-89 lexikoen berri jasotzeko eta (Nirenburg & Defrise, 93) DIONYSUS sistemaren lexikoak zertan diren jakiteko— egitura lexikal-kontzeptuala modulu bat baino gehiagoz osatzen da. KBMT-89n analisirako eta sorkuntzarako lexikoak bi ziren, baina sarrera lexikalak hain parekoak direnez lexiko bakarrean bildu dira DIONYSUS sisteman. Lexiko horrekin batera, sistemok daukaten bigarren modulua ontologia edo munduaren eredu⁷⁶ da .

⁷⁵ Esan behar da alabaina, Lori Levin-ek eta Sergei Nirenburg-ek (1994), interlingua errepresentazioaz ari direla, "Are Lexical-Conceptual Structures Language Universal?" galdera potoloa plazaratzen dutela, auzitan jarritz, besteak beste, formalismo horrek dibergentzia mota guztiak adierazteko leukakeen gaitasuna.

⁷⁶ Testuinguru honetan, munduaren eredu gero eta zabalagoak sortzera zuzendutako proiektuak garatu dira. PANGLOSS ontologia, adibidez, proiektu horietako baten emaitza dugu. Ontologia hori osatzeko iturri desberdinak erabili dira: PENMAN goi-mailako kontzeptu-sarea, ONTOS goi-mailako ontologia, LDOCE, WordNet eta Collins Spanish-English Dictionary. Iturri horiek denak batera jartzeko eman diren urratsak (Knight & Luk, 94)n deskribatzen dira.

Hori horrela, lexikoetan bi mailako informazioa bereiz daiteke: batetik, ohikoa, informazio ortografikoa, morfologikoa, sintaktikoa, semantikoa eta pragmatikoa; bestetik, mapping informazioa, hots, hitz-adiera eta ontologiako elementuen arteko korrespondentziak adieraztekoa. Berriro ere ezagutza lexikala eta kontzeptuala ondo bereizita azaltzen zaizkigu: lexikala hizkuntzari atxikia, ontologia hizkuntzatik independentea. Aplikazio eleanitzetarako hurbilpen egokia omen da horrako hori. Ikusiko dugu geroxeago ANHITZ sistemarako hobetsi den hurbilpena.

Ontologia euskarri gisa hartuta, hizkuntzetako lexikoak ontologiako kontzeptuekin parekatzean datza, ikusi dugunez, Ezagutza Oinarritutako sistemen egituraketa lexikal-kontzeptuala. Baina, horietan ez bezala, hizkuntzatik hizkuntzarako parekatze lexikalak hobesten dira beste zenbait hurbilpenetan, eskuarki transferentzian oinarritutako eredua darabilten itzulpen automatikorako sistemetan.

Shake-and-Bake hurbilpenean baterakuntza oinarritutako gramatika lexikalizatuen ekarpenak ustiatzen dira (Beaven, 92; Whitelock, 92). Lexiko elebazarretako sarrerak zeinuak dira, zeinu oso aberatsak, ezaugarri askoz hornituak. Bestalde, lexiko elebiduneko sarrera bakoitza, funtsean, erakusle-bikotea da eta bi sarrera elebazar erakusten ditu, bata sorburu-hizkuntzakoa, bestea xede-hizkuntzakoa. Baina, horrez gain, sarrera elebidunek ezaugarri edo murriztapen gehiago eduki dezakete itzulpen-erlazioen xehetasunak esplizituki adierazteko. Egituraketa horren emaitza lexiko aberats eta modulartasun handikoa da⁷⁷.

METAL[®] sisteman ere, transferentzi ereduari jarraikiz, bereizketa argia egiten da hiztegi elebazar eta elebidunen artean (Hutchins & Somers, 92: 259-278). Hiztegi elebazarren diseinuak, ezaugarri morfologiko, sintaktiko eta semantiko hornitzen denak, edozein hizkuntzatarako baliagarri izan nahi du; aldiz, hiztegi elebidunak sorburu- eta xede-hizkuntza jakinetarako antolatzen dira, parekatzeak noiz eta nola gertatzen diren xeheki adierazi nahiak behartuta. (Fontenelle *et al.*, 94)n unitate lexikal konplexuak METAL[®] sisteman nola xedatu eta errepresentatu diren deskribatzen da. Esaera idiomatikoak, unitate konposatuak eta kokakidetzak dira unitate lexikal konplexutzat hartzen direnak proposamen horretan, non, gainera, errepresentazioari buruzko xehetasunak ere ematen diren. Guri dagokigunerako iradokitzailea da esaerak, egitura konposatuak eta kokakidetzak unitate gisa tratatzea: ANHITZen halatsu ikusiko ditugu.

ANHITZ sistemaren HEBa aberasteko beste ideia interesgarria (Kawasaki *et al.*, 92)tik har daiteke. HICATS/EJ itzulpen automatikorako sisteman patroiz mailako transferentzi baliokideak erabiltzen dira, maila lexikalekoen gain. TEPT ("Translationally Equivalent Pattern Triplet") hirukotetan errepresentatuta hiru motako patroiak darabiltzate :

⁷⁷ Bideanabar, esan dezagun ACQUILEX LKBak antz handia duela egituraketa horrekin zenbait aspektutan.

sintagmatikoak ("phrasal pattern"), klausalak ("clausal pattern") eta esaera-patroiak ("sentential pattern"). Patroi hauei esker erosoago maneiatzen dira, besteak beste, esaera idiomatikoak edo ez-estandarrek edota trataera berezia eskatzen dutenak. Gurean patroia sintaktikoen komenientzia garbiro ikusi dugu.

III.3 ANHITZ Hiztegi-Ezagutza Baserako aurremugapen eta printzipioak.

Bi azterketatik jasotako ideiek gidatu dute ANHITZ HEBaren diseinua. Idei iturri bat oinarritzeko atazen espezifikazioa da, izan ere, HEBa funtzionalitate horren informazio-euskarria denez, ataza-deskribapenak erakusten baitu, beste ezerk baino hobeto, HEBek nolako informazio motak gorde behar dituen, ez noski nola antolatu behar diren, baina bai zeren gordailu —eta emaila— izan behar duen. Bigarren azterketak beste sistemen nondik norakoak ezagutzea izan du helburu, eta laburpena behinik-behin oraindik orain aurkeztua dugu —§ III.2—. Esan gabe doa beste sistemetatik ikasitakoak ideia asko iradoki dizkigula ANHITZ HEBaren antolaketaarako.

Bi arlo horiek jorratuta finkatu ditugun HEBaren errepresentazio-printzipioak ondoko puntu hauetan bildu ditugu:

a) *Errepresentazio-eredu erlazionala.*

Unitate lexikal-kontzeptualen esanahia definitzeko hiru hurbilpen bereizten ditu Martha W. Evens-ek (1988). Lehenengoan, kontzeptuek eremu semantikoak edo domeinuak konfiguratzeko dituzte, eta izaera dimentsio bakarrekoa dute: domeinu jakinetako osagaiak dira, ez beste. Horrenbestez, eremuan duen posizioak definitzen du terminoa. Bigarren hurbilpena ezaugarrien analisisian oinarritzen da: ezaugarrien analisiak ematen du aukera domeinu berekoak diren itemak ezaugarritzeko eta, hala berean, domeinu desberdinetakoak bereizteko. Bi hurbilpen hauek egokiak dira domeinu lexikal mugatuetarako, baina, baliagarri izanagatik, motz geratzen dira konplexuagoetarako. Hirugarren hurbilpenak, eredu erlazionalak, domeinu ondo zedarrituetatik areago joanez, eremu desberdinetako kontzeptuen arteko erlazioak era esplizituan deskribatzen ditu. Hitzek arteko erlazioei erlazio lexikal deritze, kontzeptuen arteko berriz erlazio semantiko edo kontzeptual.

Eredu erlazional klasikoak⁷⁸ jarraitzaile asko izan ditu hiztegi-ezagutzaren errepresentazioan, askoren ustetan adierazpide aproposa baita hiztegi-

⁷⁸ Izan kontuan lexikoiaren eredu erlazionalak ari garela, eta ez datu-baseen munduan hain ezaguna den ereduaz.

definizioetarako: (Amsler, 80; Ahlswede *et al.*, 86; Calzolari, 90; van der Vliet, 94) adibide bakanak baino ez dira beste askoren artean.

ANHITZ HEBaren diseinuan ere hurbilpen horretara lerratu gara. Xabier Artolak (1993:65-73)n hiztegi-definizioetatik erauzitako erlazio semantikoak egoki sailkatuta erakusten dizkigu. Erlazio multzo horretan badira batzuk usuak eta esanguratsuak izateagatik aipagarriak. Hara hemen aipatuak:

- Sailkapen-erlazioak edo taxonomikoak: sailkapena omen da errepresentazio-formarik oinarritzkoena, eta hala bide da hiztegi-ezagutzaren arloan ere, aspalditik dator-eta hiztegi-kontzeptuak hierarkia kontzeptualetan paratzeko joera, *genus et differentia* definizio-eredu klasikoaren isla dena seguruenik. Hierarkia kontzeptual hori *hiperonimia/hiponimia* erlazioetan datza nagusiki (Cruse, 86), baina badira horiez gain beste erlazio hierarki eratzaileak, *meronimikoak* esate baterako (Winston *et al.*, 87).

Quillian-i (1968) zor zaion "*type-token*" erlazioak aparteko garrantzia du ANHITZ HEBaren diseinuan. "type node"ek, mota-nodoek, testuingururik gabeko kontzeptua adierazten dute eta "token"ek mota-kontzeptuen hiztegi-definizioetako agerpenak errepresentatzen dituzte. Erlazio horiek izaera taxonomikoa hartzen dute gure errepresentazio-proposamenean.

- Beste kontzeptuarteko erlazioak: hierarkiarik ezartzen ez duten erlazioen artean *sinonimia* eta *antonimia* bezalakoak, edota erlazio sintagmatikoak — *eratorpena*, *kasu-* edo *tema-erlazioak*— dira nabarmenenak. (White, 88)n *kokakidetasun-erlazio* deritze honako hauei⁷⁹.

Esan dezagun erlazio taxonomikoen eta sinonimikoen garrantzia apartekoa dela. Horren adibiderik argiena Wordnet (Beckwith *et al.*, 91) sarea da: printzipio psikolinguistikoetan oinarrituta erlazio lexikal horien arabera dago egituratua sare hori.

- Metalengoaiatik erauz daitezkeen erlazio bereziak: erlazio berezitu hauek hiztegi-definizioetako osakeran maiz erabiltzen dira, normalean definizioaren gune sintaktikoa eta semantikoa bat ez datozenean. (Vossen *et al.*, 89)n bi motako erlature bereziak aipatzen dira: estekatzaileak ("linkers") eta desbideratzaileak ("shunters"), lehenak kategoria bereko kontzeptuen artekoak, eta bigarrenak kategoria desberdinen artekoak⁸⁰.

⁷⁹ Autore honek hiru erlazio lexikal-semantiko bereizten ditu: ISA erlazioa, hiperonimiaren parekoa; kokakidetasun-erlazioa, definizio berean ageri diren hitzen artekoa, eta itzulpen-erlazioa, hizkuntza desberdinetako baliokideen artekoa.

⁸⁰ "linker"ak lirateke, adibidez, 'part of', 'group of' edo 'piece of' bezalako loturak, eta "shunter"ak 'act of' edo 'quality of' bezalakoak.

b) *Informazioaren tipifikazioa.*

Kudeaketa aberats eta antolamendu txukuna helburu, beharrezko ikusi dugu informazioa maila, sail eta motatan eratzea. Asmo horrek hiztegi-informazioaren gainetik legokeen egituraketa erakartzen du. Egiturei buruzko egitura litzateke, hots, informazio lexikal-semantikoaren ezaugarriak xehetzeari zuzendutako "meta-egitura", non unitate lexikalen izaera desberdineko atributuak eta hauen propietateak esplizituki adieraziko diren.

Motaketaren xehe-larria ere hausnargai izan dugu. Eredu batzuetan motaketa xehetuaren aldeko jarrera da nabari: ACQUILEXen errepresentazio-lengoaian —LRL—, esate baterako, mota-hierarkia herentziaduna eratu dute unitate lexikalen ezaugarriei buruzkoak adierazteko. Baina, egia esan, ANHITZ HEBan ez dugu hain fin jokatu tipifikazio kontuetan, geroxeago azalduko dugunez.

c) *Maila lexikala vs. kontzeptuala.*

ANHITZ HEBan maila lexikala hiztegi-sarrerekin —lemekin— lotzen da, eta maila kontzeptuala adierekin. Batetik bestera iragatea da, ezbairik gabe, HEBaren sorkuntza eta ustiapenean topatzen den arazo latzenetakoa.

d) *Hiztegianiztasuna.*

Lehen ere ikusia dugu sistema hiztegianitzek badutela berezko problematika, sistema hiztegibakarretan topatzen ez dena. Gorabehera horietara hurbiltzeko lehen pausoa honako galdera honi erantzutea da: "Zer-nolako harremana edo lotura eduki nahi da hiztegien artean?". Hiru aukera arrunt eta orokor aipatuko ditugu hemen:

- Hiztegi independenteak: aldi berean hiztegi bat baino gehiago atzigarri izateak ematen die sistema hauei hiztegianiztasuna. Gainerakoan hiztegien bildura kontsidera daitezke eta, beraz, hauetan bai ez dela lotura arazorik sortzen. Erabilpenaren ikuspegitik aukera urrikoak dira.
- Hiztegi autonomoak: hiztegi bakoitzak mundu kontzeptual propioa dauka, nahiz eta mundu horietako unitateen artean mapaketa loturak sortzen diren elkar osatzeko eta aberasteko. Egituraketa amankomuna izan lezakete, baina adiera-banaketa desberdina dutenez mapaketa arras zaila gertatzen da.
- Hiztegi fusionatuak: hiztegi desberdinekin bat egin nahi izaten da, sare edo mundu kontzeptual bakarra eta hiztegietatik independentea sortuz.

ANHITZ HEBan bigarren hurbilpenaren aldeko apustua egin nahi izan dugu: hiztegi bakoitzari nortasun propioa duen sare erlazionale egokitu diogu. Sare horietako unitateak baliokidetasun-erlazioaren bidez lotzen dira eta "meta-egitura" mailan,

atributuak, propietateak, etab.ak konpartitu egiten dituzte, nolabait batasuna bermatuz. Gain-geruza horren pean dauden hiztegiek, hala ere, ez dute zertan izaera berekoak izan⁸¹.

e) *Eleaniztasuna*.

Hizkuntza desberdinetako hiztegiak integratu nahi izan ditugu ANHITZ HEBan. Hiztegi elebidunen erabilpenak eta hizkuntzarteko integrazio-ereduak aztergai izan ditugu puntu honetan. Bi dira, gaingiroki ikusita, eredu nabarmenenak:

- Interlingua ereduak: hizkuntzarteko loturak interlingua hiztegi bakar baten bitartez gauzatzen dira. Normalean hizkuntza batetik besterako jauzia bi loturaren bidez egiten da: sorburu-hiztegitik interlinguarakoa bata, eta interlinguatik xede-hiztegirakoa bestea. Interlingua hiztegitan alde handiak daude sistema batzuetatik besteetara, baina, ezaugarri berezkorik baldin badute, hori, hizkuntzetatik independente izatea —edo izaten saiatzea— da. Lehenago ikusi ditugun sistema lexikal eleanitzen artean badira eredu honetan oinarria dutenak, adibidez, KBMT-89, EDR edo NADIA.
- Transferentzi ereduak: hizkuntzarteko loturak hiztegi elebidun norabide bakarrekoen bidez mamitzen dira. Loturak zuzenekoak dira, hitzetik hitzerakoak, edota adieratik adierarakoak, baina, nolanahi ere, tarteko errepresentaziorik gabekoak. Konparazio batera, ACQUILEX edo MULTILEX sistemak transferentziaren molde honetara egokitzen dira.

Gure aukera transferentziaren aldera lerratu da eta, ondorioz, ANHITZ HEBan adieratik adierarako transferentzi edo itzulpen-erlazioak errepresentatu ditugu⁸². Hizkuntzarteko erlazioak zuzenekoak izateak ez du esan nahi, ordea, lotura soilak direnik. Aitzitik, lexiko mailako itzulpen-erlazioa adiera-korrespondentzia baino konplexuagoa da askoz. Lehen ere ikusi dugu nola dibergentziak, baliokidetza partzialak, hitz anitzeko unitate lexikalen itzulpenak, "hutsune" lexikalak, etab.ak errepresentatu beharreko fenomenoak diren. Gure proposamenean fenomeno horiei buruzko informazioaren kokalekua ez dago hiztegi elebazarretan, hiztegi elebidunetan baizik⁸³. Hots, itzulpen-erlazioek, lotura soil baino areago, informazio-gordailu behar dute izan. Baieztapen horretatik dator, ondorio gisa, ANHITZ HEBan

⁸¹ Oso bestalde, ezaugarri desberdinetako hiztegiak biltzea da gure xedea.

⁸² Adieratik adierako erlazioak hobestean NADIA sistemaren antza ere hartzen du ANHITZek. Hala ere, hurbilpen gisa balio badezake ere, ez da beti adieratik adierarakoa izaten itzulpen-erlazioa, § III.4.3.2n ikusiko denez.

⁸³ Literaturan badu sostengurik erabaki horrek. Adibidez, honela mintzo da Diana Santos (90:332)n: "there are problems which are bilingual in nature, and cannot therefore be properly dealt with in only one language".

Ezagutza-Base Elebidunak —transferentziarako ezagutza-base berezituak— integratzeko hartu dugun erabakia.

f) *Modulartasuna.*

Sistemaren aspektu hau, modulartasuna, diseinu-erabakien ondorio azpimarragarria da gure ustetan. Esan nahi baita informazio lexikalaren tipifikazioa planteatzeak edo hiztegien izaera autonomoa babesteak sistemaren kontzepzio modularra erakusten duela. Halaber, transferentziari buruzko informazioa ingurune elebakarretatik aldentzea modulartasuna indartzera dator. Kontu horri atxikiz, Beaven-ek (92:603) Shake-and-Bake itzulpen-sistemen propietate hori nabarmentzen du, gure asmoen isla ere baden parrafo honetan:

It offers a greater modularity of the monolingual components, which can be written with independence of each other, using purely monolingual considerations. These are put into correspondence by means of a bilingual lexicon.

g) *Itzulpenera zuzendua.*

Hiztegiak, direnak direla, itzulpenerako tresnatzat har litezke, batzuk hobeak, bestetzuk kaskarragoak, baina gutxi edo gehiago itzulpenean erabil daitezkeenak. Beste kontu bat da baina, bereziki itzulpenera zuzendutako hiztegi-sistema helburutzat hartzea. Argigarria da Ulrich Heid-ek dioena (90:108):

The comparison of bilingual dictionaries for 'human' users with those for machine translation should demonstrate that both have commonalities in the area of 'data description', differences, however, in the area of 'presentation'.

Parrafo horren harian bi erreflexio otu zaizkigu: lehena, hiztegi-sistema bati "itzulpenerako berezitasunak" erantsi nahi zaizkionean itzulpen automatikoa dela erreferente nagusia, eta bigarrena, itzulpen automatikorako hiztegien informazioa giza itzultzaileari aurkeztea ez dela nabaria.

Itzulpen automatikorako lexikoiei begiratu bat eman diegu § III.2n, batzuetatik eta besteetatik zenbait ideia jasoz. Zerrendatu egingo ditugu hemen lexikoi horietatik jaso eta ANHITZ HEBaren diseinura eramango ditugun informazio motak:

- Esanahia:
egia esan, ez da esanahiaren tratamendu sakona itzulpen automatikorako lexikoen bereizgarria. Izatekotan, alderantziz da, alegia, sarreren definizioek bigarren mailako lekua dutela portaerarekin edo itzulpen-baliokidearekin alderatuta. Dena den, ANHITZ HEBan unitate lexikalen esanahiaren deskribapena lehentasun handikoa da, a) puntuan esan dugunez.
- Transferentziari buruzko informazioa:

transferentzian oinarritutako itzulpen automatikorako sistemek aurrikusten dute transferentzi erlazioak informazioz hornitu beharra. Adibidez, (Fontenelle *et al.*, 94:7)n, METAL sistemaren gainean, "*we have seen that transfer dictionaries have to account for certain structural differences between the source language and the target language*" diotelarik, garbi adierazten dute informazio mota hori premiazkoa dena itzulpen automatikorako. Eta guk erantsi nahi dugu gisa beretsuan dela beharrezkoa giza itzulpenean laguntzeko. Ikuspuntu honetatik ere indar hartzen du transferentziarako ezagutza-base berezituaren erabilpenak.

- Portaera morfologikoa:

informazio morfologikoaren deskribapena itzulpenetarako lexikoien atal sailhestezina da. ANHITZ HEBan deskribapen morfologikoa baino areago portaera morfologikoa —aktiboa— integratzeko aukera hobetsi da, analizatzaile/sortzaile morfologikoak tresna autonomo bezala HEBan integratuz. Analizatzaile/sortzaile horien lexikoiak ez dira HEBaren menekoak.

- Portaera sintaktikoa:

Boguraev eta Levin-en hitzak dira ondokoak, ezagutza lexikalaren alderdi sintaktikoaz (93:328):

This facet of lexical knowledge needs to be fully treated in a LKB not only to ensure its completeness, but also because it plays an important part in a speaker's productive use of language.

Planteamendu horrekin ados egon arren, "*fully treated*" horri helburu handios irizten diogu gurerako. Horregatik, ANHITZ HEBan, jakitun garela zein zaila den unitate lexikalen portaera sintaktikoa osotara jasotzea, aspektu nabarmen batzuen deskribapenera jo dugu. Itzulpenaren azterketak erakutsi digu zein diren aspektu nabarmen horiek, hots: azpikategorizazioak, lokailuak, kokakidetzak eta hautapen-murritzapenak.

- Azpikategorizazioak:

testuingurupeko erabileraren berri emateko garrantzitsuak direna ez dago duda-mudarik; Pustejovsky-k, adibidez, bere teoria lexikal-semantikoan errepresentazio-maila bat eskaintzen dio argumentu-egiturari.

- Lokailuak:

unitate lexikal berezi samarrak, trataera apartekoa eskatzen dutenak. Beren erabileraren deskribapenak laguntza eskergarria ematen die itzultzaileei.

- Kokakidetzak:

testu-sorkuntzan berebiziko garrantzia dute kokakidetzek (Smadja & McKeown, 91; Heylen *et al.*, 93). Askotan, ordea, irudi lezake

arbitrarioak direla eta arau sintaktiko edo semantikoetatik at geratzen direla. Hargatik da hain inportantea kokakidetzen berri esplizituki ematea.

- Hautapen-murritzapenak:
murritzapen hauek izaera semantikoa dute baina sintaktikoki identifikatzen dira (Fass, 93). Hitzen testuingurupeko erabilera okerrak saihesteko dira baliagarriak.

h) *Inferentzi gaitasuna.*

Orain artean ukitu ditugun informazioa errepresentatzeko aukerei inferentzi mekanismoak gehitzen zaizkie ANHITZ HEBan. Mekanismo hauek, dela informazio inplizitua azaleratuz, dela erlazio lexikal-semantikoen propietateak ustiatuz, izaera aktiboa ematen diote HEBari.

i) *Ezagutzaren errepresentaziorako ingurunea.*

Aurreko puntuen sostengu izateko, ezagutzaren errepresentaziorako marko egokia aukeratu behar da nahitaez. Pustejovsky eta Boguraev-en hitzetan (94:296):

Our thesis is that a theory of lexical semantics making use of a knowledge representation (KR) framework offers a richer, more expressive vocabulary for lexical information.

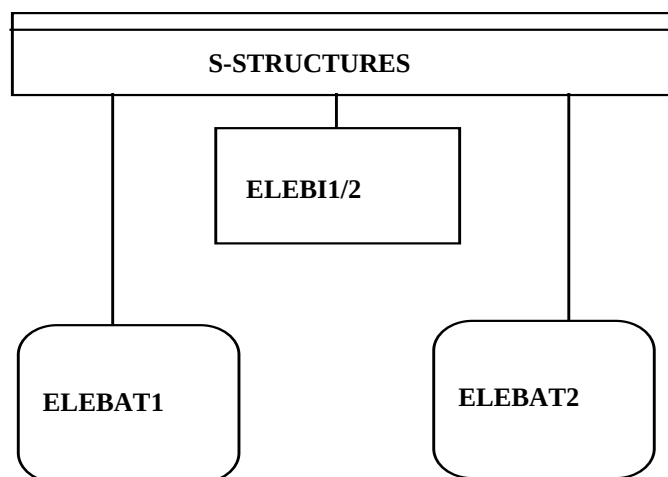
Hara non, puntu honetan, batera bildu nahi ditugun KADSeko domeinu-errepresentazioaz egindako hausnarketak eta ezagutza lexikal-semantikoari buruz xedatutakoak. Ondorioa, lehen ere aurreratu duguna, hauxe da: KEE ezagutzaren errepresentaziorako ingurunea erabiliko dugu ANHITZ HEBa maneiatzeko.

III.4 ANHITZ HEBaren deskribapena.

ANHITZ HEBaren egungo bertsioa bi hizkuntzatako —euskara eta frantsesa— hiztegietatik jasotako datuez hornitu da. Horregatik, HEBaren diseinua inguru eleanitzetarako pentsatuta badago ere, inguru elebidunera mugatuko gara hemendik aurrerako deskribapenean.

Hiru modulutan dago antolatuta HEBa: euskarazko modulu elebakarra —ELEBAT1—, frantsesezko modulu elebakarra —ELEBAT2— eta euskara/frantsesa modulu elebiduna —ELEBI1/2—, III.1 irudian adierazi bezala. S-STRUCTURES Ezagutza-Basean —EBan— hiru modulu horiek konpartitzen dituzten objektuei buruzko ezagutza errepresentatzen da⁸⁴.

⁸⁴ ANHITZ ereduaren deskribapenerako oinarri hartu dugun lengoaiak, CMLk, ontologia deritza maila horretako ezagutzari. Hara zer den (Schreiber *et al.*, 94: 16): "The ontology can be viewed as a *meta model* describing the structure of the domain knowledge".



III.1 irudia.- ANHITZ HEBaren ikuspegi orokorra.

Modulu elebatarrek sorburu- eta xede-hizkuntzetako informazioa gordetzen dute eta ELEBI1/2 modulu elebidunean adiera elebatarren arteko korrespondentziak biltegitratzen dira.

ELEBAT1 eta ELEBAT2 moduluen egituraketarako HIZTSUA sistemaren Hiztegi-Ezagutza Basearen eredua jarraitu da (Artola, 93; Agirre *et al.*, 94e). HIZTSUA hiztegi-sistema urgazle adimenduna da, elebatarra eta giza erabiltzaileari testu-ulerkuntzan eta sorkuntzan laguntzeko pentsatua.

Modulu horien deskribapenari ekin aurretik esan dezagun zein izan diren datu-horniketarako erabili diren hiztegiak: euskarazko modulu elebatarretarako I. Sarasolaren "Hauta-lanerako Euskal Hiztegia" (Sarasola, 91) baliatu da nagusiki, baina artean osatu gabeko sarrerak egokitu zaizkigunean "Euskararako Hiztegia"ra (Aurrekoetxea *et al.*, 86) jo dugu; frantsesezko EB elebatarra "Le Plus Petit Larousse" hiztegitik eraiki da, eta azkenik, modulu elebidunerako "Euskararako Hiztegia"ren euskara-frantsesa aldea erabili da Lhanderen "Dictionnaire Basque-Français" hiztegiarekin batera.

III.4.1. S-STRUCTURES Ezagutza-Basea.

S-STRUCTURES EBak erakusten duen hierarkian, OBJETS-BASIQUES klasearen azpian, hiru motako objektuen erazagupenak biltzen dira: atributuak ("ATTRIBUTS"), iratxoak ("VALEURS-ACTIVES") eta erregelak ("REGLES").

ATTRIBUTS klaseak modulu elebatarretan definitutako atributuen sailkapena erakusten du. Hiru dira atributu-klase nagusiak:

- *Informazio orokorreko atributuak:* orotariko informazioa gordetzen dute, kategoria sintaktikoa, definizio- eta adibide-testuak, etab.ak.

- *Errepresentazio-atributuak*: definizioen azaleko maila islatzen dute. Definizioen forma da horietan jasotzen dena.
- *Atributu erlazionalak*: lexikoaren ikuspegi erlazionala erakusten dute hauek. HEBaren portaera deduktiboaren sostengua dira.

VALEURS-ACTIVES klaseak modulu elebakarretako slotei lotutako "iratxoak" edo prozedura-instantziak erazagutzen ditu. Badira, beste batzuen artean, errepresentazio-atributuetatik erlazionalak iristeko iratxoak, edota erlazioak bi zentzuetan mantentzekoak.

REGLES klasean kokatzen dira dedukzio-prozesu desberdinetan erabili beharreko erregelak. REGLES motako objektuak dira desanbiguaziorako, edota osaketa taxonomiko eta sinonimikorako erregelak.

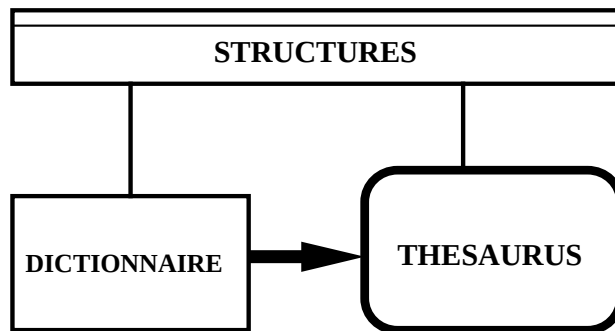
III.4.2. Ingurune elebakarrak.

Modulu elebakarrak ANHITZ HEBan hiztegi bat baino gehiagoz osa daitezke, guztiak, hori bai, hizkuntza bakar eta berekoak. Hiztegi horietako bakoitza errepresentatzeko HIZTSUAren eredua erabili da, eta ondoko atalean eredu horren laburpena doa.

III.4.2.1 Hiztegi elebakarraren errepresentazioa.

Hiru osagai desberdinetan banatu da hiztegi elebakarretako ezagutzaren errepresentazioa. Osagai horiek ezagutza-base dira aldi berean, zein bere eduki eta egituraketarekin (ikus III.2 irudia):

- THESAURUS EBa: adieren sare kontzeptuala da. Adieren arteko loturak erlazio lexikal-semantikoen bidez gauzatzen dira.
- DICTIONNAIRE EBa: hiztegi-sarreratik THESAURUSEko adieretara iristeko bidea ematen du. Hitzen eta adieren arteko lotura EB honetan adierazten da.
- STRUCTURES EBa: THESAURUS eta DICTIONNAIRE EBetako kontzeptu eta erlazioei buruzko meta-ezagutzaren gordailu da. Egiturei buruzko ezagutza hemen zehazten da, hierarkikoki zehaztu ere, slot-ak espezifikatuz eta fazeten bitartez slot edo atributu horiei buruzkoak, balio-heinak, herentzia-rolak, etab.ak, deskribatuz.



III.2 irudia.- Hiztegi elebakarra errepresentatzeko eskema orokorra.

III.4.2.1.1 THESAURUS Ezagutza-Basea.

THESAURUSEk frameen sare semantiko gisa errepresentatzen du hiztegia. Frameek bi erako kontzeptuak identifikatzen dituzte:

- kontzeptu bakunak: hitz bakarrekoak, hitz baten adierari dagozkionak
- kontzeptu sintagmatikoak: kontzeptuek hiztegi-definizioetan dituzten agerpenei dagozkie. Sintagmak adierazten dira, hots, hitz bakar baten bidez adierazi ez diren kontzeptu "konplexuak".

§ III.3n aurreratu dugun bezala, kontzeptuak erlazio lexikal-semantikoen bidez elkarlotzen dira. Erlazio horiek izaera desberdinekoak izan litezke: taxonomikoak — hiperonimia, hiponimia—, meronimikoak, erlatore meta-linguistikoek sortarazten dituzten erlazio espezifikoak, etab.

Inplementazioaren mailan erlazio horiek atributuak dira, balio gisa beste kontzeptu bat erakusten duten slotak⁸⁵. Hiperonimia eta hiponimia bezalako erlazioak KEEren erlazio hierarkikoetarako primitiboen bidez errepresentatu dira, hartara programazio-ingurunearen herentzia-mekanismoak bete-betean aprobetxatuz.

Adibide baten harian errazago ulertuko dugu unitate, atributu eta erlazioen arteko antolaketa THESAURUSen. Izan bedi ondoko definizioa:

hiriburu I I: herrialde bateko hiri nagusia

⁸⁵ Ez da gure gustoko aukera izan, programazio-inguruneak, KEEk, inposatutakoa baizik. Aukera horrek desitxuratu egiten du pixka bat kontzeptuen errepresentazioa. Bi plano daude, gure irudiko, kontzeptuen errepresentazioan: bata, barne-deskribapena, kontzeptuen ezaugarri edo *atributuen* berri ematen duena — slotetan adierazi ohi da—; bestea, kanpo-deskribapena, beste kontzeptuekiko erlazioak adierazten dituen. Erlazioak sloten bidez adierazteak nahasi egiten ditu bi mailok.

Maila definitorioan hauxe litzateke definizio horretako unitateei dagokien informazioa⁸⁶:

|hiriburu I 1|

MEMBER.OF: NOMS

GROUPE-CATEGORIEL: NOM

CLASSE-ATTRIBUT: INFO-GENERALE

TEXTE-DEFINITION: "herrialde bateko hiri nagusia"

CLASSE-ATTRIBUT: INFO-GENERALE

DEF-CLASSIQUE: |hiri I 1#2|

CLASSE-ATTRIBUT: DEFINITOIRES

RELATIONNELS-CORRESPONDANTS: DEFINI-PAR

|hiri I 1#2|

SUBCLASS.OF: |hiri I 1|

MEMBER.OF: NOMINALES

TEXTE: "herrialde bateko hiri nagusia"

CLASSE-ATTRIBUT: INFO-GENERALE

NONGO: |herrialde I 1|

CLASSE-ATTRIBUT: SYNTAGMATIQUES

DETERMINATION: BAT

RELATIONNELS-CORRESPONDANTS: LOCATIF

LOCATIF: 1.0

CARACTERISTIQUE: |nagusi I 1|

CLASSE-ATTRIBUT: SYNTAGMATIQUES

RELATIONNELS-CORRESPONDANTS: CARACTERISTIQUE

CARACTERISTIQUE: 1.0

|herrialde I 1|

...

Frantsesezko definizioei, adibidez "*capitale I 3: ville principale*" definizioari, informazio-antolaketa berdina egokitzen zaie:

|capitale I 3|

MEMBER.OF: NOMS

GROUPE-CATEGORIEL: NOM

CLASSE-ATTRIBUT: INFO-GENERALE

TEXTE-DEFINITION: "ville principale"

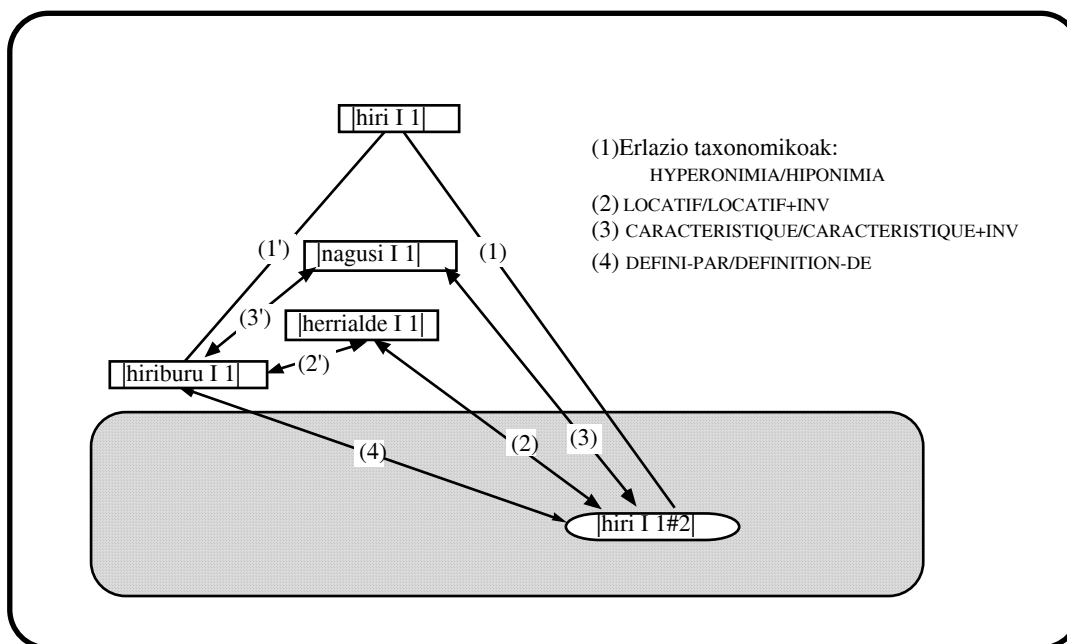
CLASSE-ATTRIBUT: INFO-GENERALE

DEF-CLASSIQUE: |ville I 1#1|

CLASSE-ATTRIBUT: DEFINITOIRES

RELATIONNELS-CORRESPONDANTS: DEFINI-PAR

⁸⁶ Unitateak letra loditan daude, slotak maiuskulatan eta fazetak edo sloten propietateak etzanetan.



III.3 irudia.- |*hiriburu I 1*| kontzeptuaren ikuspegi erlazionala THESAURUS sarean⁸⁷.

(4) erlazioak baliokidetza kontzeptuala adierazten du. |*hiri I 1#2*| kontzeptu sintagmatikoa |*hiriburu I 1*| kontzeptuaren definizioa denez, biek dituzte erlazio berdinak sare kontzeptualean, lehenarenak zuzenekoak, bigarrenarenak deduzituak.

III.4.2.1.2 DICTIONNAIRE Ezagutza-Basea.

DICTIONNAIRE EBko unitateak hiztegi-sarrerak dira. Unitate horiek atributu bakarra dute, SENS atributua, zeinek balio gisa hartzen duen sarrerari dagozkion adieren zerrenda. Horrela ezartzen da hitzen eta adieren arteko lotura. Ondoko adibidean, *hiri* hitzari — DICTIONNAIREko unitatea— *hiri I 1* eta *hiri I 2* adierak dagozkio —THESAURUSko unitateak—:

|hiri|
SENS: |*hiri I 1*|, |*hiri I 2*|

III.4.2.1.3 STRUCTURES Ezagutza-Basea.

STRUCTURES EBan DICTIONNAIRE eta THESAURUS EBetako unitateen motak definitu eta sailkatu egiten dira antolamendu hierarkiko batean. STRUCTURESek unitateak motak

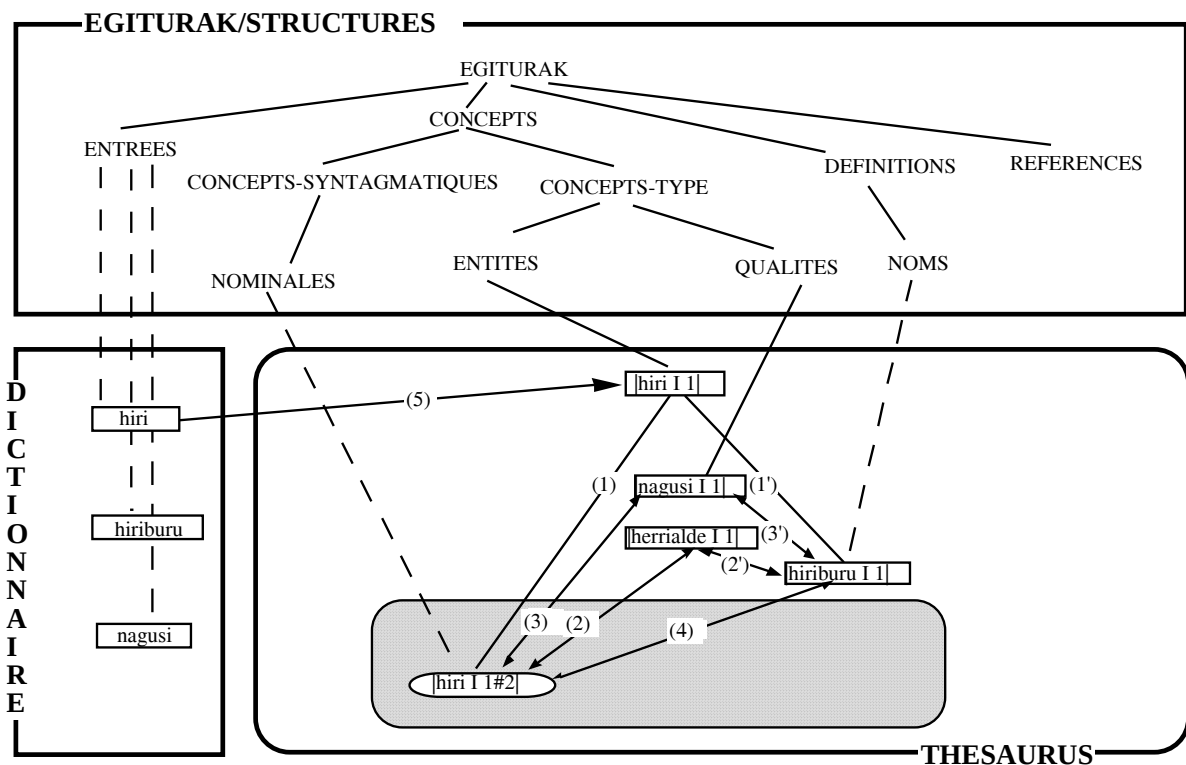
⁸⁷ Hiperonimia/hiponimiako erlazio taxonomikoak gezi-muturrik gabe marratu dira irudian. (1'), (2') eta (3') zenbakia duten erlazioak deduzituak dira. Kontzeptu sintagmatikoak plano ilunduan ageri dira.

dira, eta beren definizioko ezaugarriak mota horretako unitateek heredatu egiten dituzte. Lau azpiklase daude EGITURAK erro-motaren azpian (ikus III.4 irudiko adibidea):

- ENTREES ("sarrerak"): hiztegi-sarrerak dira, hau da, DICTIONNAIRE EBko unitateak. Bi azpiklasetakoak izan litezke sarrerak: LOKUZIOAK edo SARRERA-BAKUNAK.
- DEFINITIONS ("definizioak"): klase honek definitutako adierak biltzen ditu, bere baitan kategoria sintaktikoaren arabera sailkatuz.
- REFERENCES ("erreferentziak"): kontzeptuak dira erreferentziak, baina "definizioak" ez bezala, hauek besteren definizioetan agertzeagatik sortu dira THESAURUSen.
- CONCEPTS ("kontzeptuak"): oro har, edozein unitate kontzeptual da CONCEPTS motakoa, ez bakarrik hiztegi-adierak. Hiru sailetan edo azpiklasetan banatzen dira kontzeptuak:
 - CONCEPTS-TYPE ("mota-kontzeptuak"): Quillian-en (1968) "type nodes" direlakoei dagokie. CONCEPTS-TYPE klasekoak dira THESAURUSen errepresentatutako kontzeptu guztiak, bai bakunak, bai sintagmatikoak. Lau azpiklasetan banatzen da: ENTITES, ACTIONS/EVENEMENTS, QUALITES eta ETATS.
 - CONCEPTS-SYNTAGMATIQUES ("kontzeptu sintagmatikoak"): Quillian-en "token"ak —mota-kontzeptuen agerpenak definizioetan— maneiatzeko sortu dira. Eduki semantikoa duten kontzeptu bat baino gehiagoz osatutako egitura sintagmatikoak adierazten dituzte. Adibidez, |hiri I 1#2|-k *herrialde bateko hiri nagusia* izen-sintagma errepresentatzen du.
 - CONCEPTS-AMBIGUS ("kontzeptu anbiguoak"): erabat desanbiguatu gabeko kontzeptuak dira sail honetatik zintzilik jartzen direnak.

III.4.2.1.4 Adibidea.

Adibide batez ilustratu nahi dugu STRUCTURES, DICTIONNAIRE eta THESAURUS ezagutza-baseei buruz esandakoa. Horretarako |hiriburu I 1| kontzeptuaren definiziora joko dugu berriro ere, III.4 irudian definizio horrek sortarazten dituen ezagutza-baseen arteko erlazioak grafikoki erakusteko.



III.4 irudia.- Hiru EBetako unitateak eta beraien arteko erlazioak⁸⁸.

III.4.2.2 Hiztegi elebkarren bateratzea.

ANHITZ HEBaren antolaketan aukera ematen da hizkuntza bereko hiztegi bat baino gehiago batera errepresentatzeko⁸⁹. Aukera hori garatzeko proposatzen den hurbilpenak hiztegien autonomia babestu nahi du —ikus errepresentazio-ereduaren printzipioak § III.3n—. Hori kontuan, hiztegiarteko "mapping" erlazioa sortu da adiera guztiz baliokideen arteko mapaketa adierazteko. Jakitun izanda ere zein zaila den hiztegi desberdinen adiera-banaketak elkarrengatzea, hurbilpen horri ikusten dizkiogu arazorik arinenak. Esan nahi baita, egokiagoa dela baliokidetza erabatekoetan oinarritzea eta erdizkakoak edo partzialak saihestea. Hartara, behinik-behin, ziurra da mapping erlazioek, ekartzekotan, aberastasuna ekarriko dutena beti⁹⁰.

⁸⁸ (1), (2), (3) eta (4) erlazioak THESAURUSen barnekoak dira, III.3 irudian azaldu dugunez. (5) zenbakiarekin etiketatu dugun erlazioak sarrera-hitzaren eta adieren arteko lotura adierazten du. STRUCTURES eta THESAURUSaren arteko erlazioak bi tipokoak izan daitezke: klase/azpiklase erlazioak (etenik gabeko marraz) eta klase/instantzia erlazioak (marra ez-jarraiak).

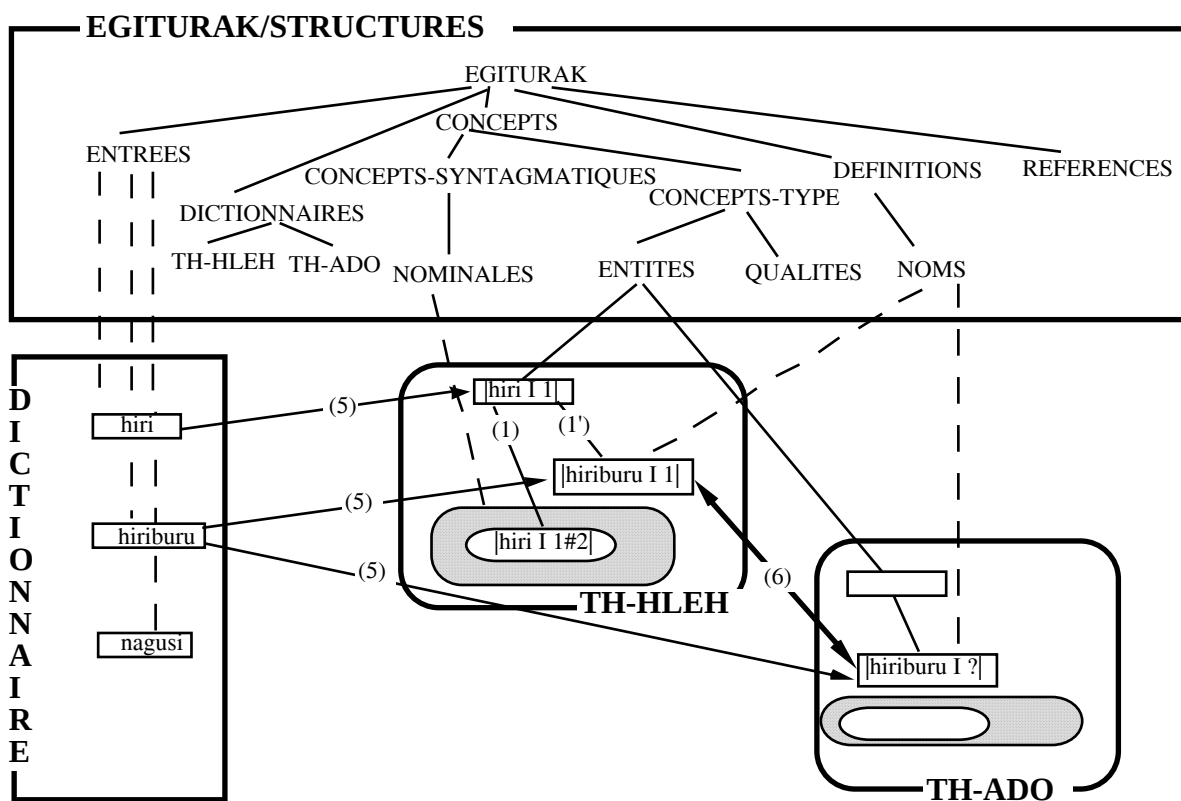
⁸⁹ Nahiz eta egungo inplementazioan hiztegi bana egon ELEBAT1 eta ELEBAT2 modulu elebkarretan.

⁹⁰ Ezagutzaren kalitatea hobesten dugu nabarmen, kuantitatearen kaltetan.

Hiztegien batera jartze eta autonomiaren arteko oreka lortzeko haxe da antolaketa-proposamena (Agirre *et al.*, 93):

- STRUCTURES bakarra hizkuntza bakoitzeko. Ezagutza-base honetan hizkuntza bereko hiztegi guztiek konpartitzen dituzten egiturak eta objektuak daude definituta. § III.4.2.1.3n deskribatutako STRUCTURES EBari DICTIONNAIRES klasea erantsi zaio hizkuntza horretako hiztegi guztien nolakotasunak adierazteko.
- DICTIONNAIRE bakarra, baita ere, hizkuntzako. EB honek bidea erakusten du hiztegi-sarreretatik hiztegi guztietan —sarrera horiei— egokitzen zaizkien adieretara.
- THESAURUS bat behar da hiztegiko. THESAURUS horietako unitateek *mapaketa* erlazioa dute definitua, ikusi dugunez. Kontuan hartzekoa da, bestalde, unitateen identifikaziorako, orain, hiztegiko erreferentzia ere adierazi behar dela.

III.5 irudian antolaketa horren ikuspegi grafikoa ematen da.



III.5 irudia.- Ingurune hiztegianitz elebakarra.

(6) zenbakiaz "mapping" erlazioa —mapaketa edo hiztegiarteko baliokidetza— dago etiketatua. TH-HLEH Hauta-Lanerako Euskal Hiztegiari dagokion THESAURUSa da, eta TH-ADO Euskararako Hiztegiari dagokiona.

III.4.3. Ingurune elebiduna.

Inguru elebidunaren osaketarako bi modulu elebakar elkarlotu dira modulu elebidun baten bitartez, transferentzian oinarritutako tradizioari atxikiz (ikus III.1 irudia). Modulu elebakarrak, dagoeneko deskribatuak, bere horretantxe utzi dira ingurune elebiduna eraikitzeko. Ez da, beraz, ezagutza-base elebakarrak aberastu edo aldatu beharrik izan ingurune elebiduna iristeko.

Noski, elebakarrak aldatu ezean, modulu elebidunak hartzen du bere gain hiztegi elebakarretako kontzeptuak lotzea. Horretarako, ELEBI1/2 deitu dugun modulu elebiduneari bi azpi-EB definitu dira, honako hauek:

- EUSK-FRANTS EBa, non euskara-frantsesa hiztegi elebidunetik erauzitako informazioa errepresentatzen den.
- S-ELEBIDUNA EBa: hiztegi elebidunen errepresentaziorako baliatzen diren egitura eta atributuak deskribatzen dira bertan.

III.4.3.1 EUSK-FRANTS Ezagutza-Basea.

EUSK-FRANTS EBan hiru klaseko unitateak biltegiratzen dira:

- *Sorburu-unitateak.* Euskarazko unitateak dira, huraxe baita sorburu-hizkuntza. Unitate hauen egitekoa euskarazko moduluko unitateekin lotura finkatzea da. Ohargarria da lotura hori "matching" edo izenen parekatzearen bidez zertzen dela. Ez dago, horretara, euskarazko moduluko unitateak erakusten dituen lotura-slot edo atributurik. Aspektu horrek, inplementazioko huskeria irudi dezakeen arren, elkar-menpekotasun "fisikorik" gabe uzten ditu modulu elebakarretako eta elebiduneko unitateak⁹¹. Horrek ez du esan nahi, ordea, moduluek elkar aberasten ez dutenik. Gurean ere, Calzolari-k eta Picci-k (86:80)n gorai patzen duten bezala, aberasketa bi zentzutan ematen da: elebakarretik elebidunera eta elebidunetik elebakarretara. Unitate hauetan definitutako BALIOKIDEAK erlazioak EB bereko unitate elebidunekin elkartzen ditu sorburukoak.
- *Xede-unitateak.* Hauek frantsesezko unitateak dira, eta sorburu-unitateen simetrikoak guztiz.
- *Unitate elebidunak.* Bi unitate elebakar baliokideen artean daudenak dira, baliokidetza buruzko informazioa daukatela. Kontuan izatekoa da baliokidetza hauek adieren artekoak direla eta bat:bat estilokoak beti.

⁹¹ Horri esker posible da modulu elebakarrek lan egitea derrigorrean inguru elebiduna kargatu gabe.

III.4.3.2 S-ELEBIDUNA Ezagutza-Basea.

Ezagutza-base elebidunetako egiturak definitzen dira S-ELEBIDUNA EBan. Egitura horiek hierarkia batean sailkatzen dira; zehazki, UNITATEAK goi-klasearen azpian hiru azpiklase deskribatzen dira:

- SORBURUKOAK: klase honetako instantziak dira EUSK-FRANTSEko sorburu-unitateak. Honetan definitzen da BALIOKIDEAK erlazioa, zeinek sorburu-unitateen baliokidetzak deskribatzen dituzten unitate elebidunak erakusten dituen.
- HELBURUKOAK: xede-unitateen gain-klasea edo mota da honakoa. Unitate elebidunekiko lotura ORDAINA+INV erlazioaren bidez adierazten da.
- ELEBIDUNAK: transferentzi unitateen klasea. Kontzeptu-baliokidetzari buruzko informazioa, transferentziaz jakin beharrekoa, hemen deskribatzen da. Egungo inplementazioan ondoko atributuak definitu dira:

- BALIOKIDEAK+INV: SORBURUKOAK klasean definitutako BALIOKIDEAK atributuaren alderantzizkoa.
- ORDAINA: unitate elebidun eta xede-unitateen arteko baliokidetzaz-loturak adierazten dira atributu honen bidez. Honen alderantzizkoa ORDAINA+INV atributua da, HELBURUKOAK klasean definitua.

BALIOKIDEAK eta ORDAINA atributuak konposatuz osatzen da sorburu-kontzeptutik xede-kontzepturako bidea. Atributu alderantzizkoen konposaketak —ORDAINA+INV eta BALIOKIDEAK+INV— alderantzizko bidea ematen du.

- BALIOKIDETZA-MAILA: era esplizituan adierazten da zenbateraino diren baliokide sorburu- eta xede-kontzeptuak. Besterik ezean, baliokidetzaz erabatekoa dela lehenesten da.
- BALIOKIDETZA-MOTA: kontzeptuen arteko baliokidetasuna landu da ANHITZ HEBan, baina THESAURUS elebazarretako kontzeptuak bi erakoak izan daitezke: bakunak edo sintagmatikoak (ikus § III.4.2.1.1). Bereizketa hori lagungarria da zenbait transferentzi erlazio deskribatzean.

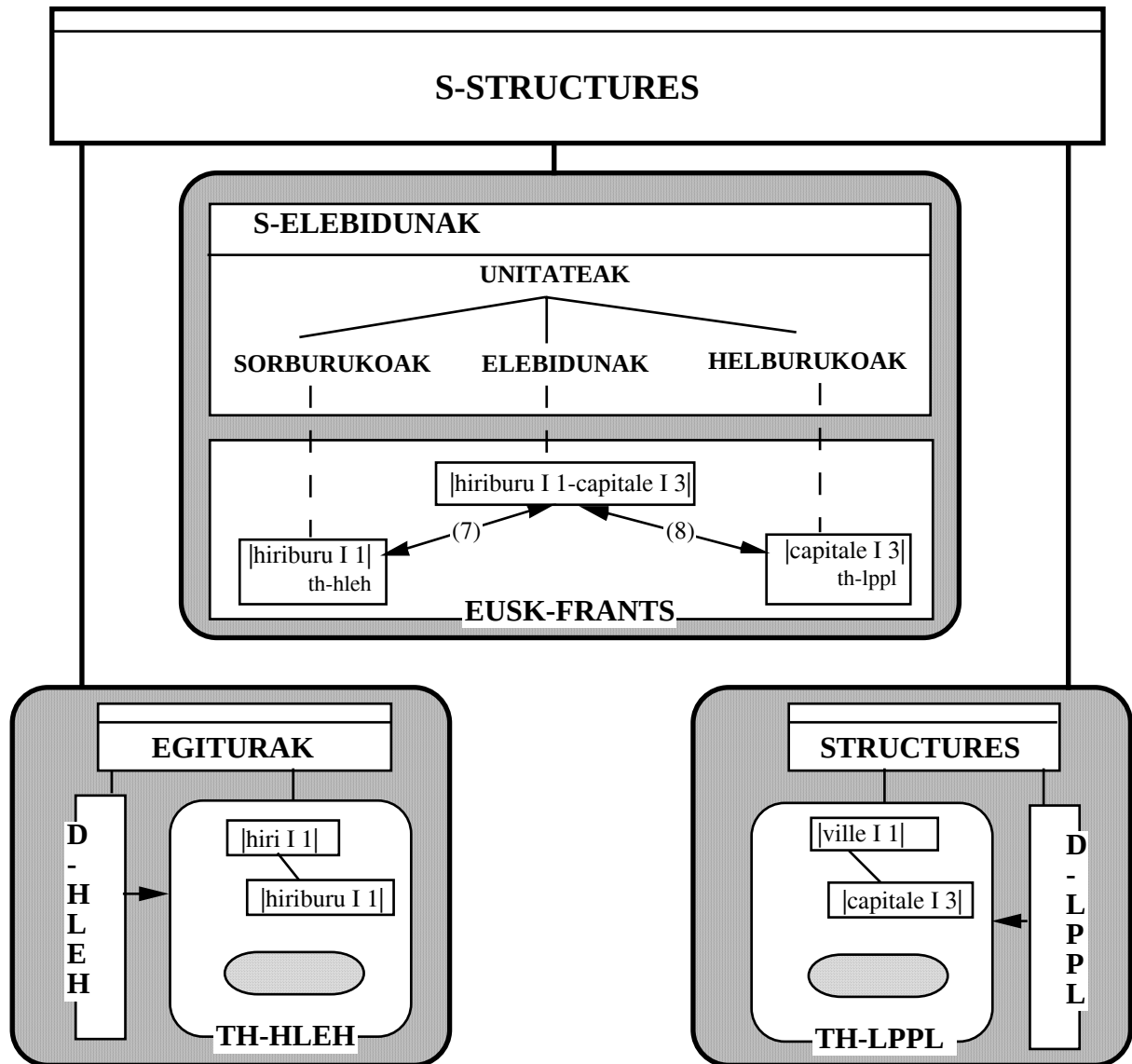
Hain zuzen ere, bereizketa hori dela bide, hitzen itzulpenak esaldiak direnean, sorburu-hitzari kontzeptu bakuna dagokio eta xede-esaldiari kontzeptu sintagmatikoa. Erlazio horiek sinpleak dira ANHITZen —bi kontzepturen artekoak—, baina deskribatutako baliokidetzak konplexuak —adieretatik esaldietarakoak edota, alderantzizko zentzuan, esaldietatik adieretarakoak—.

Horrela bada, BALIOKIDETZA-MOTA atributuak agerian utzi nahi du nolakoa den transferentzia: adieren artekoa —*kontzeptuala*— edo adiera eta esaldiaren artekoa —*sintagmatikoa*—.

- HIZT-ELEBIDUNA: erabili den hiztegi elebidunaren erreferentzia gordetzen da hemen.

III.4.3.3 Adibidea.

Ingurune elebidunerako proposatu ditugun EBen xehetasunak deskribatzen ihardun gara aurreko ataletan. Ez genuke nahi, ordea, xehetasunetan aritu eta ikuspegi orokorra galtzea. Horregatik kokatu dugu hemen III.6 irudia, zeinek ANHITZ HEBaren ikuspegi orokorra eman nahi bailuke, *hiriburu* eta *capitale* kontzeptuen arteko erlazioak adibide gisa hartuta.



III.6 irudia.- ANHITZ HEBaren ikuspegi orokorra.

Irudi honetako (7) lotura BALIOKIDEAK/BALIOKIDEAK+INV erlazioa da, eta (8)a, berriz, ORDAINA/ORDAINA+INV erlazioa. Unitate elebidunak ondoko informazioa dauka:

|hiriburu I 1-capitale I 3|

MEMBER.OF: ELEBIDUNAK
BALIOKIDEAK+INV: |hiriburu I 1|
BALIOKIDETZA-MAILA: 1
BALIOKIDETZA-MOTA: bakuna
HIZT-ELEBIDUNA: ADOREZ
ORDAINA: |capitale I 3|

III.4.4. Itzulpena zuzendutako ezagutza.

Itzultzaileekin egindako hiztegi-erabilerari buruzko azterketaren ondorioz, espreski itzulpenean laguntzeko informazioa erantsi beharra ikusi da ANHITZ HEBan. Itzultzaileek erakutsi digute, eta itzulpen automatikorako sistemek berretsi, itzulpenerako informazio deitzen dugun horrek portaera morfologiko-sintaktikoarekin duela zerikusi zuzena. Hala, sail honetan bilduko genituzkeen informazio-motak aipatu ditugu § III.3n, ezagutza-basearen printzipioez aritu garenean, eta gauden atal honetan, informazio hori errepresentatzeko proposamena jorratu nahi dugu.

Informazio morfologikoa modulu berezituetan lantzea proposatzen da ANHITZ sisteman. Modulu horiek analizatzaile/sortzaile morfologikoak dirateke, eta analizatzaile/sortzaile oso izan ere, hau da, bai lexiko mailako informazioa, bai portaerari buruzkoa, biak bilduko lirateke modulu horietan. ANHITZ HEBak ez luke, hortaz, informazio morfologikoaren gaineko ardura zuzenekorik⁹².

Azpikategorizazioak aipatu ditugu § III.3n, lexikoaren portaera sintaktikoan atal garrantzizkoa direlakoan. Iritzi hori, egia esateko, aski hedatua da, eta teoria lexikal-semantiko gehienek aurikusten dute, nolabait, azpikategorizazioen trataera. Gaiaren zailtasunak, ordea, mantso eta zuhur jokatzeko eskatzen du, eta zuhurtzia hori kausa, gure proposamena aditzetara mugatzen da. Hara zein den errepresentazio-proposamena: modulu elebazarretan, STRUCTURES EBetako VERBES klasean —DEFINITIONSen azpian—, AZPIKATEGORIZAZIOAK atributua definitzen da, zein adierazpide-eredu jarraituko den, etab.en berri emanez. Atributu hori, herentzi mekanismoen bidez, definitutako aditz-sarrera guztiek edukiko dute. Bestalde, S-STRUCTURES EBan beste atributu-klase bat sortuko da, SINTAKTIKOAK, eta horren ume izango da AZPIKATEGORIZAZIOAK definitu berria.

Lokuzioen, edo esaera idiomatikoaren, antolamendua eta errepresentazioa buruhauste askoren zio izan da (Dobrovolskij, 94). ANHITZ HEBan hartu den bideak ondoko printzipioa du oin: esaera idiomatikoak kontzeptu bakunak dira, ez sintagmatikoak. Horri

⁹² Euskararako, XUXEN analizatzaile/sortzaile morfologikoaz balia gintezke (Agirre *et al.*, 92). XUXENek darabilen datu-base lexikalari buruz informazio gehiago nahi izatera, ikus (Agirre *et al.*, 94a).

atxikiz, atzipen-arazoetara lerratu da gure arreta, hitzen mailatik THESAURUSEko kontzeptu idiomatikoetara erarik erraz eta zuzenekoenean iristeko aukerak emateko intentzioarekin. Horretarako egin den aberasketa sinplea da oso: DICTIONNAIRE EBetako unitateei — ENTREES motako hiztegi-sarrerak— bi atributu gehitu zaizkie. Bata, LOKUZIOAK izenekoa, lokuzio ez diren sarrerei dagokie eta sarrera agertzen deneko lokuzio guztien erakusle da; bestea, SARRERA-BAKUNAK deitzen dena, lokuzio-sarreraren ezaugarria da eta lokuzioa osatzen duten adiera guztiak erakusteko usatzen da.

Itzultzaileen esanetara makurtuz, bada ukitu beharreko beste gai bat: juntagailuen trataera. Juntagailuak unitate lexikal defnigaitzak dira, horregatik hiztegieta azalpenak eman ohi dira beren erabilerari buruz. Gure ustetan, juntagailuen antolamendu hierarkikoak erruz lagun diezaioke itzultzaileari, bai beste hizkuntzetako baliokideak bilatzean, bai hizkuntza bereko formulazio desberdinak nahi direnean. Sailkatze-eredu hori berezkoa du ANHITZ HEBek. Nahikoa da DEFINITIONS unitatean beste azpiklase bat definitzea, JUNTAGAILUAK, eta horren ume gisa juntagailu-mota guztiak. JUNTAGAILUAK motako kontzeptuen atributurik esanguratsuenak eraketa sintaktikoa deskribatuko du, hartara, itzultzaileari esaldiaren idazketarako informazio eskergarria ematen ahal diola.

Kokakidetzak ere hizpide izan ditugu itzulpenaren, eta bereziki sorkuntza lexikalaren, gorabeherak landu ditugunean. Egia esateko, kokakidetzak eta sorkuntza lexikala batera ageri zaizkigu autore askoren lanetan (Heid & Raab, 89; Smadja, 91; Smadja & McKeown, 91; Heylen *et al.*, 93; Heylen *et al.*, 94). Honako honetan galdera da, alabaina, ea nola adieraz ditzakegun kokakidetzaz-erlazioak gure ANHITZ HEBan. Aipatu ditugun autoreak bat datoz kokakidetzari buruzko informazioa unitate lexikaletan jartzean, eta gehienek Mel'cuk-en funtzio lexikalak ("lexical functions") darabiltzate horretarako. Gu ere Mel'cuk-en eredura hurbilduko gara (Mel'cuk, 82; Mel'cuk & Zholkovsky, 88)⁹³. Funtzio lexikalak erlazio bihurtuko ditugu gurean, hau da, THESAURUSEko kontzeptuek funtzio lexikalei dagozkien slot erlazionalak edukiko dituzte, eta slot horien balioak kontzeptu kokakideak izango dira. Horrek esan nahi du, besteak beste, atributu erlazionalen klasean beste azpiklase bat erazagutuko dugula: FUNTZIO-LEXIKALAK klasea⁹⁴. Atributu hauek STRUCTURESeko DEFINITIONS klasean definituko dira, definitutako kontzeptu guztiek, edozein dela ere beren kategoria, herentziaz jaso ditzaten.

⁹³ Mel'cuk-en teorian hirurogei bat funtzio lexikal definitzen dira unitate lexikalen arteko kokakidetzamenpekotasunak adierazteko. Funtzio horiek "oinarri" unitateei aplikatzen zaizkie eta unitate "kokakide"ak itzultzen dituzte. Badago funtzio hauen erabilera-adibiderik § II.4.4.4n, *kokakide-lexikalen-galdea* ataza primitiboa espezifikatu denean.

⁹⁴ Ohartaraztekoa da, batetik, funtzio lexikalak horrela adierazteak ez dakarrela funtsezko aldaketarik errepresentazio-eskeman, eta bestetik, FUNTZIO-LEXIKALAK motako erlazioak ez direla simetrikoak.

Itzulpenera zuzendutako sail honetan hautapen-murritzapenak aipatu nahi ditugu azkenik. Esan dugu lehen erabilera okerrak saihesteko direla erabilgarriak, azken batean, bateragarritasun semantikoa adierazten baitute. Hautapen-murritzapenak maneiatzeko bi aukera ikusten ditugu. Lehen, ANHITZ HEBa dagoenean utzi eta, erlazio lexikal-semantikoak ustiatuz, bateragarritasun semantikoari buruzko dedukzioak egitea. Bigarrena, hautapen-murritzapenak esplizituki adieraztea, aurretik erlazio aproposak definituta. Lehenengo aukeraren bideragarritasunak zalantza handia sortarazten digu, bai baitakigu zein zaila den, kasu bakanetan izan ezik, dedukzio horiek ziurtasunez burutzea. Bigarren aukerak, itxuraz erabilgarriagoak, hautapen-murritzapenak aurretik eskuratzea eskatzen du, eta hori ez da lan makala, nahiz eta azken aldion saio esperantzagarriak burutzen ari diren.

III.4.5. Inferentzi geruza.

KADS metodologian domeinu-geruzaren gaineko inferentzien multzoari *inferentzi geruza* deritza. Guri dagokigunez, hiztegien aktibotasuna aldarrikatzen dugun heinean, inferentzi mekanismoak hiztegi-ezagutza dira, izaera desberdinekoa nahi bada, baina hiztegiari atxikitutako ezagutza, edozein moduz.

ANHITZ HEBan inferentzien erabilpena bi estadiotan gertatzen da, HIZTSUA sistemak erakutsitako molde berean (Artola, 93; Agirre *et al.*, 94e):

- HEBaren sorkuntza-fasean.

Maila honetan zenbait aberasketa-prozesu aktibatzen da egitate berrien sorrera eraginez. Prozesu hauek bi aspektu ustiatzen dituzte: batetik, sinonimiaren simetrikotasuna eta trantsitibotasuna, eta bestetik, taxonomiaren eta sinonimiaren konposagarritasuna. Desanbiguatze-heuristiko batzuk ere aplikatzen dira fase honetan.

- Informazio-eskariei erantzutean.

Erabiltzaileak edo atazek eskatutakoari ihardestean⁹⁵, ANHITZ HEBak zenbait dedukzio egin dezake, ezagutza, esplizitua nahiz implizitua, unean prozesatuz. Bi maila bereziko ditugu:

- *Dedukzio dinamikoak.* Erregeletan oinarritutako arrazonamenduaren bidetik, maila honetako mekanismorik aipagarriena *erlazio lexikalen konposaketa* da. Mekanismo horrek erlazio lexikalen propietateak ustiatzeko aukera ematen du; horretarako, (R_1 R_2 R_3) ereduko hirukoteak erazagutzen dira, non R_i -ak

⁹⁵ Horra nola kateatzen diren atazak eta inferentziak.

erlazio lexikalak diren. Hirukote horiek trantsibitate-erregelen forma eta esanahia hartzen dute:

if X R₁ Y **and** Y R₂ Z **then** X R₃ Z

Era horretan adierazten dira, konparazio batera:

- erlazio meronimikoen trantsitibotasuna (Cruse, 86; Winston *et al.*, 87):
(PARTIE-DE PARTIE-DE PARTIE-DE)
Ez, ordea, beste hau: (PARTIE-DE MEMBRE-DE PARTIE-DE)
- erlatore bereziekin osatutakoak:
(CARACTERISTIQUE QUI-A POSSESSION)
(OBJECTIF CE-QUI OBJECTIF)
- erlazio meronimiko eta ez-meronimikoen konbinazioak:
(PARTIE-DE LOCATIF LOCATIF)
(MEMBRE-DE HYPERONYME MEMBRE-DE)
(LOCATIF HYPERONYME LOCATIF)

Adibidez, azkeneko hirukoteari dagokion erregelaren instantzia osatzen da *hiriburu* eta *herrialde* kontzeptuen propietateak konbinatuz. Erregela-instantzia horretatik deduzi daiteke *eskualde* kontzeptua *hiriburu* kontzeptuaren lokatiboa dela, hain zuzen, *herrialde hiriburu*-ren lokatibo izan eta *eskualde herrialde*-ren hiperonimo delako. Hau da:

if |hiriburu I 1| LOCATIF |herrialde I 1| **and**
|herrialde I 1| HYPERONYME |eskualde I 1|
then |hiriburu I 1| LOCATIF |eskualde I 1|

- *Oinarrizko funtzioak. Objektuei zuzendutako programazioaren* moldeen arabera, ANHITZ HEBko objektuei beren portaerari buruzko propietateak atxikitu zaizkie. Propietate horiek metodoen bitartez implementatzen dira, hots, objektuen atributuetan dauden LISP funtzioen bitartez. Oinarrizko funtzioak estuki lotuta daude ataza primitiboekin (ikus § II.4.4)⁹⁶ eta, erabiltzaileak edo sistemak eskatuta, mezu bat igorritz exekutatzen dira.

Oinarrizko funtzioen artean, EQUIV da gako ingurune elebidunaren funtzionalitatean. Lehen ere nabarmendu dugu ez dagoela lotura-erakuslerik modulu elebazarretako kontzeptuak erlazionatzeko, eta ondorioz, diogu orain, batetik besterako iragateak EQUIV metodoaren bidez gauzatzen dira. Kontzeptu baten EQUIV oinarrizko funtzioak burutuko du EB elebazar eta elebidunen arteko identifikadore-parekatzea, eta funtzio horren bitartez jakingo da zein diren kontzeptuen transferentzi baliokideak. Hori horrela, EQUIV bilakatu da

⁹⁶ Esan liteke, xehetasunen bat gorabehera, oinarrizko funtzioak ataza primitiboen implementazioak direla.

beste hainbat oinarrizko funtzioen metodo lagungarria. Halaber da ohartaraztekoa, EQUIV funtzioak bereziki lantzen dituela baliokidetza partzialak eta sintagmatikoak⁹⁷, eta emaitzak bueltatzean, baliokideekin batera, baliokidetza-maila ere ematen duela. EQUIV gai da, gainera, BALIOKIDEAK+INV eta ORDAINA+INV erlazioak ustiatuz bi zentzuetako baliokideak lortzeko. Esan nahi baita, hiztegiak "simetrikotu" egiten dituela, (Byrd *et al.*, 87:239)n adierazten den legez.

III.4.6. ANHITZ HEBaren egungo prototipoaren egoera.

ANHITZ HEBaren osaketan hiru hiztegi erabili dira nagusiki: euskarazko "Hauta-lanerako Euskal Hiztegia", frantsesezko "Le Plus Petit Larousse" eta "Euskararako Hiztegia"ren euskara-frantsesa aldea⁹⁸. Sarrera-hitzen aukeraketa eskuz egin da: hasteko, erabileremu bat finkatu da, udal-administrazioa, eta segidan eremu horretako hitzak aukeratu dira hiru hiztegi horietatik.

Horrela osatutako HEBaren dimentsioak jakinarazteko datu batzuk emango ditugu jarraian:

- Euskarazko modulua.

D-HLEHk 168 unitate dauzka. Unitate horietako bakoitzak hiztegiko hitz bat errepresentatzen du eta —SENS atributuaren bidez— TH-HLEHko kontzeptuak erakusten ditu. TH-HLEHan 305 kontzeptu daude denetara:

- 89 DEFINITIONS klaseko kontzeptu
- 90 REFERENCES klaseko
- 126 CONCEPTS-SYNTAGMATIQUES klaseko

TH-HLEHko mota-kontzeptuen hierarkiak —CONCEPTS-TYPEren azpiklaseak— honela antolatu dira:

- 172 azpiklase ENTITESen
- 117 ACTIONS/EVENEMENTSen
- 13 QUALITESen
- 2 ETATSen

Atributu erlazionalen kopuruei dagokienean, behin hasierako eraikuntza eta aberasketa burututa datuok dira azpimarragarri:

⁹⁷ Gogoan izan horiei buruzko informazioa EB elebiduneko unitateetan paratzen dela.

⁹⁸ Horiekin batera, § III.4n aipatu dugun bezala, "Euskararako Hiztegia" eta Lhanderen "Dictionnaire Basque-Français" erabili dira lagungarri gisa.

- ENTITESen azpiko unitateek batezbeste 5.5 arku erlazional dituzte
- ACTIONS/EVENEMENTSen batezbeste 5.9 lotura dituzte beste unitateekin
- QUALITESen 3.3 lotura batezbeste
- ETATSen 2.5 lotura
- Frantsesezko modulua.

D-LPPLn 541 unitate daude, eta TH-LPPLn 1139 kontzeptu:

- 418 DEFINITIONS klaseko kontzeptu
- 513 REFERENCES klaseko
- 208 CONCEPTS-SYNTAGMATIQUES klaseko

TH-LPPLko mota-kontzeptuak:

- 535 azpiklase ENTITESen
- 434 ACTIONS/EVENEMENTSen
- 173 QUALITESen
- 2 ETATSen

Ikuspegi erlazionalak datuok erakusten ditu:

- ENTITESen azpiko unitateek batezbeste 3 arku erlazional dituzte
- ACTIONS/EVENEMENTSen batezbeste 2.9 lotura dituzte beste unitateekin
- QUALITESen 2.8 lotura batezbeste
- ETATSen 1.5 lotura
- Modulu elebiduna.

ANHITZ prototipoan zentzu bakarreko hiztegi elebidunetik ateratako informazioa jaso da. Alderantzizko erlazioak ustatuz lortu da modulu horri bi zentzuko erabilera ematea. Tamainari dagokionez, EUSK-FRANTS EBan 556 unitate bildu dira, horietatik:

- 158 SORBURUKOAK motakoak
- 208 ELEBIDUNAK motakoak
- 190 HELBURUKOAK motakoak

Egoerari buruzko laburpentxo honetan beste pare bat aspektu azpimarratu nahi ditugu. ANHITZ HEBaren egungo prototipoak ez du jaso bere baitan hiztegi bat baino gehiago hizkuntzako. Hortaz, § III.4.2.2n proposatutakoak ez dira zertu gaurdaino. Bestalde, itzulpenenerako ezagutza espezifikorik ere ez da bildu HEBetan oraindik.

III.5 Gogoeta eta oharpen zenbait.

Errepresentazio-ereduaren inguruko gogoeta batzuk ekarri nahi izan ditugu atal honetara. Hasi, (Boguraev & Levin, 93)n jasotako iritzi batzuekin hasiko gara, hausnarrerako gai aproposak direlakoan. Hona bada lehena:

A network of relationships among word sense nodes —where nodes have internal structure, but the relationships hold over entire nodes and not over fractions of them— is topologically equivalent to the two-dimensional matrix we introduced earlier as an abstract model of a lexical database.

Hitz horiek (93:336) ANHITZ HEBa bezalako sare kontzeptualen eta database lexikalen arteko baliokidetza azpimarratzen dute. Zergatik, orduan, ezagutza-baseak inplementatu? Datu-baseekin ez ote genuen aski? Erantzuna, lehendik ere adierazi duguna, nabari samarra da, baina autore horiei kopiatuko diegu (337):

Given a word and its class membership, it should be possible within such an LKB model to infer not only a set of properties associated with this sense, but also related senses and their properties.

Horra, espero genuena, inferentzi geruza da ANHITZ HEBari izaera propioa ematen diona, datu-base lexikal erlazioaletatik bereizten duena. Beraz, ez bedi pasa oharkabea inferentzi geruzaren garrantzia.

Boguraev eta Levin-ek zalantzan jartzen duten beste kontu bat hiztegien rola da. Beraien esanetan (335): *"If an LKB is to be instantiated solely on the basis of the information available, either explicitly or implicitly, in MRD sources, the resulting data structure is going to be significantly impoverished, since it will lack essential generalizations which are missing from dictionaries"*. Uste horrekin bat gatozenez, gure proposamenean hiztegiak datu-iturri dira ezagutza-baserako eredu baino areago, eta horregatik markatu dugu bestelako informazioa —bereziki itzulpena zuzendutakoa— eskuratzeko premia.

Diote Boguraev eta Levin-ek beste pasarte batean (337):

We propose that the representational framework of a computational lexicon should explicitly indicate the semantic class membership of each word sense it includes, as well as the syntactic properties associated with that word sense.

Ados. Bai propietate sintaktikoak, bai klase semantikoak dira adierazi beharrekoak. Are:

And when a new syntactic property is posited, its values can be computed for each relevant semantic class and then inherited by the members of these semantic classes; they need not be computed for each individual word sense.

Gureari begiratuta, STRUCTURES EBa bada, hein batean, klase semantikoen eta morfosintaktikoen egitura; eta THESAURUSEko adierek heredatzen dituzte klaseari dagozkion ezaugarriak, direnak direla. Esan behar da, alabaina, klase semantiko horiek finago beharko luketenaren susmoa dugula, hau da, STRUCTURES EBaren eraketan klase gehiago bereizteak mesede egingo liokeela gure ANHITZ HEBari.

Zenbait definizio errepresentatzeko zailtasunak

- Arazo klasikoak. Zernahi hizkuntza eta hiztegitan aurkitu ohi direnak dira, gehienetan, lokailu edo erlatore zenbaitek zer esan nahi duen garbi ez dagoenean:
 - "eta", "edo", ",", ";", "lokalak
 - "bereziki"
 - "ikusi"
- Euskarari estuki lotuak. Egitura sintagmatiko batzuek adierazten duten erlazio lexikal semantikoa finkagaitza denean:
 - "-ko kide", "-ko banako", "-ko bakoitza"
 - "-tzat hartua"
- Erabilera eliptikoak. Maiz, zorigaitzez maizegi, genus-a eliptikoa izaten da:
 - "duena", "dagoena" bezalako esaldi-bukaerak

Gure proposamenean argi eta garbi defenditu dugu hiztegi bakoitzak bere sare, edo zabalago, unibertso kontzeptuala osatzen duela. Ontologia unibertsalarekin eszeptiko gara bi arrazoiengatik: lehena, praktikoa, hiztegietako adiera-banaketak bateraezinak direla, eta

127

bigarrena, ez dugula uste kultura desberdinen ezaugarriak —hiztegietan ere jasotzen direnak— ontologia bakarrean bil daitezkeenik.

Horrek ez du esan nahi, ordea, elkarren arteko erlazioak sortuz hiztegiak aberasten ez direnik, eta bide horretatik joz, § III.4.2.2n mapping problemari eman diogun trataera azaldu dugu. Harako hartan mapaketa erabatekoak egitea hobesten genuen, mapaketa partzialak landu gabe utziz. Hizkuntzarteko transferentziei eman diegun trataera ikusita —baliokidetza partzialak lantzeari garrantzi handia ematen diogu— pentsa liteke gure proposamena kontraesankorra dela, edo azalpenik behar duela, behinik behin.

Gure ustez, baliokidetza elebakarrak aberasgarriak baizik ez dira, horien presentziak onura ekartzen dio funtzionalitateari, baina horien gabeziak ez du funtzioen exekuzioa galarazten. Baliokidetza elebidunak, aldiz, funtsezkoak dira, premiazkoak. Horiek gabe sistema elebidunaren funtzionalitatea nabarmen murrizten da. Horra bada zergatik jaso duten trataera desberdina.

Ingurune elebiduna

Hiztegi elebidunetatik jasotako baliokidetzak errepresentatzean, maiz, sorburu- edo xede-hiztegian ez zeuden hitzak edo adierak aurkitu ditugu. Horrelakoetan erantsi egin ditugu. Baina, hori egitean gure printzipio bat auzitan jarri dugu, kontzeptuen eransketa modulu elebakarrak aldatzea ekarri baitu halabeharrez, eta gogora dezagun gure nahia zela modulu elebakarrak ukitu gabe uztea ingurune elebidunean. Printzipio-hauste hori konpontzeko bereizi egin ditugu STRUCTURESen unitate jatorrizkoak eta elebidunetik etorritakoak.

Modulu elebidunaren diseinua aberastea ere gogoan dabilkigu. Errepresentazio-ereduaren aldaketa handirik gabe, soilik S-ELEBIDUNAK egokituz, hobekuntza batzuk egin daitezkeelakoan gaude, adibidez:

- kategoriaren arabeko klase elebidunak definitzea
- dibergentzien klasea sortzea
- itzulpen-adibideak eranstea

Ezagutzaren eskuratzea

Esan gabe doa kapitulu honetan ez diogula tarterik eskaini ezagutzaren eskuratzeari, ez da eta, hori izan gure helburua. Bihoaz, hala ere, ohar hauek argibide gisa.

Euskarazko modulua eta elebiduna eskuz kargatu dira osorik. Frantsesezkoa, ostera, erdi-automatikoki osatu da, (Artola, 93; Agirre *et al.*, 94c; 94e)n azaldutako eran. Gure

taldean garatutako eraikitze-metodologia horrek hiztegi-definizioen azterketa eskatzen du eta, euskarazko definizioak lantzen ari garen arren, oraindik burutu gabeko lana dugu horrako hori (Arriola *et al.*, 95).

Itzulpenera zuzendutako informazioaren eskuratzek oraindik eta zailtasun gehiago dakartza. Azterketa linguistikoaren eta metodo erdi-automatikoaren beharra nabaria da ANHITZ HEBa horrelako informazioez hornituko bada.

IV. ANHITZ Sistemaren Geruza Estrategikoa eta Arkitektura.

IV. ANHITZ Sistemaren Geruza Estrategikoa eta Arkitektura.

ANHITZ eredu kontzeptualetik diseinura

ANHITZ eredu kontzeptuala deskribatu dugu II. kapituluan, eta esan dugu harako hartan emaitza dela eredu kontzeptuala, behatze- eta modelizatzeko emaitza, hain zuzen. Honako honetan, berriz, eredu hori bera iturburu izatea nahi genuke, ANHITZ sistemaren diseinurako oinarri.

Voß eta Karbach-ek "Making KADS Run" idatzi zuten (Voß & Karbach, 93) gure oraingo asmo beretsua zebilkiela. Izan ere, kapitulu honetan ANHITZ eredua martxan jartzeak ekartzen dituen gorabehera heldu nahi baitiegu. Eredua deskribapen soiletik — ikus § II.4.3— eredua exekutagarri bihurtzera pasa nahi dugu, eta jauzi horretan sortzen diren inplementazioari eta konputagarritasunari buruzko galderei erantzuten saiatu.

Arrazonamendu estrategikoa

ANHITZ HEBan inferentzi geruzako arrazonamenduak ikusi baditugu ere, ANHITZ sistemaren diseinua osatzeko goi-mailako arrazonatzea landu eta jasoko dugu orain: arrazonamendu estrategikoa. Maila estrategikoan atazei buruz arrazonatzen da, atazak dira oinarritzko objektuak eta beraiei buruz modelizazioan ikasitakoa erabiltzen da arrazonatzean. Estrategi mailan atazen kudeaketa eta kontrola da gunea. Nola konbinatu atazak, nola erabaki unean uneko soluziorik egokienak atazen deskribapenean oin hartuta, horrelako kontuei egiten zaie aurre strategi mailan. Hori dela eta, arrazonatze-estrategikoa meta-mailako ihardueratzat hartzen da.

Arrazonamendu estrategikotik arkitekturara

Jakina, arrazonamendu estrategikoaren trataerak errotik baldintzatzen du sistemaren diseinu eta arkitektura. Ez dago arrazonamendu estrategikoari aurre egiterik ez bada horretara egokitzen sistemaren diseinu orokorra, hots, ez badira modulu aproposak integratzen eta haien arteko erlazioak egoki tratatzen. Ez da harritzekoa, hortaz, KADS hurbilpeneko strategi geruza eta sistemaren arkitektura elkar-lotuta egotea.

Ondorengo ataletan ANHITZ sistemaren estrategi geruza aurkeztu ahala arkitekturari buruzko argibideak aurkituko ditu irakurleak. Hasi, atazetara zuzendutako arkitektura batzuen aipamenetik hasi gara —§ IV.1—, ANHITZekin zerikusia badutelako ustean. ANHITZ sistema deskribatu dugu segidan —§ IV.2—, moduluz modulu, datu-moduluak lehenik —§ IV.2.1—, eta modulu funtzionalak hurrena —§ IV.2.2—. Eta, azkenik, § IV.3n, inplementazioaren nondik norakoak bildu ditugu hitz bitan.

IV.1 Atazen Kudeaketarako Arkitekturak.

Atazei zuzendutako diseinuetan —ikus § II.4.1—, ezagutza egoki interpretatu eta erabiltzea arkitektura mailako trataera desberdinen eragile da. Edozein dela ere hurbilpena, ataza-egituren araberako helburuak finkatzeari eta horiek iristeko metodoak aukeratu eta aktibatzeari egin behar izaten diote aurre arkitekturek.

Ataza Generiko ("Generic Task", GT) ereduan (Chandrasekaran, 87, 90; Chandrasekaran & Johnson, 93) badira bi hurbilpen orokor:

- Ataza espezifikotarako arkitekturak ("task-specific architectures"): GTen aurreneko belaunaldiak molde honi egokitzen zaizkio. Atazei buruzko ezagutza modulu espezifikotan dago txertatua arkitekturan, eta ataza hori erabilteza suertatzen da beste inguruneetan. Horrela, arkitektura monolitikoak eratzen dira, atazari atxikiak eta ataza burutzeko metodo bakarra aktibatzen ahal dutenak.
- Arkitektura orokorrak ("general architectures"): malguago izan nahi dute hauek, atazei aurre egiterakoan metodo bat baino gehiago aukeratzeko adina malgu. Hurbilpen honetako arkitekturek berezko dute atazatik independente izatea, eta, malgu izango badira, gauza izan behar dute metodo desberdinak ebaluatu, horien artean bat aukeratu eta exekutatu, azpihelburuak finkatu eta prozesua errepikatzeko. Metodo-aukeraketan¹⁰⁰ bi ezagutza mota erabiltzen dira: metodoaren ezagutza —espezifikoa— eta ezagutza estrategiko orokorra. Ez dago esaterik, beraz, arkitektura orokorrekin ezagutza espezifikoa muzin egiten diotenik. Oso bestalde, ataza espezifikotarako arkitekturen ekarpenak eta metodoen aukeraketa dinamikoaren malgutasuna konbinatu nahi dira. Arkitektura orokorren eredu zabal horretara saiakera bat baino gehiago egokitzen da, besteak beste, BB1 (Hayes-Roth, 1985)¹⁰¹, Soar (Laird *et al.*, 1987), TIPS (Punch, 1989) eta DSPL++ (Herman, 1990) arkitekturak dira aipagarri.

¹⁰⁰ Chandrasekaran-ek hiru irizpide nabarmentzen ditu metodoak aukeratzeko: soluzioaren propietateak, prozesuaren mugapenak eta metodoa aplikatzeko ezagutzaren eskuragarritasuna.

¹⁰¹ Geroztik B. Hayes-Roth-ek (1995) sistema adimendun moldagarrietarako agente-arkitektura aurkeztu du, non hiru prozesuri dagozkien moduluak biltzen dituen: pertzepzio-, ekintza- eta kognizio-moduluak.

(Chandrasekaran & Johnson, 93)n Soar inplementazioen egokitasuna goraipatzen da. Agi danez, Soar arkitekturaren ezaugarri batzuk, bereziki *problema-espazioak* ("problem space") eta *azpihelburutze unibertsala* ¹⁰² ("universal subgoalng"), dira horretarako egokiera ematen dutenak. Autoreak aski esplizituak dira (254)n:

In the Soar implementation of GT's, we get the best of both worlds: the task-level leverage of the GT perspective and the flexibility and opportunism of the Soar architecture.

Metodo *rol-mugatzaileek* ("Role-limiting"), ataza generikoez ez bezala, ataza-egituran barreneko bideak identifikatzen dituzte. Ataza, metodo eta azpiataza guztiak mailarik baxuenean deskribatzen dira, zuzenean inplementagarriak diren azpiatazak eta hauek egikaritzeko ordena espezifikatuz. Batetik, metodo rol-mugatzaileak abstraktuak dira, domeinutik independenteak, eta hortaz, domeinu bat baino gehiagotan aplikagarriak. Bestetik, metodoen espezifikazioetan erabat zehazten da metodoaren "portaera" edo kontrol-ezagutza. Ezaugarriok arras baldintzatzen dute estrategi mailako portaera eta, hala berean, arkitektura¹⁰³.

Ataza generikoetan eta rol-mugatzaileetan oinarritutako ereduez gain badira beste hurbilpen batzuk ataza-ezagutza kudeatzeko. KADS metodologia da horietarik bat, baina KADSi ondorengo puntuan eskainiko diogu aparteko tarte.

IV.1.1 Estrategi geruza KADSen. Arkitektura-eredu bat.

Kapitulu-sarreran esan dugun bezala, Karbach eta Voß-en hitzak erabiliz orain (93:721), "the strategy layer of KADS is supposed to control and possibly repair the activities being modeled by the lower layers".

Hori horrela, domeinu-, inferentzi eta ataza-geruzak kudeatzeko kontrol-sententziak eta geruzen arteko konexioak landu beharra suertatzen da KADSen. Baina, egia esan, aspektu hauek gutxitxo jorratu dira. Oro har, jakinduri ereduaren espezifikazioa ezezik exekugarritasuna ere aintzat hartzen duten lengoaiak dira estrategi geruzaren gorabeherari aurre egiten dietenak. OMOS, MoMo, FORKADS, KARL eta (ML)² lengoaiak ez dute eraikitzaile propiorik estrategi geruzarako, nahiz eta geruza horren zenbait egiteko nola edo hala lantzen duten. QIL lengoaiak estrategi eta ataza-geruzak batera biltzen ditu: ataza-

¹⁰² Problema-espazio eraikuntzak ondo babesten du atazaren, metodoaren eta metodoa burutzeko ezagutzaren arteko harreman estua. Bestalde, azpihelburutze unibertsala era aberatsa da metodoak aukeratzeko exekuzio-denboran, gerta litezkeen impasse edo arazoak emendatuz.

¹⁰³ Ez bitez oharkabea pasa metodo rol-mugatzaileen berezitasun hauek, ANHITZ sistamarako hurbilpenak izango du eta, horien antzik.

deskonposizio finkorik gabe, azpiataza sekuentziak sortzen ditu dinamikoki, horrela estrategi mailako zeregina betez.

Ezbairik gabe, MODEL-K da estrategi geruza eratzeko egin den saiorik finena. Ikus dezagun zertan den saio hori.

MODEL-K: estrategi geruza

Arrazonamendu estrategikoak nagusiki kontrola ukitzen badu ere, lengoaia honen estrategi geruzan beste aspektu batzuk ere jaso nahi izan dira, dela problemaren zailtasun eta ebazgarritasunaren azterketa, dela problemaren ebazpenerako iturburuaren kudeaketa, problema sinplifikatzeko edo eraldatzeko gaitasuna, denbora-mugapenen kontrola, etab. Ataza-egituren exekuzioari dagozkion iharduera horiek guztiak meta-portaera adierazletzat har daitezke. Hortik dator MODEL-K lengoian arrazonamendu strategikoa lantzeko meta-maila erabiltzea.

Meta-maila horren domeinua ez da beste ezagutza-basea baizik, hots, estrategi-geruza sostengatzen duten objektuez osatutako basea. Domeinu horren gainean, KADS tradizioari eutsiz, inferentzi eta ataza-geruzak eratu dira MODEL-Kn eta, hara non, meta-mailan, hiru geruzako antolamendua errepikatzen den.

Planteamendu horren harian *problema-ebazpide islatzailea* ("reflective problem-solving") modelatu da REFLECT proiektuan (Reinders *et al.*, 91), honelaxe:

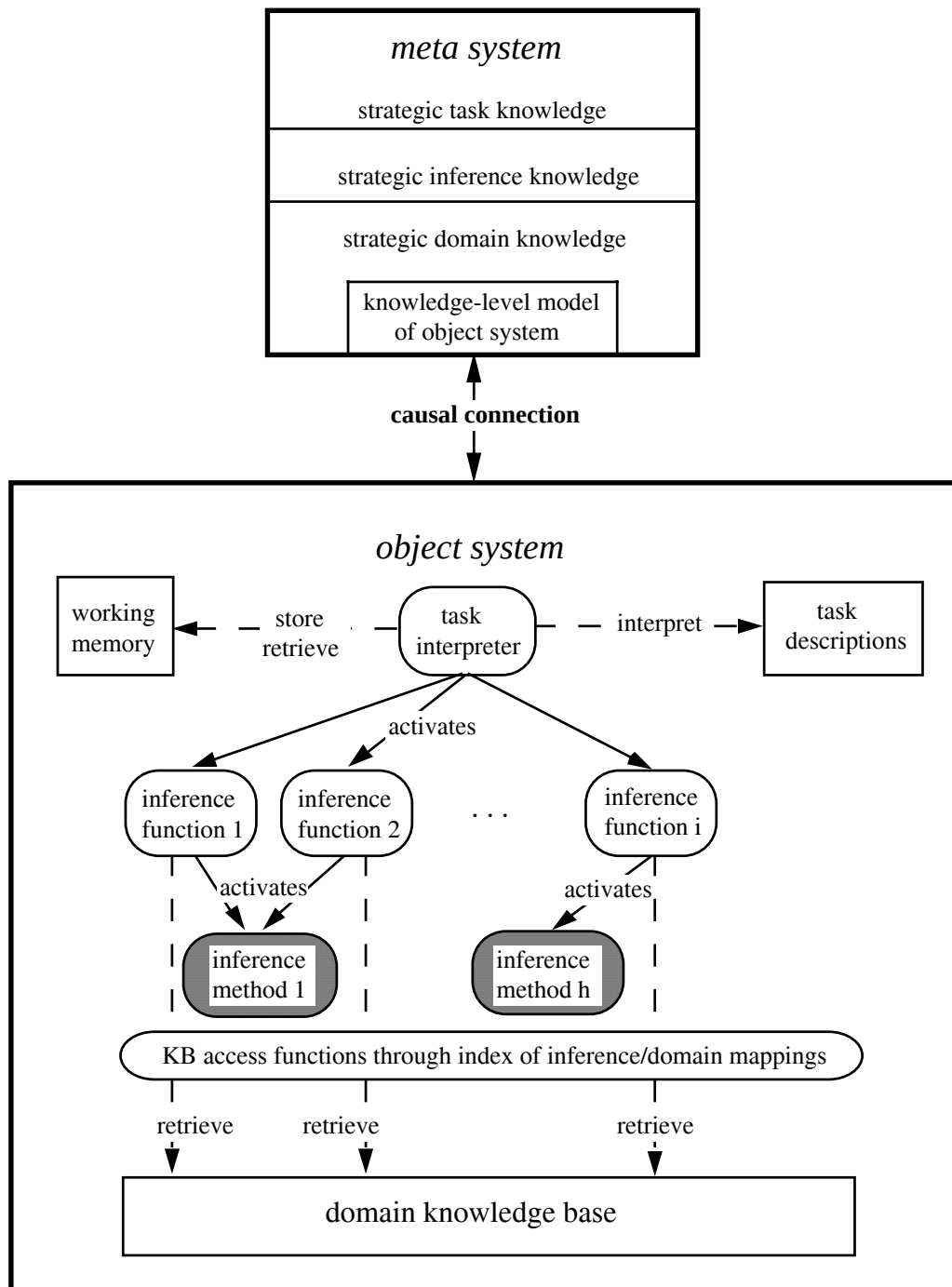
- Objektu-sistema hiru geruzatan modelatzen da: domeinu-, inferentzi eta ataza-geruza.
- Meta-mailan modulu islatzaile bat edo bat baino gehiago espezifika daitezke, eta hauek, berririo ere, hiru geruza horietan atontzen dira, domeinu-geruzak objektu-sistemaren islak direla.
- Modulu antolatzaile batek definitzen du moduluen arteko kontrol-fluxua.

Egituraketa islatzaileak bide zuzena ematen du ataza-egituraz balorazioak egiteko, eta komeni denean egitura hori doitu eta egokitzeko. Baina, aukera hori aprobetxa dadin, badira diseinu-mailan babestu beharreko propietateak. Batetik, objektu-sistema dinamikoki egunera daitekeela kontuan izanda, sistema horretako edozein aldaketak isla eragin behar du behetik gora, objektu-sistematik meta-mailara, eta halaber islatu behar dira problema-ebazle islatzaileak sortzen dituen aldaketak goitik behera. Bestetik, bi mailen arteko konmutazioak planteatzen duen kontrol-arazoa landu behar da. Konmutazio-paradigma desberdinak erabil litezke:

- Meta-simulazioa: kontrol guztia meta-mailan dago.

- Komunikazio asinkronoa: meta-sistema eta objektu-sistema batera doaz, paraleloan exekutatuz.
- Krisi-kudeaketa: objektu-sistemak lan egiten du krisiren bat detektatu arte, eta orduan kontrola meta-sistemari pasatzen dio, hark egin beharreko aldaketak egin ditzan.
- Isla-eta-ekintza: objektu-sisteman inferentzi pauso bat emandakoan kontrola meta-sistemari pasatzen zaio.
- Azpiatazen gestioa: meta-sistemak, meta-atazak ezezik, objektu-sistemako atazak ere aktibatzeke eskumena dauka. Onartzen da meta-atazek objektu-sistemako azpiatazei dei egitea.
- Kanpotiko konmutazioa: kanpoko antolatzaile batek kontrolatzen du meta-sistema eta objektu-sistemaren arteko konmutazioa. Paradigma hau da REFLECT proiektuan erabili dena.

IV.1 irudiak KADS familiako arkitektura islatzaileen ikuspegia ematen du, non nabarmena den ezagutza estrategikoa bereiz operazionalizatzen dena, meta-sistema batean.



IV.1 irudia.- Meta-mailako arkitekturaren ikuspegi orokorra, (Schreiber *et al.*, 93:141)tik hartua.

Hurbilpen honen baliagarritasuna eta orokortasuna azpimarratzen da (Reinders *et al.*, 91)n, arkitektura ezagun batzuen —NEOMYCIN/HERACLES, MOLE, SOAR, ...— ezaugarriak egoki jasotzen direla argudiatuz¹⁰⁴.

¹⁰⁴ Adibidez, SOAR arkitekturako "impasse"en emendatzea gertu dago kontzeptualki meta-mailaren egitekoetatik.

IV.2 ANHITZ Sistemaren Arkitektura.

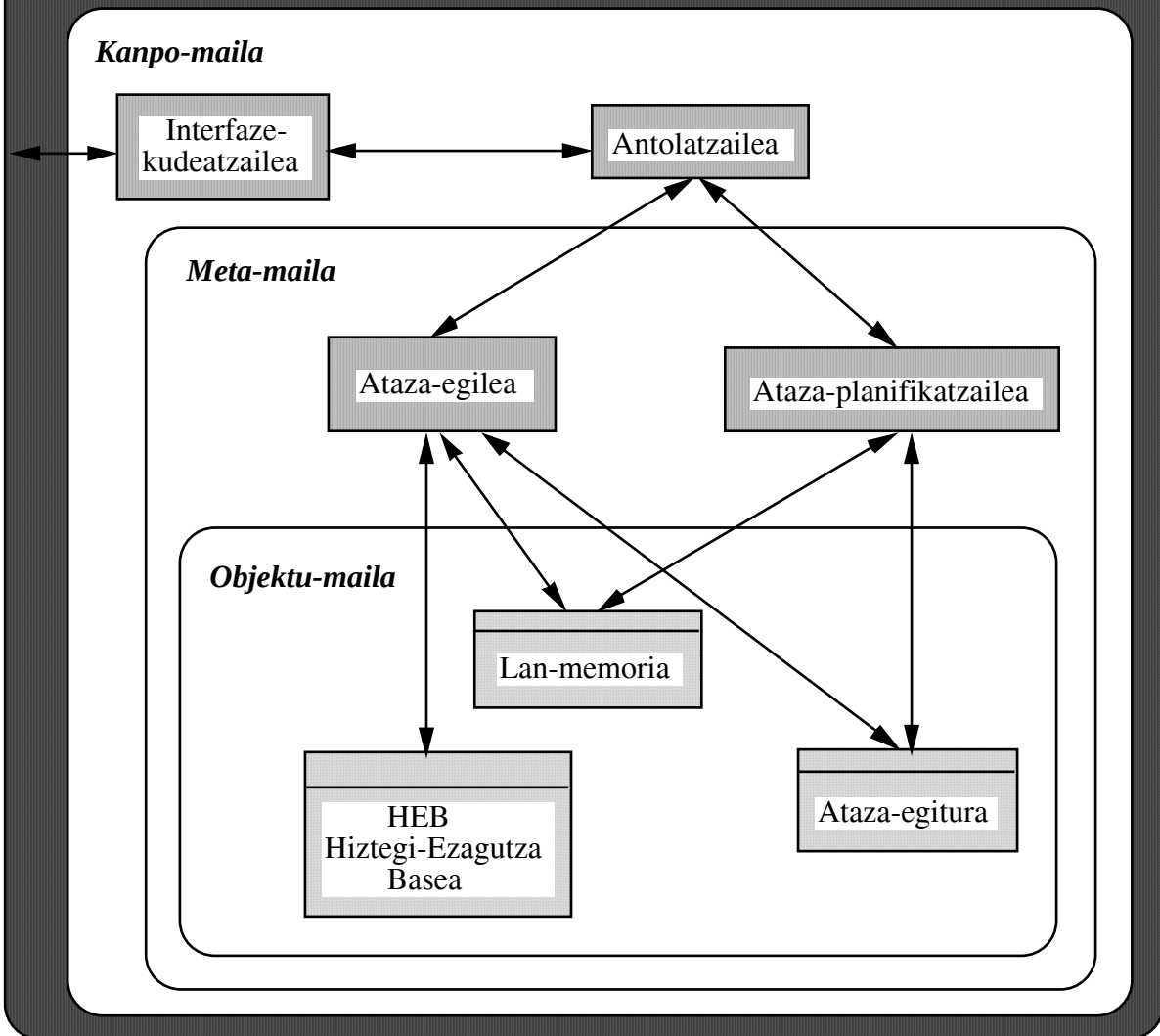
Interesgarri iritzita eman diegu begiratu azkarra atazen gaineko estrategi ezagutza sostengatzen duten arkitektura batzuei; izan ere, horien itzalpean diseinatu baita atal honetan deskribatzera goazen ANHITZ sistemaren arkitektura. Hala ere, itzalak itzal, ANHITZ arkitekturak baditu tasun bere-bereak. Horren erakusle, hona hemen ezaugarri batzuk:

- *Espezifikoa*: erabilpen jakin bati zuzendua, § Ven deskribatuko den funtzionamenduari ondo eusteko doian diseinaturik. Ez da aurrikusten liburutegi orokorren edo modulu generikoen erabilpena.
- *Berrerabilgarria*: ANHITZ arkitektura erabilgarria da beste inguru eta zereginetan, derrigorrean itzulpen lexikalera mugatu gabe.
Bi ezaugarriok, espezifikotasuna eta berrerabilgarritasuna, elkarren aurkakoak dirudite baina maila desberdinekoak dira: sistema espezifikoak da arrazoi estrategikorako aukera espezifiko bat markatzen duelako —ez da § IV.1en ikusi ditugun arkitektura batzuk bezain orokor—, baina aldi berean aukera hori erabileremu desberdinetan izan daiteke egoki eta, hortaz, berrerabilgarria ere bada.
- *Modularra*: batetik, ezagutza-baseak eta modulu funtzionalak ondo bereiziak daudelako; bestetik, ezagutza-motak modulu banatuetan kokatzen direlako. Diseinuak ondo babesten du KADS eredu kontzeptualetako mailaketa.
- *Hibridoa*: ez da ataza generikoen estiloko arkitektura orokorra, baina badu, haien antzera, metodo desberdinen artean aukeratzeko malgutasuna. Ukan, badu bestetik, arkitektura rol-mugatzaileen ezaugarriak, bai ataza-egituraren barrena bideak korritzen direlako, bai atazen kontrol-ezagutza erabat zehaztu behar delako. Nolanahi ere, nonbaitetik hurbil egotekotan, KADS moldeko arkitekturetatik dago hurbil ANHITZ, eta zehatzago, MODEL-Ktik. ANHITZen ere ezagutza estrategikoa meta-mailan lantzen da, eta kanpo-antolatzaile batek kontrolatzen du moduluen arteko konmutazioa¹⁰⁵.

Gatozen, bada, ezaugarriak gainbegiraturik eman ostean, ANHITZ arkitekturaren barrualdea arakatzera. IV.2 irudiak agerian uzten ditu modulu eta erlazio nagusiak.

¹⁰⁵ Nahiz eta ez den inola ere arkitektura islatzailea.

ANHITZ arkitektura



IV.2 irudia.- ANHITZ arkitektura.

Diseinu horretan zazpi modulu desberdintzen dira, lau modulu funtzional eta hiru datu-modulu. Datu-moduluak honako hauek dira:

- HEB, Hiztegi-Ezagutza Basea: ezagutza lexikal eta semantikoaren gordailua. ANHITZ eredu kontzeptualeko domeinu- eta inferentzi geruzak bertan daude.
- Ataza-egitura: atazei buruzko ezagutza biltzen da modulu honetan. § II.n deskribatu dugun itzulpen lexikalaren modelizazioak, ataza-geruzak, hemen aurkitzen du kokalekua.
- Lan-memoria: iharduneko gorabeheren biltokia da, erabiltzailearen eskariak eta atazen exekuzioak aldarazten dutena.

Modulu funtzionalak, berriz, hauek dira:

- Antolatzailea: sistemaren gainbegirale eta kudeatzaile nagusia da. Kontrol- nahiz datu-fluxuak gobernatzen ditu.
- Interfaze-kudeatzailea: sarrera/irteera maneiatzen du, bereziki erabiltzailearekiko elkarriketa.
- Ataza-planifikatzailea: modulu honek erabakitzen du goi-ataza bat egiteko zein azpiatazatatik hasi eta zein azpiataza-sekuentzia jarraitu, horretarako lan-memoriak erakusten duen egoera aintzat hartuz.
- Ataza-egilea: eginkizun bikoitza du. Batetik, atazak abian jartzeaz arduratzen da, eta bestetik, lan-memoria kudeatzen du, atazen exekuzioko gorabeherak memoria horretan gordez.

IV.2.1 Datu-moduluak.

Informazioaren edo ezagutzaren biltoki hauek modulu funtzionalek atzitzen dituzte: modulu atzituak dira, ez atzitzailleak, ez baitute ahalmen operatiborik. Modulu horietan dago atazen egikaritzapenerako behar den informazio guztia, hala domeinuari lotua, nola atazen exekuzio-moduari buruzkoa. Informazioa era deklarati boan errepresentatzen da nagusiki, framet an oinarritutako errepresentazio-eskema erabiliz.

Badira datu-moduluen artean informazio estatikoa dutenak, hau da, behin sistemaren eraikuntzan sortuta, ez atazen exekuzioek, ez erabiltzailearekiko elkarrekintzek aldatuko ez dituztenak. Modulu estatikoak dira ANHITZ sisteman HEBa¹⁰⁶ eta ataza-egitura. Horiek ez bezala, lan-memoria informazio dinamikoz osatzen da, hots, aldagarria da eta aldakorra izan behar luke.

IV.2.1.1 Hiztegi-Ezagutza Basea.

III. kapitulu an deskribatu dugu sistemaren ezagutza-modulu hau. Gogoan izatekoa da bi mailako ezagutza errepresentatzen dela HEBan: domeinu- eta inferentzi geruzetako a.

Inferentzi geruzan dedukzio dinamikorako mekanismoak eta oinarritzko funtzioak biltzen dira, besteak beste. Pentsatzerik badago inferentzi geruzaren mekanismo horiek ez datozela bat oraintsu aipatu ditugun informazioaren izaera deklarati bo eta estatikoarekin. Baina, esan behar da dinamikoki deduzitutako ezagutza ez dela gordetzen ezagutza-basean eta, beraz, HEBak lehengoan segitzen duela dedukzioen ondoren.

¹⁰⁶ HEBari buruz grosso modo esandako hori § IV.2.1.1 en ñabartzen da.

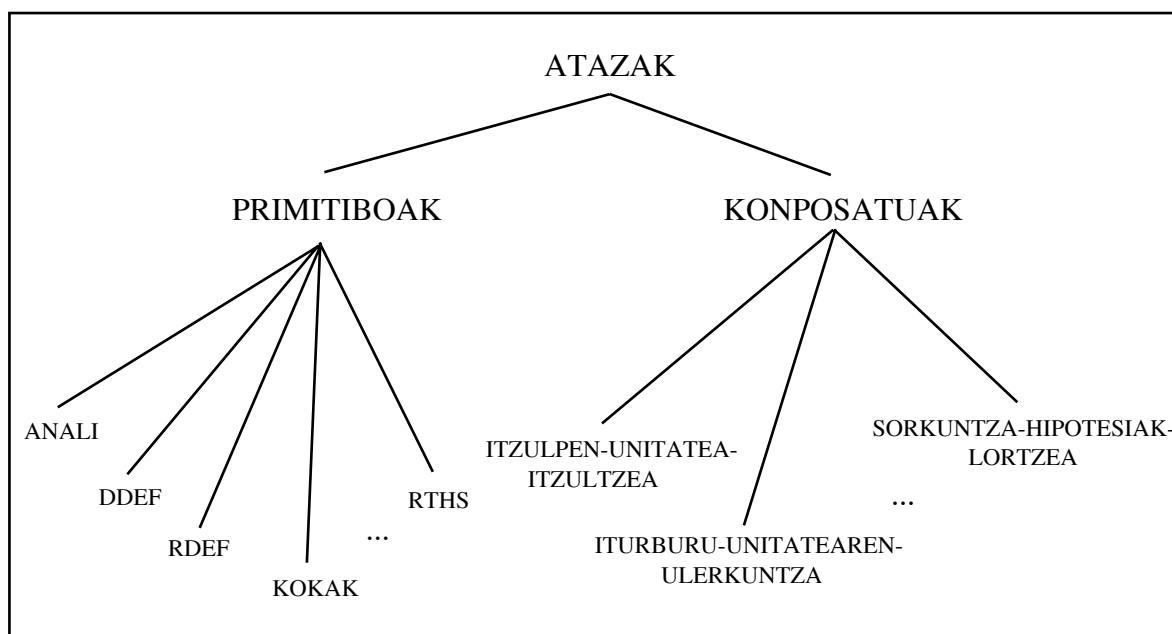
IV.2.1.2 Ataza-egitura.

Ataza-egituran deskribatzen da unitate lexikalen itzulpena eta prozesu horretan hiztegia noiz eta nola erabiltzen den. II. kapituluari ataza-ezagutzaren eskurapenean erabili dugun metodologia, adierazpidea eta espezifikazioa xehetu dira. Harez geroztik ikuspegi zabalagoaren jabe egin garenez, ataza-ezagutzak gainerako geruzekin dituen harremanak ere ukitu nahi ditugu orain, atazen espezifikazioa osatu eta aberasteko.

Atazei buruzko ezagutza adierazteko CML lengoaia erabili dugu —ikus § II.4.2—, lengoaia hori espezifikazio egituratu eta finduak egiteko aproposa baita. Espezifikaziotik diseinura igaro garen honetan ez dugu, haatik, CML lengoaia albora utziko. Bi printzipio argiri atxiki gatzaizkie ataza-egitura diseinatzerakoan:

- ataza-egituraren oinarritzko unitatea ataza da, konposatua edo primitiboa¹⁰⁷.
- ataza-ezagutza CMLk finkatzen dituen klausuletan antolatzen da.

Printzipio horiek erakutsitako bidetik egitura hierarkikoa paratu da, atazak framentan errepresentatuz eta CML-klausulei slotak egokituz. Atazen arteko erlazio hierarkikoak sub-tasks eta parents klausulen arabera osatu dira.



IV.3 irudia.- Ataza-egituraren antolamendu hierarkikoa.

¹⁰⁷ Ataza konposatua da azpiatazatan deskonposatzen bada, eta primitiboa soilik inferentzi geruzako oinarritzko funtziotan xehetzen bada.

Ataza-unitateetan antolatzen den ezagutzak izaera deskribatzailea du berezkoa. Alabaina, § II.4.3.3n esandakoa berretsiz, slot batzuk, deskribatzeko ezezik, kontrolatzeko ere erabiltzen dira¹⁰⁸. Kontrol-ezagutza hori, operazionalizazioari dagokiona, ANHITZ ereduko beste geruzek —bereziki estrategi geruzak— interpretatu eta ustiatzen dute, eta hala zertzen da geruzen arteko harremana. Kontrol-slot horiek § II.4.1en aurkeztutako acceptance-conditions (onartze-baldintzak) eta control-structure (kontrol-egitura) klausulei dagozkienak dira. Duten garrantziarengatik beta hartuko dugu klausula horiei buruzko argibideak emateko.

IV.2.1.2.1 Onartze-baldintzak.

Ataza martxan jarriko bada gertatu behar duten aurreko baldintzak finkatzen dira slot honetan. Itxuraz, rolen gaineko formula logikoak dira baldintzok, hau da, emaitza boolearrak bueltatzen dituzten formulak. Ohiko interpretazio hori eman lekieke onartze-baldintzei, baina egia esan, hurbilpen gisa balio badezake ere, ez dator bat egiazko erabilerarekin, orain ikusiko dugunez.

Ataza-planifikatzaileak onartze-baldintzen arabera erabakitzen du zein den aktibatu beharreko ataza. Hori horrela, onartze-baldintzen emaitza boolearrak ez lirateke guk nahi bezain diskriminatzaile, emaitza boolearrek aukeren espazioa bakarrik bitan banatzen baitute, ataza aktibagarriak eta aktibaezinak bereiziz. Ezerk ez du bermatzen ataza aktibagarrien multzoa bakuna izango denik, eta horixe da ataza-planifikatzaileak nahiko lukeena. Arazo horren aurrean bi irtenbide daude: bata, onartze-baldintzez gain beste heuristiko edo mekanismo diskriminatzaileak erabiltzea; bestea, onartze-baldintzak aberastea atazen artean lehentasun-banaketa finagoak egiteko. Noski, bigarren aukera horrek halabeharrez eskatzen du baldintzen izaera boolearra aldatzea.

Guk bigarren bide horretatik jo dugu. Horrenbestez, eta kontraesankorra badirudi ere, baldintzek zenbakizko balioak, *egokiera-balioak*, itzultzen dituzte. Hala, egokiera-balio handieneko ataza izango da hobetsia, betiere balio hori minimo baten gainetik badago.

"Zenbakizko onartze-baldintza" hauen idazkeran lengoaia formal berezitua erabili da. Lengoaia horretan bi tipoko baldintzak biltzen dira:

- *Derrigorrezkoak*: gutxienezko betebeharrak finkatzen dituzte, nahitaez gertatu beharrekoak. Baldintza derrigorrezkoak faltsuak direnean egokiera-balioa minimoa eta onartezina da.

¹⁰⁸ Esan liteke kontrola deskribatzen dutela.

- *Aukerakoak*: baldintza osagarriak dira, egokiera-balioa puzteko erabiltzen direnak. Ataza bat martxan jartzeko abagunearen neurria adierazten dute. Aukerako baldintzak faltsuak izan arren, derrigorrezkoak egiazkoak direnean atazak aktibagarriak dira.

Unitate_lexikaletik_hiztegi_sarrerara atazaren onartze-baldintza da hemen ondokoa¹⁰⁹:

```
( obl( testu_eredua sorburu_unitatea/=nil and
      parametro_oinarrizkoa sorburu_sarrerara =nil )
  +
  ( itzultzaile_eredua gogoko_hizkuntza =
    hizkuntza_eredua xede_hizkuntza )
  +
  (itzultzaile_eredua hizkuntza_ezaguera = txikia )
)
```

Onartze-baldintza horretako derrigorrezko aldean —obl etiketaz markatua— esaten da testu-ereduko sorburu-unitateak balioen bat behar duela eta sorburu-sarrerak, aldiz, ez duela baliorik. Hori gertatu ezean ez dago unitate_lexikaletik_hiztegi_sarrerara ataza martxan jartzerik. Gainera, aukerako bi baldintza gehitu zaizkio derrigorrezko aldeari, + etiketaz nabarmenduta: lehenengoan itzultzailearen gogoko hizkuntza itzulpen-prozesuko xede-hizkuntza izatea baloratzen da, eta bigarrenean itzultzaileak hizkuntza-ezaguera txikia izatea. Bi baldintzok betez gero unitate_lexikaletik_hiztegi_sarrerara atazari egokiera-balio handiagoa legokioke.

IV.2.1.2.2 Kontrol-egitura.

Onartze-baldintzekin batera, kontrol-egitura da atazen planifikazioan zeresan handien duen klausula. § II.4.3.3n deskribatu da gaingiroki kontrol-egituran erabilitako lengoaia. Nolanahi ere, harako hartan behar beste ukitu gabeko aspektu batzuk hemen jorratu nahi ditugu.

Sententziei dagokienean, esan dezagun badirela sententzia hutsa, ataza konposatu, ataza primitibo eta transferentzi atazez gain beste bi motako aginduak, ANHITZ arkitekturarekin zerikusi zuzena dutenak:

- *oinarrizko-funtzioei eginiko deiak*: ataza primitiboen kontrol-egituretan ageri dira eta ataza primitiboei inferentzi geruzako zein oinarrizko funtzio dagokien adierazten dute.

¹⁰⁹ Baldintza horietako objektuak ANHITZ ereduko rolak dira. Laster ikusiko dugu rol horiek lan-memorian gordetzen direla.

- *lan-memoriaren atzipen- eta aldaketa-aginduak*: lan-memoriaren ustiapena eta eguneraketa kontrolatzeko sententziak dira hauek. read eta update aginduak erabiltzen dira lan-memoriaren eremu desberdinak irakurri edota eguneratzeko¹¹⁰.

Adibide gisa, *analisi_morfologikoa* (*anali*) ataza primitiboaren kontrol-egituran funtzio-dei bat eta lan-memoriaren gaineko sententziak ikus ditzakegu:

```
control-structure:
begin
  read (out sorburu_unitatea);
  f_anali (in sorburu_unitatea) -> analisi_morfologiko_lista;
  update (in analisi_morfologiko_lista);
end
```

Kode horrek erakusten duenez, *analisi morfologikoa* egiteko sorburu-unitatea irakurri behar da lan-memoriatik, ondoren inferentzi geruzako *f-anali* oinarritzko funtzioari dei egin eta, azkenik, lan-memorian gorde *analisi morfologikoen* lista.

Kontrol-mekanismoek ere merezi dute xehetasunik gauden kapitulu honetan. § II.4.3.3n, ANHITZ eredia deskribatzean, sekuentziazioa, baldintzazko konposaketa, iterazioa, aukerako eta noiznahiko kontrol-mekanismoak ikusi ditugu. Orain, eredia martxan jartzeko kontuetan murgildu garen honetan, bada zer aipa mekanismo horiei buruz:

- *Case* aginduaren semantika:
ohiko esanahia eman diogu lehenago *case* aginduari, betetzen den lehenengo baldintzari doakion sententzia exekutatzen dela esanez. Interpretazio hori, ordea, aberastu egiten da —eta gure beharretara hobeto egokitu— zenbakizko onartze-baldintzak baliatzen baditugu. Hartara, *case* aginduko baldintzak lehian dauden onartze-baldintzen gisakoak dira. Exekuzioaldi bakoitzean, baldintza horiek guztiak ebaluatu eta egokiera-balio handienekoa hobesten da, hari doakion sententzia egikarituz. Kontrol-mekanismo honek egoera etengabe eguneratzea, baldintza guztien ebaluaketa eta atazen planifikazio dinamikoa ekartzen du.
- *Loop* iterazioaren bukaera:
loop iterazioaren irteera-baldintza B dela, B hori gertatzean *loop*-aren exekuzioa amaitu egiten da. Baina, ez dago seguru jakiterik B hori gertatuko denik, ez eta iteratzeak ezagutza-rolen egoera aldatuko duenik ere. Iterazioaren bukaera ziurtatzeko kanpo-etenak dira beharrezko.
- Aukerako eta noiznahiko kontrolak:

¹¹⁰ MODEL-K lengoaiak oso antzeko sententziak erabiltzen ditu ezagutza-rolen balioak kontsultatu eta aldatzeko.

bi mekanismo hauek nolabaiteko ez-determinismoa ekartzen dute kontrol-egiturara —ikus § II.4.3.3—. Ez-determinismoa ez da ohikoa KADS lengoaietan¹¹¹ eta ANHITZ sisteman ere, deskribatzean bai, baina exekutatzean ez dira erabiltzen oraingoz aginduok.

Adibide hartu dugun itzulpen_unitatea_itzultzea atazaren kontrol-egituran aipatu ditugun kontrol-sententzia horiek ageri dira:

```
control-structure:
begin
  [testuingurua_eskuratzea (out testuingurua, out hizkuntzak);]
  [itzultzaile_eredua_eskuratzea (out itzultzailea);]
  loop
    case
      not ?(ulertua(sorburu_definizioa)):
        sorburu_unitatearen_ulerkuntza (in sorburu_unitatea,
                                         out sorburu_definizioa);
      not ?(aurkitua(xede_kontzeptua)):
        baliokidea_bilatzea (in sorburu_unitatea, out xede_kontzeptua);
      ?ulertua(sorburu_definizioa):
        xede_unitatearen_sorkuntza (in xede_definizioa,
                                     out xede_unitatea);
      ?aurkitua(xede_kontzeptua):
        xede_unitatearen_sorkuntza (in xede_kontzeptua,
                                     out xede_unitatea);
    end-case;
    exit when ?sortua(xede_unitatea);
  end-loop;
  .-> itzul_lex (in sorburu_unitatea, out xede_unitatea);
end
```

Badira aspektu ohargarriak adibide horren gainean:

- testuingurua_eskuratzea eta itzultzaile_eredua_eskuratzea sententziak —makoen artean— aukerakoak dira eta erabiltzailearen esku dago horiek exekutatzea, ANHITZ sistemak ez baitu horien exekuzioa sostengatzen¹¹².
- itzul_lex ataza noiznahi exekuta daiteke, esan nahi baita, ez dagoela exekuzio-ordena finko batean txertatzerik. Askotan ataza lagungarriei esleitzen zaie kontrol-mekanismo hau, noiz egikaritu behar diren zehatz esaterik ez dagoenean. ANHITZ sisteman, gaur egun, ez da lantzen mekanismo hau.
- case aginduan lau hautabide desberdin biltzen dira. Lau baldintzen ebaluaketak erakutsiko digu zein den egokiena une horretan eta hari dagokion sententzia exekutatuko da.
- loop sententziak ?sortua(xede_unitatea) baldintza betetakoan izango du bukaera normala. Hori gertatu ezean, kanpo-agente batek amaiarazi beharko luke iterazioa.

¹¹¹ Soilik MoMo, (ML)² eta K_{BS}SF lengoaiak ahalbidetzen dituzte espezifikazio ez-deterministak.

¹¹² Itzultzaileari erakutsi beharko litzaioke ataza horiek burutzeak onurak ekartzen dizkiola, horrela sistemaren erantzunak zuzenagoak eta aberatsagoak izango baitira.

IV.2.1.3 Lan-memoria.

Ihardunaren egoera gordetzen du lan-memoriak ANHITZ sisteman. Dinamikoa da, sentikorra, aldaketak jaso eta sistemaren uneko prozesua zertan den islatzen duena.

Uneko egoera adierazteko lan-memoriak ANHITZ ereduko objektuen edo, nahi bada, ezagutza-rolen balioak biltzen ditu. Hurbilpen hau aski hedatua dago KADS lengoaietan: *"The dominant view on how to represent the state is clearly by the current contents of the knowledge roles"*, diote argiro (Fensel & van Harmelen, 1994:138)n.

Lan-memoriaren beste alderdi azpimarragarriak geruzen arteko komunikazioarekin du zerikusia. Ataza-geruzako onartze-baldintzetan eta kontrol-egituran oinarritzkoak dira lan-memoriako objektuen erreferentziak. Objektu horiek, aldi berean, inferentzi geruzako oinarritzko funtzioek darabiltzate, bai datu, bai emaitza bezala. Estrategi geruzak, bestalde, lan-memoriako objektuen kontsulta eta eguneraketa kontrolatzen du. Hitz gutxitan esanda, geruzen arteko informazio-fluxuak lan-memorian aurkitzen du euskarria, hantxe egoten dira datuak kontsultagai eta hara eramaten dira lortutako emaitzak.

Objektuen edukiez gain, ihardunaren ikuspegi zabalagoa edukitzeko, gertatutako gorabeherak ere gordetzen dira lan-memorian, dela martxan jarritako atazak, dela lortutako helburuak, eta abar.

Idea horiek guztiak jasotzeko lan-memoria eremuka diseinatu dugu eta, gainerako informazio-baseak bezala, frametari inplementatu da egitura hori. Hara hemen segidan eremu horien berri:

- *<itzultzaile-eredua>*:

ezagutzen oinarritutako sistema moldagarrietan ohikoa da erabiltzailearen eredua deskribatzen duten moduluak erabiltzea (Brajnik & Tasso, 94). ANHITZ sisteman itzultzailearen eredua paratu dugu, itzultzaileari buruzkoak eredu horretan gorde eta, ahal balitz, iharduna haren beharretara egokitzeko. Bost ezagutza-rol bildu ditugu eredu honetan:

- *gogoko hizkuntza*: itzultzaileak itzulpeneko sorburu- edo xede-hizkuntzen artean gogokoen duena.
- *hiztegi hobetsia*: itzultzailearen hiztegiarik gustukoena.

Ondorengo hiru rolek itzultzailearen iaioitasun edo trebeziaren berri ematen dute:

- *sorburu-hizkuntzaren ezaguera*: *bikain, ..., eskas*
- *xede-hizkuntzaren ezaguera*: *bikain, ..., eskas*
- *esperientzia*: *handia, ..., txikia*

- *<testu-eredua>*:

itzultzen ari den testuaren nolakoa deskribatzen da eredu honetan. Nahiz eta unitate lexikala izan itzulpenaren zuzeneko gaia, unitate hori dagoen testuak badu eraginik azken emaitzan, § II.3n nabarmendu dugunez. Testuaren ezaugarriok jaso ditugu:

- *mota: literarioa, teknika, ..., periodistika*
- *ezkerreko testuingurua*
- *eskubiko testuingurua*

- *<oinarrizko rolak>*:

itzulpenean zehar munta handikoak dira oinarritzko rolak, itzulpenaren nondik norakoak markatzen dituztenak eta uneko egoerak deskribatzean saihestezinak. Horrenbestez, onartze-baldintzetan eta kontrol-egituretan usu ageri dira. Hara zein diren ANHITZ lan-memoriako oinarritzko rolak:

- *sorburu- eta xede-hizkuntza*
- *sorburu- eta xede-unitatea*: unitate lexikalak
- *sorburu- eta xede-sarrera*: unitate lexikalei dagozkien hiztegi-sarrerak
- *sorburu- eta xede-kontzeptua*: sarrerei egokitzen zaizkien adierak
- *sorburu- eta xede-definizioa*
- *sorburu- eta xede-morfologia*
- *sorburu- eta xede-sintaxia*

- *<rol lagungarriak>*:

rol hauek ez dute oinarritzkoek adinako nortasunik eta esanahirik itzulpenean. Gehienak, ondoren zerrendatzen direnak adibidez, inferentzi geruzako oinarritzko funtzioen parametro gisa usatzen dira:

- *erabiltzailearen definizioa*
- *kontzeptu-arteko aldea*
- *kontzeptu-arteko harremana*
- *kontzeptu-atributua*
- *funtzio lexikala*
- *konnotazioak*
- *erabileremua*

Badira beste rol lagungarri batzuk atzera jotzeko prozesuak sostengatzeko erabiltzen direnak, horietan baitaude gordeak itzulpen-iharduneko urrats desberdinetako aukera guztiak. Rol lagungarriak dira, konparazio batera, honako hauek:

- *hiztegi-sarreraren lista*
- *kontzeptuen lista*

- *analisi morfologikoen lista*
- *definizioen lista*
- *baliokideen lista*
- *<agenda>*:

agendan itzultzaileak eragindako gorabeherak eta gertakizunak gordetzen dira. Historiaren gordailua da, batez ere, agenda: itzulpenean zehar egindakoa eta, neurri apalago batean, etorkizuneko asmoak erregistratzeko memoria. Agendako rolak ditugu honakoak:

- *uneko ataza*: martxan dagoen ataza, hasia baina amaitu gabea
- *egindako atazak*: dagoeneko burutu diren atazen historia
- *lortzeke dauden helburuak*
- *lortutako helburuak*
- *<barne-agenda>*:

barne-agendak agenda arruntaren eraketa berdina du, baina barne-agendako rolek ez dute erabiltzaileak eragindakorik islatzen, ANHITZ sistemak bere kabuz eragindakoak baizik. Barne-agendaren zentzua ulergarriagoa gertatuko da V. kapituluan, sistemaren funtzionamendua ikustean.

IV.2.2 Modulu Funtzionalak.

ANHITZ eredu kontzeptuala exekutatzen duten moduluak ditugu honakoak. Modulu hauen artean badira batzuk ANHITZ sistemaren meta-mailakoak, KADS hurbilpeneko estrategi geruzari lotuak, eta beste batzuk, aldiz, kanpo-mailakoak, sistemaren goi-kontrolaz eta kanpo-alderako portaeraz arduratzen direnak —ikus IV.2 irudia—. Meta-mailakoak dira ataza-planifikatzailea eta ataza-egilea, eta kanpo-mailakoak, berriz, antolatzailea eta interfaze-kudeatzailea.

IV.2.2.1 Ataza-planifikatzailea.

Plana, zabal hartuta, eragiketa-sekuentzia bat zein ordenatan egikaritu kontrolatzen duen prozesua da (Miller *et al.*, 1960). Planen kudeaketa, planifikazioa, adimen artifizialeko saio askoren gai izan da, eta saio horietan domeinu anitz landu eta hurbilpen desberdinak urratu dira.

ANHITZ sistemaren estrategi geruzan atazen planifikazioa landu da, bereziki ataza-planifikatzaile deitu dugun moduluan landu ere. Planifikatzeak, ANHITZen, ataza-espazioan

bide bat aukeratzea esan nahi du, hau da: goi-ataza baten burutzapena helburu hartuta, zein azpiataza-sekuentzia exekutatu behar den erabakitzea, betiere uneoroko egoerak aintzat hartu eta horien arabera jokatzuz.

Planifikazioak, horretara, badu zenbait alde bereizi beharreko:

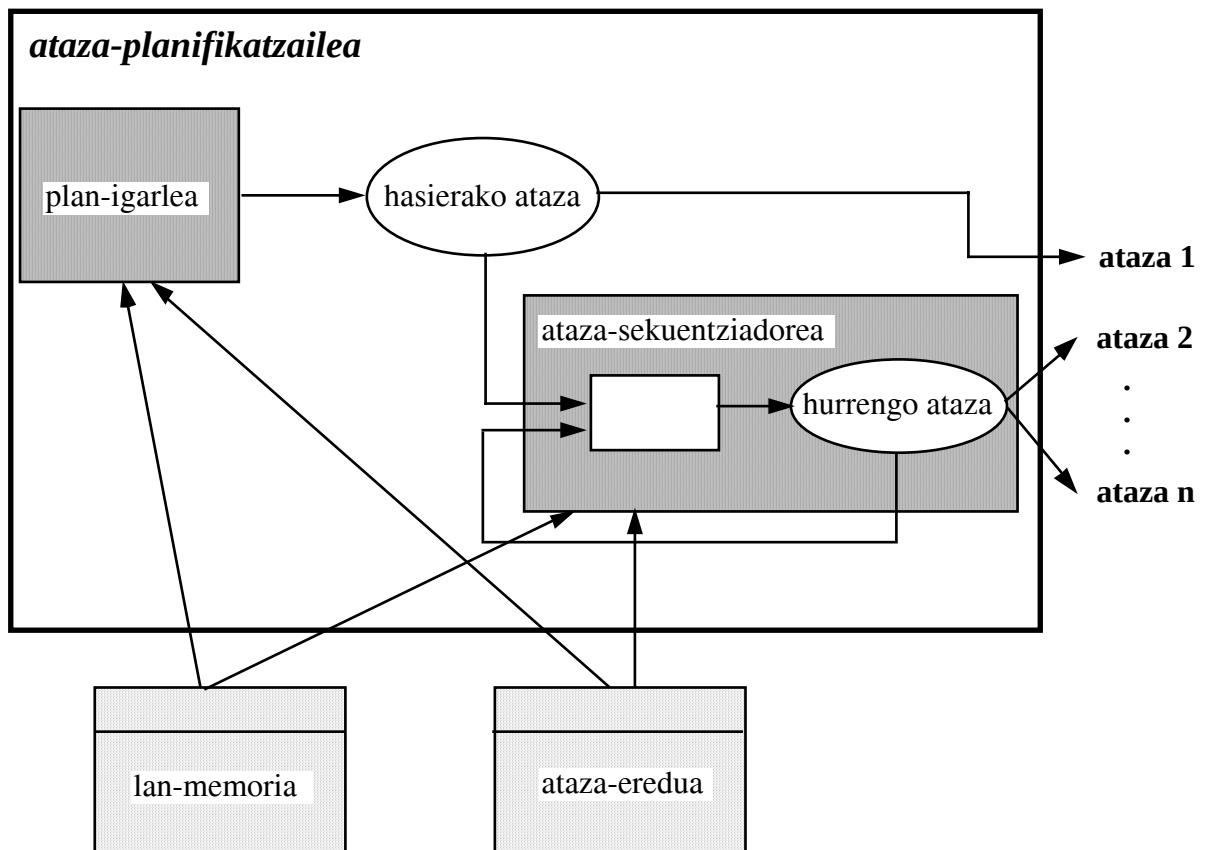
- *Helburuaren finkapena*: helburu-egoera zein den ezagutzea, alegia. Atazen deskribapenean jaso den konpetentzi baldintzak —ikus § II.4.2— adierazten du ataza bakoitzak zein egoera sortaraz beharko lukeen.

ANHITZ ereduan, alabaina, azken helburua beti bat da: unitate lexikala itzultzea. Beraz, dagoen egoeran dagoela, azken helburu hori iritsi bitartean ataza-planifikatzaileak ihardunean jarraituko du. Helburua noiz lortu den jakiteko lan-memoria kontsultatzen da.

- *Hasierako azpiatazaren aukeraketa*: nondik hasi, nola abiatu, horiek ere planifikazioaren aspektuak dira. Horretarako hasierako egoera zein den igarri eta egoera horretan gehien komeni den ataza aurkitu behar da. Aukeraketaren ardura planifikatzailearen azpi-modulu batek, ataza-igarleak, hartzen du bere gain.
- *Azpiataza-sekuentziaren osaketa*: behin hasierako azpiataza erabakita, ataza-planifikatzaileak hurrengo atazen aukeraketari ekiten dio, bata bestearen atzetik egikaritzuz sortzen den bidea helburura iristeko biderik onena datekeelakoan. Atazen aukerakera urratsez urrats egiten da, aurreko atazak utzi duen egoera kontuan hartuz hurrengo ataza zein izango den erabakitzeko.

Azpiatazen sekuentziazioa ataza-sekuentziadore deitu moduluaren egitekoa da. Ataza-sekuentziadoreak ataza-ereduko kontrol-egiturak analizatu eta kudeatzen ditu bere eginkizuna betetze aldera.

Azpimodulu horien deskribapen xeheagoari heldu aurretik begiratu bat emango diogu ataza-planifikatzailearen diseinuari —IV.4 irudia—:

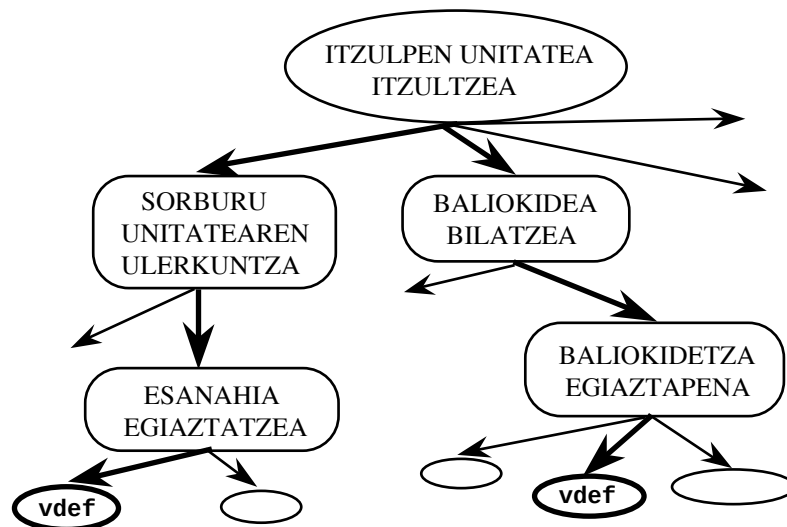


IV.4 irudia.- Ataza-planifikatzailearen diseinu orokorra.

Irudian ikus daitekeenez, plan-igarleak erabakitzen du zein den planaren lehenengo ataza, lan-memoria eta ataza-eredua oinarri hartuta. Hasierako ataza, halaber, sekuentziadorearen sarrera izango da, eta hortik aurrera, sekuentziadoreak aukeratutako atazek bi rol jokatuko dituzte: sekuentziako hurrengo atazarena eta, atzeraelikatuz, sekuentziadorerako sarrera-atazarena.

Gauden honetan komeni da ataza kontzeptuaren erabilera desberdinak ez nahastea. Ikusi dugu II.2 irudian nola errepresentatzen den atazen espazioa arbola baten antzeko deskonposizio-diagrama batez. Diagrama hartan badira behin baino gehiagotan agertzen diren atazak. Ataza-ereduan behin definitzen dira —bat dira—, baina planifikazioaz ari garela agerraldiak desberdintzea beharrezkoa da. Hots, sekuentziadoreak sortutako atazak ataza-ereduko instantziak edo agerraldiak dira. Instantzia horiek ezaugarritzeko deskonposizio-arbolaren adarrak erabili dira. Sekuentziako atazak, hortaz, arbolako adarren bidez adierazi dira.

Errepara diezaiogun deskonposizio-arbolako zatitxo honi:



(*itzulpen_unitatea_itzultzea*, *sorburu_unitatearen_ulerkuntza*, *esannahia_egiaztatzea*, *vdef*) eta (*itzulpen_unitatea_itzultzea*, *baliokidea_bilatzea*, *baliokidetza_egiaztapena*, *vdef*) adarrei bi ataza desberdin dagozkie, nahiz eta adarren bukaeran *vdef* ataza primitibo bera egon. Horra zergatik erabiltzen diren adarrak atazen planifikazioa lantzen denean.

IV.2.2.1.1. Plan-igarlea.

Sistema moldagarrietan ohikoa da sistemaren portaera erabiltzailearen edota, oro har, kanpo-agenteen ezaugarrietara egokitzea (Krause *et al.*, 93). Horretarako beharrezkoa da, ordea, agente horien asmo, behar edo planak igartzea (Carberry, 83; Goodman & Litman, 92; Villegas & Eberts, 94).

Gurean, ANHITZ sisteman, itzultzailea zertan ari den eta hiztegietatik zer jaso nahiko lukeen asmatu nahi da. Zertan ari den asmatzeko lan-memoria behatzen da, eta zer nahiko lukeen erabakitzeko ataza-eredura jotzen da, eredu horrek erakusten baititu itzulpenaren nondik norakoak.

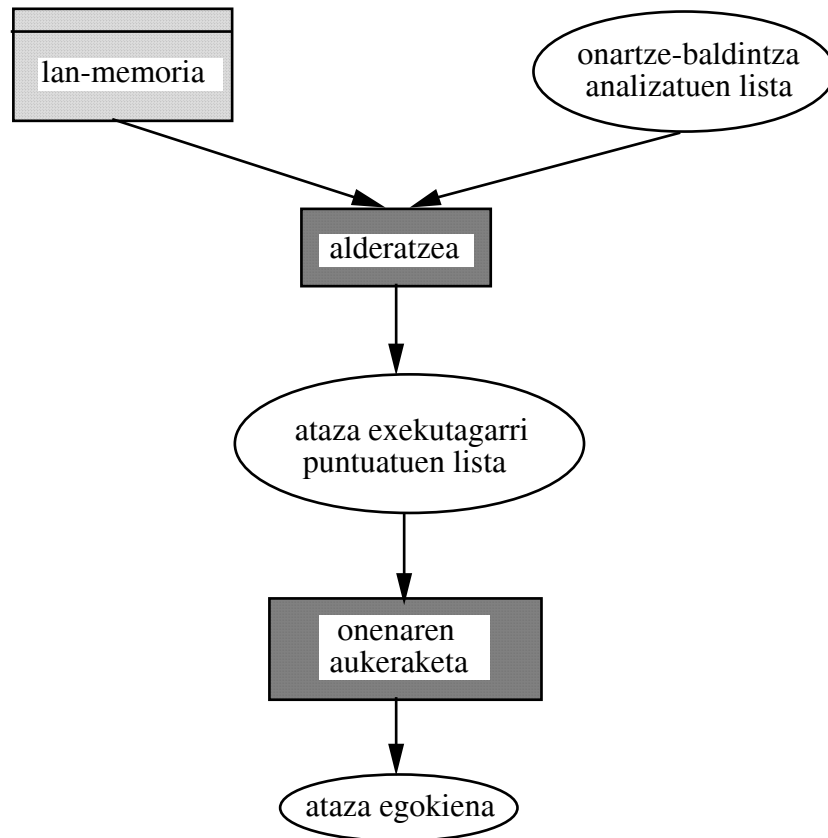
Bi dira plan-igarlearen baitako moduluak; bata, onartze-baldintzak analizatzea du eginkizun eta, besteak, atazen artean egokitasun-konparaketak egitea.

IV.2.2.1.1.1 Analizatzailea.

Analizatzaileak atazen onartze-baldintzak sintaktikoki analizatu eta egokitu egiten ditu, alderatzaileak onartze-baldintza analizatuak erabil ditzan.

IV.2.2.1.1.2 Alderatzailea.

Onartze-baldintza analizatuak lan-memoriaren egoerarekin alderatzea eta ataza egokiena aukeratzea da alderatzailearen betekizuna. Egiteko hori IV.5 irudian erakutsitakoaren arabera gauzatzen da.



IV.5 irudia.- Ataza-alderatzailearen diseinua.

Bi urratsen buruan lortzen da ataza egokiena zein datekeen erabakitzea:

1. *alderatzea*: onartze-baldintza analizatuak jaso eta baldintza horiek guztiak lan-memoriaren uneko egoeran ebaluatzen dira. Baldintza horien derrigorrezko partean gutxienezko betebeharrak adierazten dira, eta horiek bete ezean, ez da egokiera-balio onargarria lortzen. Onartze-baldintzak gutxienezko baliora ailegatzen ez direnean, horiei dagozkien atazak bazterrean geratzen dira urrats honetan bertan, ez dira-eta exekutagarriak. Baldintza onargarrien egokiera-balioa kalkulatzeko, derrigorrezko parteari aukerakoaren ebaluaketa gehitzen zaio. Saio horretan zenbakizko funtzio bat, puntuatze-funtzioa, usatzen da. Puntuatze-funtzioak, zenbat eta baldintza gehiago bete, orduan eta puntu gehiago esleituko dizkio onartze-baldintzari, eta zeharka, baldintza hori duen atazari.

2. *onenaren aukeraketa*: ataza puntuatuen lista ordenatu egiten da urrats honetan, eta puntu gehien duena da aukeratzen dena. Gerta liteke ataza batek baino gehiagok puntuaketa —egokiera-balioa— berdina jasotzea. Horrelakoetan, hainbanaketak ebazteko, ausaz aukeratzen da ataza garailea.

IV.2.2.1.2. Ataza-sekuentziadorea.

Modulu honek plan-igarlearen hasierari jarraipena ematen dio. Ataza-sekuentziadoreak hiru informazio mota darabiltza atazak bata bestearen atzetik paratu eta sekuentzia osatzeko; hara hemen informazio horien zertarakoak —ikus IV.4 irudia—:

- ataza-eredua: oinarria da, atazak nola kateatu daitezkeen finkatzen duen oinarria. Sekuentziadorea atazen kontrol-egituraz baliatzen da ataza baten atzetik zein beste ataza egin daitekeen jakiteko. Kontrol-egiturek ataza-espazioa definitzen dute, hau da, ataza-konbinazio guztien espazioa, eta sekuentziadoreak bideak urratuko ditu espazio horretan, pausoero ataza bakarra aukeratu eta gainontzeko kandidatuak baztertuz.
- lan-memoria: sekuentzia finkorik ez izateak, atazak egin ahala sekuentzia dinamikoki osatzeak, prozesuaren egoera kontrolatzea eskatzen du. Esan nahi baita, une jakin batean ataza bat baino gehiago egon daitekeela egingarri¹¹³, eta ataza kandidatuen artean diskriminatzeko egoerari buruzko informazioa erabili behar dela. Uneko egoerak, ANHITZ ereduan, lan-memoriak islatzen ditu.
- sekuentziako azken ataza: abiapuntua da, hori ezagututa erabaki daiteke zein egin hurrena.

Sekuentziadorearen arierak kontrol-egituren ulerkuntza du ardatz. Interpretatu egiten ditu, nolabait, atazen kontrol-egiturak¹¹⁴, eta horien arabera nondik nora joan eta zein atazatik zein atazatara mugitu erabakitzen du. Kontrol-egituren barreneko korritze horretan bi modulu lagungarri baliatzen ditu sekuentziadoreak:

- *kokatzailea*: modulu honi dei egiten zaio sekuentziako atzeneko atazak hurrengorik ez duenean dagoen kontrol-egituran —egitura horren exekuzioa bukaerara iritsi da—. Hurrengo aurkitzeko ataza gurasoaren kontrol-egiturara jo behar da eta han ediren ataza bukatua. Kokatzaileak ataza gurasoaren kontrol egiturara darama hurrengoaren bilaketa, atazen diagraman goranzko prozesua simulatuz.

¹¹³ Aukera horretaz ohartzeko aski da kontrol-egituretako case aginduak gogoratzea.

¹¹⁴ Gogoan izan kode algoritmikoaren tankera dutela kontrol-egiturek, non sententziak atazak diren nagusiki.

- *aurreratzailea*: hurrengo ataza atzenekoaren kontrol-egituran dagoenean erabiltzen da aurreratzailea. Noski, horretarako kontrol-sententzien —sekuentzi kateamendua, case baldintzazkoa eta loop iterazioa— ulerkuntza nahitaezkoa da. Hurrengo ataza konposatua denean, ataza horren kontrol-egiturara jotzen du aurreratzaileak, ataza-diagraman beherantz eginez.

Sekuentziadorearen iharduna hiru kasutan amaitzen da:

- helburua lortutakoan: hau da, goi-atazaren —itzulpen-unitatea itzultzea— kontrol-egituraren exekuzioa bukatzen denean.
- aurrera segitzerik ez dagoenean: ezinezkoa izango da aurrera jotzea lan-memoriaren uneko egoeran ataza guztien onartze-baldintzak onartezinak badira.
- kanpo-eten bat gertatzen denean: kanpo-mailan eragindako etenek amaiairazi egin dezakete sekuentziadorearen iharduna.

IV.2.2.2 Ataza-egilea.

Ataza-egilea meta-mailako modulu funtzionala da —IV.2 irudia—. Modulu hau ataza primitiboak martxan jartzeaz eta horien egikaritzapena kontrolatzeaz arduratzen da. Ataza-egileak objektu-mailatik/ra zuzentzen du bere ekimena: ataza-egituratik hartzen ditu ataza primitiboen kontrol-egiturak, non adierazten den zein ekintza¹¹⁵ egin atazak aurrera eramateko; ANHITZ HEBko oinarritzko funtzioei deitzen die, hartara, ataza- eta inferentzi geruzak exekuzioan erlazionatuz, eta azkenik, lan-memoria eguneratu egiten du, memoria hori uneko egoeraren isla izan dadin beti.

Esandakoaren haritik, bi azpimodulu bereiz biltzen dira ataza-egilean. Horiek ezagutzera goaz segidan.

IV.2.2.2.1. Ataza primitiboen aktibatzailea.

Funtsean, mezu-igorlea da ataza primitiboen aktibatzailea. III. kapituluan ikusi dugu metodoen bidez inplementatzen direla inferentzi geruzako oinarritzko funtzioak, bada, metodo horiei mezuak igorritz jartzen ditu martxan aktibatzaileak ataza primitiboei lotutako funtzioak.

¹¹⁵ Ekintza horiek arkitekturarekin lotura estua dute.

Modulu honen ahalmena eskusiboa da: aktibatzaileari, eta ez beste inolako moduluri, dagokio HEBko inferentziak martxan jartzea eta, halaber, emaitzak gestionatzea. Aktibatzailearen bitartez ez bada, ez dago HEBa atzitzerik.

IV.2.2.2.2. Lan-memoriaren kudeatzailea.

Kudeatzaileak bi erako ekintzak burutzen ditu lan-memoriaren gainean:

- irakurketak: ataza-egileak, oinarritzko funtzioei deitzera doanean, parametroak eguneratu egin behar ditu ezagutza-rolen uneko balioak kontsultatuz. Irakurketa-ekintzek ematen dute horretarako aukera.
- eguneraketak: oinarritzko funtzioen exekuzioek ezagutza-rolen aldaketa eragin dezaketenez, lan-memoriak aldaketa horiek sentitu eta baitaratu egin behar ditu. Eguneraketa-ekintzen bitartez jasotzen dira aldaketak.

Beste kontu bat da, ordea, zein rol kontsultatu edo eguneratu behar den jakitea. Normalean atazen kontrol-egituretan —read eta update sententzien bidez— adierazten da zein rol erabili, baina badira edozein oinarritzko funtzio burututakoan eguneratzen diren rolak ere, adibidez, agenda beti aldatuko da.

IV.2.2.3 Antolatzailea.

Meta-mailatik kanpoko antolatzaile batek kontrolatzen du ANHITZ arkitekturan sistemaren funtzionamendua. Antolatzailearen eginkizunak dira honakoak:

- Moduluen arteko kontrol- eta datu-fluxua kudeatzea.
- Etenak, bai itzultzaileak, bai sistemak eraginak, gestionatzea.
- Krisiak, erroreak, impasse-uneak tratatzea, adibidez, kontrol-egituretako loop iterazioek bukaerarik ez dutenean.

IV.2.2.4 Interfaze-kudeatzailea.

Interfazeek sistemen diseinu orokorretan duten eragina gero eta handiagoa dela (Hurley, 93), garrantzizkoa da ANHITZ bezalako sistema batean gai horri heltzea.

ANHITZen kanpo-mailako modulua da interfaze-kudeatzailea eta itzultzailearekiko elkarrekintza gestionatzea du helburu. Interfazeak sistemaren funtzionamendua sustatu egin behar du, eroso eta aberatsa gerta dadin bitartekoak jarri. Baliabide horiek onuragarri beharko

lukete bai sistemaren barrualdera begira, gainerako moduluen funtzionamendua hobetzen lagunduz (Goodman & Litman, 92), bai itzultzaileari begira, erabilera erraztuz. Beraz, beharrezkoa da sistemaren diseinu orokorraren, funtzionamendu-oinarrien eta interfazearen arteko komunztadura lortzea.

Sistemaren diseinu orokorra kapitulu honetan harrotu badugu ere, ANHITZen ikuspegi funtzionala aztertu gabea dugu oraindik. Komeni da, hortaz, V. kapitulurako, funtzionamenduari heltzen diogunerako, uztea interfazeari buruzko xehetasunak, han ulergarriagoak izango baitira.

IV.3 Implementazio-egoera eta landu beharrak.

Inplementazio-zehazkizunetara jo gabe, moduluen egoera aipatzea da honakoan gure asmoa eta, bidenabar, modulu horiek osatzeko edo hobetzeko ideia batzuk planteatzea.

Esan dezagun, bada, objektu-mailako hiru moduluak, HEBa, ataza-egitura eta lan-memoria, jadanik inplementatuta daudela:

- HEBa: ezagutza-base honen egoeraz luze ihardun dugu III. kapitulan.
- Ataza-egitura: modelizazio-prozesuan espezifikatutako ataza guztiak jasota daude ataza-egituran. Hala ere, bereziki onartze-baldintzak eta kontrol-egiturak doitu eta aberastu beharra ikusten dugu.
- Lan-memoria: esan gabe doa lan-memoriaren uneko egoera, § IV.2.1.3n deskribatua, nahi beste aberats daitekeela. Kontuan izan behar da, ordea, lan-memoriaren aberasketei probetxua ateratzeko, ataza-eredua ere egokitu egin behar dela berrikuntza horietara. Eta, komeni da betiere, aberastasunaren eta konplexutasunaren arteko oreka bilatzea.

Meta-mailan, ataza-planifikatzailea inplementatuta dago, ez ordea ataza-egilea. § IV.2.2.2n espezifikatu den ataza-egilea inplementatzeak ez luke, egia esan, alde azpimarragarri handirik. Aldiz, ataza-planifikatzailea martxan jartzeak badu bai aspektu erakargarriak, eta horrexegatik eman diogu azken modulu honi lehentasuna, ataza-egilearen kaltetan.

- ataza-planifikatzailea: modulu honen baitan definitutako puntuatze-funtzioa birdefinitzeak hobekuntzak ekar ditzake igarpenetan. Egungo implementazioan aukerako baldintza guztiei puntu kopuru berdina dagokie, baina, guk uste, baldintzen artean ere badagoela bereizketak egiterik, eta horien arabera puntuak ere alda

daitezkeela batetik bestera. Egokitzapen horretarako beharrezkoa da onartze-baldintzen sintaxia aldatzea eta ataza-egituran baldintzak doitzea.

Kanpo-mailako bi moduluak ez daude inplementatuta gaur egun. Etorkizun hurbilean horiei heldu nahi diegu, sistema osatu eta funtzionamendua aztertzeko.

- antolatzailea: modulu honen oinarrizko betekizunez gain, aurrikus daitezke beste batzuk, ez hain oinarrizkoak, baina dezente konplexuak eta beste modulu batzuen parte hartzea eskatzen dutenak. Adibidez, interesgarria dateke planifikatzaileak darabilen strategiari buruz azalpenak ematea erabiltzaileari, beste sistema batzuetan egiten den bezala (Chandrasekaran *et al.*, 89). Edota, egindako saioetatik ikastea.

V. Ikuspegi funtzionala.

V. Ikuspegi funtzionala.

ANHITZ itzultzaileari lagunduko dion hiztegi-sistema da. Horretarako zuzendu ditugu orain arteko pausoak, bai ANHITZ eredu kontzeptualaren modelizazioa, bai HEBaren errepresentazioa, bai, nola ez, sistemaren arkitektura. Hiru pauso horien ostean badugu sistemaren berri, erakutsi dugu ANHITZen eraketa, baina, egia esan, ezer gutxi hitz egin dugu sistemaren funtzionamenduaz.

V. kapitulu honetan, eraikuntza-urratsak dagoeneko deskribatuta, heldua da tenorea itzultzailearen begietatik sistema martxan ikusteko. Parada aproposa dugu aztertzeke zertan diren, I. kapitulan, motibazioez ihardundakoan, ANHITZ sistemari atxiki dizkiogun ezaugarriak. § V.1en gogora ekarriko ditugu ANHITZen tasun horiek eta, funtzionamenduoinarriak deskribatuz, sistemaren portaerari lotuko gaitzaizkio. ANHITZek noiz eta nola laguntzen duen, laguntza-estrategien azpiatalean nabarmenduko dugu bereziki.

Are, gauden kapitulan, kanpora aldeko portaerarekin batera, egokiera dugu orobat sistemaren barne-funtzionamendua jorratzeko. Hala, § V.2n sistemaren moduluak lanean jarri eta beraien arteko koordinazioa eta komunikazioa deskribatzen da. Puntu horren garapenerako bi sostengu baliatu ditugu: bata, sistemaren arkitekturari buruzko IV. kapitulua, non barne-moduluen berri ematen den; bestea, V.1 puntua, izan ere, kanpora aldeko portaera baita moduluen ariera justifikatzen duena.

§ V.3n itzultzailearen eta sistemaren arteko komunikaziora lerratuko gara. Beharrezkoa da, izan ere, ikuspegi funtzionala hizpide ukanda, interfazeari arreta eskaintzea, ez bakarrik sistemaren kanpora aldeko itxura erakusteko, baizik eta gainera, interfazea erabiltzailearen eta sistemaren portaera baldintzatzen duen modulua delako.

Portaeraren erakusgarri, erabilera-adibide batzuk paratu ditugu § V.4n. Adibide horiekin sistemaren ikuspegi funtzionala osatu eta biribildu egin nahi genuke.

V.1 ANHITZ sistemaren portaera.

ANHITZek praktikotasunaren erronkari egin nahi dio aurre. Horretarako, egoera nolakoa den eta itzultzailearen iharduera nondik nora doan, jokabide desberdinak hartzen ditu. Jokabide horiek garatzeko baliabideak ezagunak dira: ezagutza lexikal-semantikoa, itzulpen-ihardueraren eredua eta, bestalde, itzultzailearen atazak igarri eta planak erabakitze gaitasuna.

Besterik ezean, baliabide horien erabilerak ematen dio nortasuna ANHITZi eta ahalmen horiez mamitzen da ANHITZen izaera. Sistema praktiko eta atsegina izan nahita, ondokoak dira ANHITZen portaera-ezaugarriak behinenak:

- *Pasiboa eta aktiboa.*

ANHITZek itzultzailearen eskari zuzenekoei erantzuten die. Itzultzaileak eskatutakoan erantzuten ahalegintzen da ANHITZ. Pasiboa da, beraz.

Bestalde, ANHITZ agente aktiboa da: itzultzailearen iharduera behatzen du etengabe eta abagunea duenean laguntza eskaintzen du, itzultzaileak eskatu ez badu ere.

Eman lezake kontraesankorra dela batera biltzea portaera pasiboa eta aktiboa, baina hara zer dioten horri buruz C. Kemke-k (86:7):

Combining active and passive modes of aiding the user, incorporated into one system which could answer questions about the domain as well as give unsolicited advice, approaches an ideal help system.

Beraz, kontraesankorra ez, baizik eta ideala dateke laguntza pasiboa eta aktiboa konbinatzea.

- *Urguria.*

ANHITZek ez du interferentziarik eragin nahi, ez da endredamakila. Zuhur jokatzeko joera hori agerikoa da beste antzeko sistemetan ere: "...the *"discretion"* of COMFOHELP1, which from a *"nonobtrusive point of view"* is a central quality feature of an intelligent help system.", diote (Krause *et al.*, 93:276)n.

ANHITZentzat itzultzaileak du lehentasun erabatekoa, hark agintzen du uneoro eta sistemaren ariera, pasiboa nahiz aktiboa, haren menekoa da beti. Itzultzaileak inizatiba hartutakoan sistemak otzan jotzen du, egoskorkeriaz bereari eutsi gabe. ANHITZek ez du inoiz itzultzailearen okerrik zuzendu nahi, hark ez baitu sekula okerrik egiten¹¹⁶. Beraz, ANHITZen erantzunak iruzkinak edo proposamenak dira, eta ez, inola ere, obeditu beharreko baieztapen agintzaileak.

- *Adeitsua.*

Eta ez batere gogaikarria. Kontu handiz landu behar da ezaugarri hau, ez da komeni eta, lagundu beharraren lagundu beharraz, itzultzailea alferrikakoekin gogaitzea. Kategorikoa da, zentzu horretan, (Krause *et al.*, 93:258)ko baieztapen hau:

¹¹⁶ Jakina, hori lan-premisa da, errealitatea baino gehiago.

Nothing is more annoying, however, than an "intelligent" help system offering unsolicited advice which either is useless for the user or misinterprets his intentions.

- *Langilea.*

Erraz kilikatzen da eta lanean jartzeko prest egon ohi da. Itzultzaileak aukera ematen duenerako martxan jarri eta egitekoak bete artean ez dio lanari uzten. Noiz eta nola ekiten dion lanari § V.1.1en zehaztuko dugu.

- *Moldagarria.*

Egoeretara egokitzen saiatzen da. Lan-memorian gordetako egoerak darabiltza, horien arabera portatu eta erantzuteko.

- *Aurre hartzailea.*

Ahal duen guztietan itzultzailearen asmoak igarri eta aurrea hartzen saiatzen da. Planifikatzeko gaitasuna erabiltzen du zeregin horretan, itzultzaileak egingo zukeen hori aldeztu aurretik egiteko.

- *Adimentsua.*

Aurrea hartzeko gaitasuna adimenaren adierazlea da, bistan denez. Baina, horrez gain, itzultzailearen zuzeneko eskariei erantzutean ere nabari zaio ANHITZi portaera adimentsua, bereziki hiztegi-ezagutzaren ustiapenean.

Bi mailako portaera adimentsuak bereizten dira: aurrea hartzea estrategi eta ataza-mailetako portaerarekin dago lotua eta hiztegi-ezagutzaren ustiapena, berriz, domeinu- eta inferentzi mailekin.

- *Inperfektua.*

Edo, behintzat, ez derrigorrean perfektua; baina, hala ere, lagungarria. Ezaugarri hau ez da ohikoa software-aplikazioetan, baina sistema urgazleek beren egiten dute. Bereizgarri hau dela eta, aipatzekoa da (Krause *et al.*, 93:257)n esaten dena:

An intelligent help system does not have to be perfect. But it must be more helpful than alternatives like traditional passive help systems ...

ANHITZen erantzunak proposamen gisa hartzeak arindu egiten du inperfekzioaren zama, eta bide batez mesede egiten dio sistemaren erabilgarritasun eta fidagarritasunari.

V.1.1 Laguntza-estrategiak.

Jokabide bat baino gehiagoz baliatzen da ANHITZ itzultzailearen begietara oraintsu deskribatu dugun bezala agertzeko. Jokabide horietako bakoitza estrategia jakin bati dago lotua eta egoera desberdinetan aplikatzen da. Denek, ordea, helburu bera dute: itzultzaileari, ahal den neurrian, hitza itzultzen laguntzea. Hiru estrategia bereiziko ditugu:

- *Galdera-erantzuna estrategia.*

Estrategia hau erabiltzen da itzultzailearen zuzeneko eskarien aurrean, hau da, laguntza pasiboan. Horrelakoetan, itzultzaileak ANHITZen beharra izaten du eta zuzeneko galderak egiten dizkio sistemari. Eskari horiek hiztegi-kontsulta arrunten parekoak dira eta, hortaz, ANHITZ eredu kontzeptualean espezifikatutako ataza primitiboei —ikus § II.4.3— dagozkie. Ez dago inongo ataza-planifikazioaren beharrik eskari horiei erantzuteko, inferentzi geruzako oinarritzko funtzioak gauza baitira erantzunak lortzeko.

Zuzeneko galdera-erantzun komunikazioak ez du esan nahi, ordea, sistemak itzultzaileari eskariak egitea erraztu behar ez dionik. Bistan dago interfaze eroso batek dezenteko anparoa emango liokeela itzultzaileari, bai eskariak formulatzean, bai erantzunak ulertzean. Kontu horiei § V.2n helduko diegu.

- *Elkarrekintza bidezko estrategia.*

Honetan ere, sistemak itzultzailearen eskariz jartzen ditu martxan oinarritzko funtzioak, baina, galdera-erantzuna estrategian ez bezala, sistema ez da gauza funtzioaren exekuzioa burutzeko. Zailtasunen aurrean sistemak, paradoxikoa bada ere, itzultzailearengana jotzen du laguntza eske, eta horrela mamitzen da informazio-trukea itzultzailea eta sistemaren artean. Sistematik itzultzailearenganako laguntza-eskeek aukera-mugapenekin dute zerikusia: dela anbiguate-kasuen ebazpena, dela kontzeptu edo erlazio ugarien artean diskriminatu beharra, ...

Estrategia honen egokitasunak zalantzak sortzen ditu: ez ote da horrela galtzen sistemaren adeitasun eta urguritasuna? Itzultzailea gogaitzeko arriskuak ezbaian jar lezake estrategia hau, baina esan behar da itzultzailearen esku geratzen dela betiere sistemarekin elkarriketatzea. Sistemak, erantzunik jaso ezean, bi jokabide har ditzake: aurrera jo ezinean ahalegina bertan behera uztea, edota aukeren artean bat hartu eta harekintxe saiatzeari, nahiz eta ez eduki segurtasunik aukeraren egokitasunaz.

Ez dago zertan esan, bestalde, interfazearen garrantzia are eta nabariagoa dela elkarrekintzazko estrategia honetan.

- *Iharduera autonomia.*

Estrategia hau laguntza aktiboaren moldekoa da: itzultzaileak eskatu zai egon gabe, sistemak bere kabuz laguntzea erabakitzen du, itzultzaileak premia baduelakoan.

Ohikoa da, sistema urgazle aktiboetan, laguntza erabiltzailea okerbideetik doanean eskaintzea (Fischer *et al.*, 85; Breuker, 90). Horretarako, noski, antzeman egin behar zaio erabiltzailearen iharduerari eta, hark eskatu gabe bada ere, bide zuzenera eramaten saiatu. Batzuetan zuzendu beharrak eragiten dio sistema urgazleari, beste batzuetan, berriz, erabiltzailearen iharduera hobetu edo aberastu nahia izaten da bultzatzailea.

Baina, gureru bueltatuz, noiz du laguntza-premia itzultzaileak? Esana dugu aurrez itzultzaileak ez duela okerrik egiten hiztegiak kontsultatzean, edo hori hartuko dugula lan-hipotesizat behinik-behin. Hori horrela, ANHITZen laguntza aktiboek aberasketara zuzenduak izan behar dute; baina, noiz eta nola aberats litezke itzultzailearen hiztegi-kontsultak?

Galdera horren erantzunean datza sistemaren portaera autonomia. Itzultzailea kontsultak egiten hasi eta, azkeneko helburura iritsi gabe, egonean geratzen baldin bada —eskaririk egin gabe, asetuta, hausnarrean edo nora jo ezean—, egonaldi horrek eragiten du laguntza aktiboa. ANHITZek orduantxe erabiltzen ditu itzulpen-prozesuaren uneko egoeraz eta hiztegien erabilpenaz dakizkienak, eta hala, itzultzailearen rola bere gain hartuz, iruzkinak eta proposamenak aurkezteari ekiten dio. Proposamen horiek lagungarri edo osagarri gerta litezke eta, are, zenbaitetan, bere kasa lortu ezin zuena ANHITZen eskutik etor lekitoke giza itzultzaileari.

Iharduera autonomoko proposamenak bipilki aurkezten dira: ez da azalpen gehigarririk ematen, ez dira asmoak iragartzen, ez da inola ere galderarik egiten, ez zaio itzultzaileari egiaztapenik eskatzen. Informazioa eman egiten da, lortu ahala eskaini, itzultzaileak horietatik probetxua atera dezakeelakoan. Baina, haatik, erantzun horiei ulergarriak, sinesgarriak eta koherenteak izatea eskatu behar zaie.

Dena dela, estrategia honek ezin du itzultzailearen ariera eragotzi, printzipio gisa onartzen baita estrategia autonomoaren gainetik zuzeneko galderek dutela lehentasun osoa. Hortaz, itzultzaileak beti, edozein dela ere ANHITZen egoera, eskariak egiteko aukera du, eta eskari horiek iharduera autonomia eten egiten dute,

galdera-erantzun dinamikari berehalakoan ekiteko. Hau da, iharduera autonomoa besterik ezean jartzen da martxan, eta besterik ezean irauten du lanean.

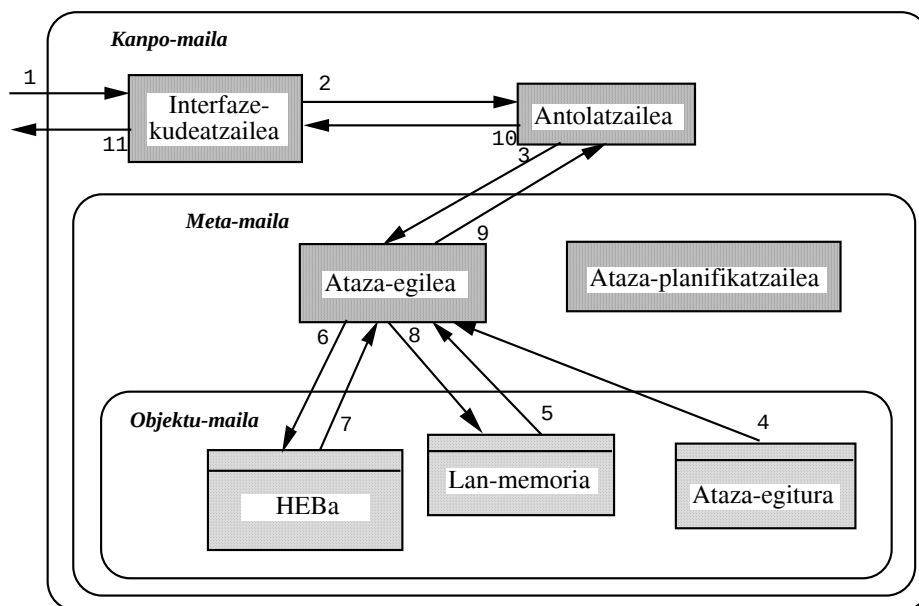
V.2 Sistemaren barne-funtzionamendua.

Portaeraren berri jaso ondoren, sistemaren barne-funtzionamendua deskribatuko dugu honakoan, § IVn azaldutako sistemaren barne-egituran oin hartuta.

Esan bezala, hiru dira ANHITZek sostenga ditzakeen jokabideak, itzultzailearen iharduerak aktibatu eta baldintzatzen dituenak hirurak. Barne-funtzionamenduaren deskribapenerako komeni da hirurak bereiz aztertzea.

Galdera-erantzuna estrategia

V.1 irudiak erakusten du kontrola eta informazioa nondik nora doazen itzultzailearen eskariei zuzeneko erantzuna ematen zaienean.



V.1 irudia.- Galdera-erantzuna strategiari dagokion funtzionamendua¹¹⁷.

Irudiko ordena-zenbakiak baliatu ditugu segidakoan, barne-urratsak xeheago azaltzeko. Hara zein den itzultzailearen eskariak eragiten duen ekintza sekuentzia:

1. Itzultzaileak eskaria —ataza primitiboa— helarazten dio ANHITZeko interfaze-kudeatzaileari.

¹¹⁷ Geziek kontrol- eta informazio-paseak adierazten dituzte. Gezien abiapuntuetan jarri diren zenbakiak ordenaren adierazle dira.

2. Eskaria antolatzaileari pasatzen zaio, hark egoeraren kontrola har dezan.
3. Antolatzaileak, eskaria jasota, ataza-egileari eskatzen dio atazaren egikaritzapena aurrera eramatea. Une horretan beste gainerako prozesuak eten egiten ditu, lehentasuna itzultzailearen eskariak dute eta.
4. Burutu beharreko ekintzak zein diren ataza-eredutik jasotzen du ataza-egileak, horretarako atazaren kontrol-egiturara joz.
5. Oinarrizko funtzioaren parametroak osatu behar direnean, ataza-egileak lan-memoriatik jasotzen ditu horiei buruzkoak. 4. urratsean kontsultatu den kontrol-egiturak erakusten du zein diren lan-memorian atzitu beharreko ezagutza-rolak.
6. Ataza-egileak oinarrizko funtzioari deitzen dio.
7. Funtzioaren emaitzak ataza-egilera pasatzen dira.
8. Exekutatutako funtzioa eta lortutako emaitzak lan-memoria eguneratzeko erabiltzen dira.
9. Emaitzak antolatzailearen esku uzten dira.
10. Interfaze-kudeatzaileak jasotzen ditu emaitzak.
11. Emaitzak erantzun bihurtu eta itzultzaileari aurkezten zaizkio.

Elkarrekintza bidezko estrategia

Funtzionamendua galdera-erantzuna estrategiaren molde berekoa da, eta ematen diren urratsak ere ez dira asko aldatzen elkarrekintzazko iharduera honetan. Galdera-erantzun zuzenekoan ematen diren hamaika urratsak oinarritzat har litezke elkarrekintzazkoan, baina, bi gertaerak aldaraz dezakete urrats-sekuentzia hori:

- 5. urratsean, parametroak egoki osatu ezin direnean:

ANHITZek itzultzaileari galdegiten dio parametro horiek instantziatzeko informazioa, hori gabe ezin baitu ataza burutu. Eskaera honela gauzatzen da:

- 5.1. Ataza-egileak antolatzaileari pasatzen dio informazio-eskea.
- 5.2. Interfazeak eskea jaso eta itzultzaileari jakinarazten dio.
- 5.3. Itzultzaileak parametroak osatzeko informazioa ematen dio sistemari.
- 5.4. Informazio hori interfazetik antolatzaileari eta handik ataza-egilera pasatzen da.

Behin parametroak osatuta prozesuak aurrera jarraitzen du 6. urratsetik aurrera lehen ikusi bezala.

- 7. urratsean, oinarrizko funtzioaren exekuzioak emaitza ez, baizik eta galdera bat bueltatzen duenean. Galdera hori, nolabait, aurrera segitzeko jakin nahi —behar?— litzatekeen zerbaiti buruzkoa da. Honela jokutzen du sistemak horrelakoetan:

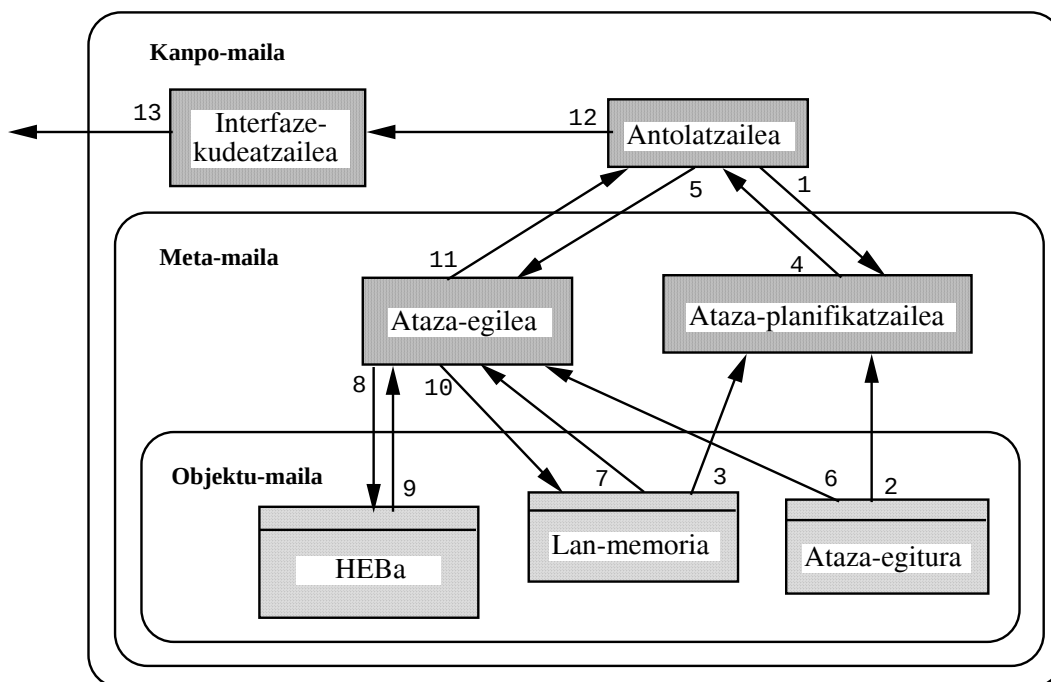
- 7.1. Galdera ataza-egilera iristen da HEBtik.
- 7.2. Ataza-egiletik antolatzaile goratzen da.
- 7.3. Antolatzaileak itzultzailearen esku hartzea eskatzen du interfazearen bitartez.
- 7.4. Itzultzaileak erantzuna idazten du.
- 7.5. Erantzuna interfazetik antolatzaileari eta ataza-egilera iragaten da.

Ekintza-sekuentzia hau burututakoan 6. urratsari heltzen zaio berriro.

Ohargarria da 7.4 urratsean itzultzailea ez dagoela behartuta¹¹⁸ erantzutera. Erantzuten ez duenean, antolatzaileak kontrola berreskuratu eta ataza-egileari ahal duen bezala ekiteko agindua pasatzen dio —dauden aukeretatik bat hartuz—.

Iharduera autonomoa

Hirugarren estrategia honetan ANHITZ sistemako modulu guztiek hartzen dute parte. Ikus dezagun nola koordinatzen diren moduluen ihardunak —ikus V.2 irudia—:



V.2 irudia.- Iharduera autonomoaren ikuspegia.

1. Antolatzaileak iharduera autonomoari ekitea erabakitzen du eta, horrenbestez, ataza-planifikatzaileari hasteko agindua igortzen dio.

¹¹⁸ Sistemaren ikusmoldearen aurkakoa litzateke itzultzailea behartzea.

Iharduera autonomoarekin noiz hasi erabakitzeko itzultzailearen azkeneko kontsulta burutuz gero iragandako denbora hartzen da kontuan. Denbora-tarte egokia iragandakoan¹¹⁹, sistemak itzultzaileari laguntzea erabakitzen du, hark zer egin edo nondik abiatu ez dakielakoan.

2. Planifikatzaileak atazen deskribapena kontsultatzen du, bereziki onartze-baldintzak eta kontrol-egiturak.
3. Planifikatzaileak itzulpenaren egoera behatzen du lan-memorian, hartara, atazen espezifikazioak aintzat hartuz, uneko ataza egokiena zein datekeen erabakitzeko.
4. Planifikatzaileak antolatzaileari jakinarazten dio zein den egin beharreko ataza primitiboa.
5. Antolatzaileak ataza hori exekutatzea agintzen dio egileari.
6. Ataza primitiboaren kontrol-egitura atzitzen du egileak, eta handik jasotzen du egikaritzapenari buruzko informazioa.
7. Oinarrizko funtzioaren parametroak osatzeko, egileak lan-memoriako ezagutza-roletatik jasotzen du informazioa.
8. Ataza-egileak oinarrizko funtzioari deitzen dio.
9. Emaitzak ataza-egileari pasatzen zaizkio.
10. Burututako ataza eta emaitzak lan-memorian gordetzen dira, nabarmenduz iharduera autonomoan lortutakoak direla.
11. Emaitzak antolatzaileari pasatzen zaizkio.
12. Emaitzak interfaze-kudeatzailerari helarazten dira, era autonomoan erdietsita daudela ohartaraziz.
13. Itzultzaileari erakusten zaizkio emaitzak. Beharbada arreta eskainiko die, eta, nork daki, agian helburua lortzeko lagungarriak gertatuko zaizkio. Txarreanean, ez die jaramonik egingo.

2.etik 13.erako urratsak errepikatu egiten dira harik eta iharduera autonomoari bukaera ematen zaion arte. Ariera honen bukaera, § IV.2.2.1.2n zehaztu denez, hiru kasutan gertatzen da: helburua lortutakoan, aurrera segitzetik ez dagoenean edota kanpo-eten bat gertatzen denean. Kanpo-eten arruntenak itzultzailearen eskariak dira. Bukaeren kudeaketa antolatzailearen esku dago.

Aipagarria da, bestalde, prozesamendu honetan ezinezkoa dela itzultzailearekin elkarriketatzea. Horrek esan nahi du, besteak beste, elkarrekintzazko iharduera ez dela eramangarria estrategia autonomoan eta, beraz, oinarrizko funtzioen exekuzioa itzultzailearen

¹¹⁹ Erabaki gabe dago zein denbora-tarte litzatekeen egokiena. (Card *et al.*, 83)n bost segundo edo gehiagoko egonaldiak hartzen dituzte helburu berriei ekiteko prestakuntza mentalaren seinaleztat, eta (Villegas & Eberts, 94)n tarte horren erreferentzia darabilte edizioko egonaldiak lantzeko. Gure kasuan, ordea, aztertu eta probatu beharra ikusten dugu denbora-tarte hori finkatu aurretik.

laguntzarik gabe burutu behar dela. Ahal bada, emaitzak aurkeztuko dira, eta ezin bada, ez da ezer egingo¹²⁰. Elkarriketatu ezinak, gainera, transferentzi atazak ez lantzea¹²¹ eragiten du.

V.3 Interfazearen xehetasunak.

Gizaki-ordenadore elkarrekintzaren arloak ("Human-Computer Interaction") interes handia piztu eta zabalkunde handia izan du azken urteotan. Emariak ere bistakoak dira: interfaze "lagunkoiak", interfaze "berezituak", interfaze "adimentsuak", multimediarako interfazeak, etab.

Gaiaren zabala eta sakona kontuan izanda, atal honetara apunte soilak baizik ez genituzke ekarri nahi, besterik ez bada ere, ANHITZ interfazearen ezaugarriak markatzeko.

Laguntza aktiborako interfazeak

Laguntza-sistema aktiboek erabiltzailearen asmoak igarri eta asmo edo helburu horien arabera erantzun ohi dute. Berezitasun horrek badu eraginik interfazean. Eskuarki, nahiago izaten da interfazeak erabiltzailearen aukerak mugaraztea, hartara errazagoa gertatzen baita asmoak igarri eta planifikatzea. (Krause *et al.*, 93:256)n interfaze grafikoen komenientziaz ondoko hau diote:

This technique restricts the range of possible non-interpretable user actions and of subsequent goals that cannot be deduced.

Era bertsuan, lengoaia naturaleko elkarrekintzak ekartzen dituen zailtasunak azpimarratuz, mintzo da P. Schefe (83: 419)n:

Despite some advances in Artificial Intelligence, natural language processing systems have not gained (and will not gain) the capabilities of recognizing intentions, context and knowledge of the user necessary for "understanding" the user's input. On the other hand, new interactive technologies such as window systems offer new possibilities of a rich and powerful interaction.

¹²⁰ Betiere hobe izango da deus ez egitea, galderekin aspertzea baino.

¹²¹ Transferentzi atazak dira, adibidez, itzultzaile_eredua_eskuratzea, gogoko_hizkuntza_aukeratzea edo konnotazio_finkatzea.

Hiztegi elektronikoen interfazeak

Hiztegi elektronikoen munduan ere interfazeek erakarri dute ateneziorik. Gai honen inguruan emandako zenbait iritzi eta proposamenek (Fox *et al.*, 80; Raymond & Tompa, 88; Marchionini, 89) hiztegi-interfazeek eduki beharreko ezaugarriak azaleratzen dituzte:

- *Dinamikotasuna*: paper gaineko sistemen egonkortasunari kontrajarrita aipatzen da ezaugarri hau. Eskariak, sinpleak zein konplexuak, aise formulatzeko aukera desberdinak ematea da garrantzitsua.
- *Sendotasuna*: nabarmenki ortografiaren aldetiko sendotasuna azpimarratzen da; hau da, kasu baterako, hitz baten idazkera zehatza ez ezagutzea ez dadila oztopo gaindiezin gerta hitz hori hiztegian aurkitzeko. Erabiltzaileari pattern-matching-eko teknikak eskaintzea inportantea da testuinguru honetan.
- *Hostokagarriak ("browsable") izatea*: hots, arin kontsultatzeko erraztasunak ematea, bilaketa-estrategia sinple eta interaktiboak sostengatuz.
- *Atzeraelikatze eta elkarrekintzarako gaitasuna*: erabiltzaile-sistema elkarriketak ahalbidetu behar dira, elkarriketa horiekin, gainera, sistema atzeraelikatzen dela.

Ezaugarri horiek jasotzeko interfaze mota desberdinak aipatu izan dira: menuen bidezkoak, grafikoak, hipertestuen teknologian oinarrituak eta lengoaia naturalezkoak¹²², besteak beste. Hipertestuak direla eta, aipatzekoa da teknika horrek oso ondo eusten diola "hostokagarritasunari", azkar eta eroso nabigatzeko aukerak ematen dituela eta erabiltzailearen atazen testuinguruan ondo integra daitekeela.

ANHITZ interfazea

Esan behar da, hasteko, proposatzera goazen interfazea ez dela era isolatuan ikusi behar. Jeneralean itzultzailearen lanpostu zabalago baten baitan integratuko litzateke ANHITZ interfazea, esan nahi baita, ANHITZ interfazearekin batera, normalki beste batzuk ere egongo direla: ediziorako, itzulpenaren prestakuntzarako, corpus-en kontsultarako, etab. Atal honetan ANHITZ interfazera mugatuko bagara ere, komeni da ikuspegi hori ez galtzea.

Komeni da orobat ANHITZ interfazeari buruzko azalpenak diseinurako hurbilpen gisa hartzea, oraindik gauzatu gabeko moduluaz arituko baikara. Ondokoak, hortaz, gure asmoen berri emateko dira, baina eginkizun daude gaur egun.

¹²² Bistan da interfaze mota hauek ez direla elkar kontrajarriak, elkar osagarriak baizik.

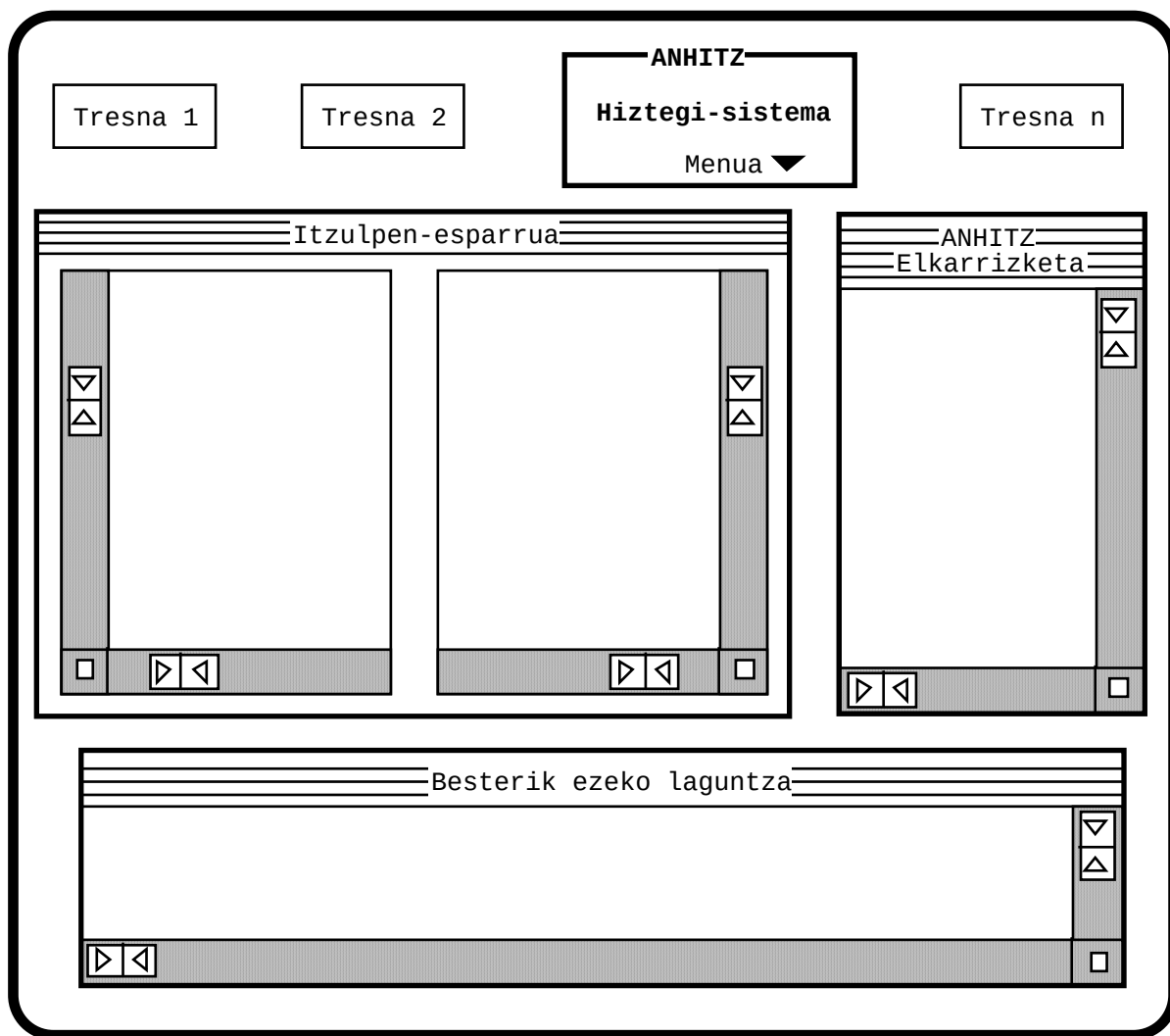
ANHITZ interfazearen nolakotasunak xedatzeko HIZTSUA sistemarako (Artola, 93: 223-235) proposatutakoa hartu dugu erreferentzia nagusitzat. Horretan oinarri hartuta asmatu den ANHITZ interfazeari ondoko ezaugarriak nabarmenduko zaizkio:

- *Grafikoa*:
 - Leihoanitzekoa: leiho batzuk aktiboak eta bestetzuk pasiboak —soilik irakurtzekoak— izango dira.
 - Menuak izango dituen: leiho berezitu horiek baliatuko ditu sistemak itzultzaileari zer egin dezakeen erakusteko.
 - Hipertestua: desiragarria da hipertestu-sistemen ezaugarriak jasotzea interfazean¹²³, bereziki hitzen agerpenak eta hitz horien hiztegiko sarrerak lotzeari dagokionean.
- *Lagunkoia*, itzultzaileari lana erraztuko diona¹²⁴. Atsegina izateko ANHITZ interfazeak gai izan beharko du:
 - Eskariak zehazten laguntzeko: besteak beste, menu dinamikoen bitartez itzultzaileari dituen aukeren berri eman, pattern-matching-eko teknikak eskaini eta, parametroak osatzean, sistemak berak lehenetsitako parametroak aurkeztuko ditu.
 - Errakuntzak antzemateko: itzultzaileak, eskariak formulatzean, ortografi hutsak edo errore kontzeptual nabariak egiten dituenen, interfazeak errakuntza horiez ohartarazi egingo luke.
 - Hitzetik adierarako bidea gainditzen urgazteko: adieren berri emanez eta aukera lehenetsiak erakutsiz.
 - Erantzunak errepresentazio ulerterrazean emateko: lengoaia naturalean ez bada ere, barne-errepresentaziotik aldendu eta itzultzaileak aise ulertuko duen idazkera erabiliko da.

V.3 irudian balizko ANHITZ interfazearen kanpo-itxura ikus daiteke:

¹²³ Hipertestuaren maneiak informazioaren antolamendu egokia eskatzen du. ANHITZ HEBaren errepresentazioak trataera hori eutsiko lukeelakoan gaude.

¹²⁴ Kontuan hartu behar da lagunkoitasun hori sistemaren mesederako ere badela. Sistemak, itzultzailea "lagundu", gidatu eta eramaten duen neurrian, aiseago igartzen ditu haren asmoak.



V.3 irudia.- ANHITZ, itzulpenera zuzendutako interfaze orokor batean.

Irudian ageri diren leiho guztiak ez dagozkio ANHITZ sistemari. Goiko aldean itzulpenarako tresna lagungarriak daude, botoi banaz adieraziak —ANHITZ horietatik bat besterik ez da—. *Itzulpen-esparrua* deitu leihoa itzultzailearen lan-leiho nagusia da, hor maneiaturko du itzultzaileak sorburu-testua eta hortxe osaturko du xede-testua. Beste bi esparruak, *Elkarrizketa* eta *Besterik ezeko laguntza*, ANHITZ tresnaren interfazeko leiho berezituak dira.

ANHITZ botoian dagoen gezi-muturrak ematen du bidea —klik eginez edo— menu nagusia bistaratzeko. Menu horretako aukerak ekinbide desberdinen adierazle dira. Horiek hautatuz oinarrizko funtzioak aktibatzea eskatzen da, batzuetan zuzenean menu nagusitik, beste batzuetan turrustan leihoratzen diren azpimenuen bitartez. Aukera horiek aurkezteko ataza-egituran finkatutako ezagutza erabiltzea komeni da, uneoro eragiketa batzuei lehentasuna emanez, baina, bestalde, edozein aukeraketarako bidea libre utziz. Menu nagusian ondoko aukerak jasotzen dira.

a. *Itzulpenari ekin.*

Aukera honen bitartez hitz bat nola itzuli asmatu ezinik dabilela jakinarazten dio ANHITZi itzultzaileak. Oharpen gisako aukera horrek ANHITZen lan-memoriaren eguneraketa dakar, ordura arteko ihardunak beste itzulpen-unitate bati zegozkion eta.

b. *Transferentziak.*

Menu-aukera honek itzultzailearen eta ANHITZen arteko informazio-paseari hasiera emateko balio du. Egia esan, beti norabide berean gertatzen da informazio-pasea: itzultzailearengandik sistemara. ANHITZ eredu kontzeptualean nabarmendu diren transferentzi funtzioak —ikus § II.4.3.1— dira aukera honen azpimenuan biltzen direnak, eta horien bidez itzulpen-ingurumariaren berri ematen zaio ANHITZi, hartara sistemaren funtzionamendua zehatzagoa izan dadin. Beraz, sistema prestatzeko eta jazteko dira baliagarri funtzio hauek.

c. *Hitz baten esanahia.*

Aukera nagusi hau hautatuz gero kontzeptuen esanahia azaltzeko funtzioak jar daitezke martxan. Zer-nolako azalpenak nahi diren finkatzeko azpimenu bat pantailaratzen da, non eskatzen ahal diren kontzeptuaren definizioa, kontzeptuaren ezaugarriak, sinonimoak, erlazionatutako kontzeptuak, azalpen-hizkuntza, etab. Eskaera horiek parekagarriak dira beste hainbat oinarritzko funtzioekin.

d. *Kontzeptu baten itzulpen-baliokidea.*

Adiera jakin baten itzulpena lortzeko bidea ematen du honek.

e. *Hitz baten bilaketa.*

Ideiak edo murriztapenak edukita, horiek betetzen dituzten hitzak bilatu nahi direnerako dago pentsatuta aukera hau; hots, azalpenak eman eta hitzak bilatu nahi direnerako. II. kapituluan espezifikatu dugun kontzeptu_bilaketa_tesaurikoa funtzioari deituz gauzatzen da.

f. *Hitzen arteko erlazioak/aldeak.*

Bi aukera jasotzen dira: elkarbide_bilaketa eta diferentzien_galdea funtzioetan oinarritutako aukerak, hain zuzen ere.

g. *Zerbaiten egiaztapena.*

Aukera honen azpimenuan sailkatzen dira egiaztapen-funtzioak, hau da, hitzaren definizioa, ezaugarriak, erlazioak, aldeak eta baliokideak egiaztatzen dituzten funtzioak.

h. *Hitza testuinguruan.*

Hitzaren portaerarekin zerikusia duten eskariak biltzen dira honakoan. Ortografia, morfologia, sintaxia edo kokakidetasunari buruzko galde edo egiaztapenek hemen dute beren lekua.

i. *Adibideak.*

Hitzaren erabilpen-adibideak erakusteko jarri da aukera hau.

ANHITZ elkarrizketa leihoan itzultzaile-ANHITZ elkarrekintzak bistaratzen dira, baina, menu-nagusian ez bezala, elkarrizketa-esparruan sistema da elkarrekintza-eskea egiten duena. Zailtasunen edo aukera anitzen aurrean erabaki beharrak eragiten du sistemaren jokabide hori.

ANHITZen hirugarren esparru propioa *besterik ezeko* laguntzarakoa da. Leiho horretan sistemak bere kabuz, inork zuzenean agingu gabe, lortutako emaitzak erakusten ditu. Irakurleioa da soilik, ez dago elkarrekintzarik.

V.4 Erabilera-adibideak.

Sistema nola erabiliko litzatekeen ilustratzeko ekarri ditugu atal honetara erabilera-adibideak. Esan dezagun garbiro, adibideotan, elkarrekintzari eta erabilera autonomoari lotutako alderdi batzuk asmatuak direla —ez denak baina—. Beraz, irudi bat jasotzeko bai, baina ez dira nahikoa sistemaren ahalak neurtzeko, ez eta ere itzultzaileari emandako laguntzaren egokiera eta zehaztasuna ebaluatzeko. Modulu guztiak inplementatu eta integratu artean ezingo dugu ANHITZen funtzionalitateari buruzko ebaluaketarik egin, nahiz eta horretara jo nahi dugun luze baino lehen.

1 adibidea: eskari sinplea

Itzultzaileak ANHITZ menu nagusitik abiatuta eta turrustan ageri diren azpimenuetan barrena¹²⁵ oinarritzko definizio-galdea egiten dio sistemari¹²⁶. Egin dezagun kontu "zein da *etxalder*en definizioa?" galdera egiten duela. *Etxalde* hitzaren adiera nagusia hartuta, galdera horren azken formulazioa —sistemak ulertzen duen lengoia—, hauxe izango da:

(ddef 'etxalde I 1 | 'herite 'th-hleh 'euskara)

¹²⁵ Aukeraketa astunegia gerta ez dadin beti dago zuzeneko botoi lagungarriak jartzea, ukitu bakarrez eskaria zehazteko. Botoi horiek maizen eskatzen diren aukerak adieraziko dituzte.

¹²⁶ Eskari horren parametroak osatzeko interfaze grafikoaren erraztasunak baliatuko ditu itzultzaileak.

Sistemak erantzun-leiho bat zabaldu eta han erakutsiko ditu definizioari buruzko azalpenak:

Etxaldea etxea da,
familia bat bizi da,
lurrak ditu,
lurrak nekazaritzako edo abelzantzako dira.

2 adibidea: elkarriketa

Itzultzaileak, azalpenekin asetu ez eta zalantzak argitzera jotzen du segidan. Jakin nahi luke, konparazio batera, zein den etxaldea eta baserriaren arteko erlazioa, eta halaxe jakinarazten dio ANHITZi, berriz ere menuen bidez. Sistemak, ordea, ez du berehalakoan erantzunik bueltatzen, ez baitaki baserri hitzaren zein adiera den itzultzaileak gogoan duena¹²⁷. Adieren berri eman eta itzultzaileari horien artean aukeratzeko eskatzen dio elkarriketa-leihoan.

Elkarriketa-leihoan baserriren bi adiera bistaratuko dira:

baserri I 1: $\bar{l}$ urraldea
baserri I 2: $\bar{e}$ txea

Itzultzaileak erantzuten baldin badu, agudoago burutuko da elkarbide-bilaketa, bestela, itsuaren moduko saiakerak egin beharko ditu sistemak, batzuetan itzultzailearen asmoei zuzen erantzunez, hurrena berriz, haren nahietatik desbideratuz.

Itzultzaileak |baserri I 2| adiera aukeratuko balu, ANHITZek ondoko elkarbide_bilaketa egingo luke:

(drap ' |baserri I 2| 'etxalde I 1| 'th-hleh 'euskara)

Eta erantzuna:

|baserri I 2| lurrak dituen da eta
|etxalde I 1| lurrak dituen da,

ondoko elkarbide-espresioari dagokionez:

((and (|baserri I 2| PREDICATION |ukan I 1#4|)

¹²⁷ Arazoa da elkarbide_bilaketa funtzio konplexua dela eta adieren artean erabakitzeak erruz lagunduko liokeela ANHITZi. Beste funtzio errazagoetan ez litzateke elkarriketarik abiaraziko hitza-kontzeptua jauzia gainditzeko: adiera bat baino gehiago dagoenean denen berri emango litzateke.

(|ukan I 1#4| PREDICATION+INV |etxalde I 1|)),
non |ukan I 1#4| kontzeptu sintagmatikoaren testua "lurrak ukan" den.

3 adibidea: prozesu autonomia

Itzultzaileak halako batean eskariak egiteari utzi egin dio, eta ez du beste hitz baten itzulpenari ekin dionik adierazi —*itzulpenari ekin* aukera menu nagusian—. Sistemak bere gain hartzen du sorburu-unitatearen itzulpena lortzea, edo behintzat, eginkizun horretarako argibideak ematea. Une horretan lan-memoriak erakusten du zein izan den prozesuaren egoera. Eman dezagun horrela dagoela memoria hori, partzialki bada ere:

<itzultzaile-eredua>:

gogoko hizkuntza: euskara
hiztegi hobetsia: HLEH ("Hauta-Lanerako Euskal Hiztegia")
sorburu-hizkuntzaren ezaguera: ona
xede-hizkuntzaren ezaguera: eskasa
esperientzia: txikia

<testu-eredua>:

mota: haur-liburua
ezkerreko testuingurua: <...>
eskuineko testuingurua: <...>

<oinarrizko rolak>:

sorburu-hizkuntza: euskara
xede-hizkuntza: frantsesa
sorburu-unitatea: etxaldean
sorburu-sarrera: etxalde
sorburu-kontzeptua: |etxalde I 1|
sorburu-definizioa: "etxaldea etxea da, familia bat bizi da, lurrak ditu, lurrak nekazaritzako edo abeltzantzako dira"
sorburu-morfologia: (izena, inesiboa, mugatua, singularra)
xede-kontzeptua:
xede-unitatea:

<rol lagungarriak>:

kontzeptu-arteko harremana:
((and (|baserri I 2| PREDICATION |ukan I 1#4|)
(|ukan I 1#4| PREDICATION+INV |etxalde I 1|))),

<agenda>:

uneko ataza:
egindako atazak: ddef(|etxalde I 1|), drap(|etxalde I 1| |baserri I 2|)
lortzeke dauden helburuak: aurkitua(xede_kontzeptua), sortua(xede_unitatea)
lortutako helburuak: ulertua(sorburu_definizioa)

Une horretan, ANHITZen modulu antolatzaileak iharduera autonomia jartzen du martxan, eta ataza-planifikatzaileak egikaritzeko den hurrengo ataza aukeratzen du. Itzultzaileak esperientzia txikia eta xede-hizkuntzaren ezaguera nolabaitekoa duenez, onartze-baldintzak ebaluatuta baliokide_galdea ("equiv") ataza hautatzen da, hark ateratzen baitu

egokiera-baliorik handiena¹²⁸. Ondoren, parametroak osatu —lan-memorian daude— eta ataza burutzen da itzultzailearen inongo partehartzerik gabe. Emaitzak besterik ezeko leihoan erakusten dira.

Besterik ezeko leihoan:

Egindako ataza: (equiv 'etxalde I 1| 'th-lppl)

Emaitzak: (('|ferme I 4| 0.5)
(|ba^timent I 1| 0)
(|maison I 1| 0))

Hiru aukera bistaratzen dira, horietatik bat baliokide zuzena: |ferme I 4|. Agindu horren exekuzioak lan-memoriako barne-agenda aldarazi du, ez ordea agenda. Orain:

```
<barne-agenda>:
  uneko ataza:
    egindako atazak:
      ddef (|etxalde I 1|),
      drap (|etxalde I 1| |baserri I 2|),
      equiv (|etxalde I 1|)
    lortzeke dauden helburuak:
      egiaztatua(xede_kontzeptua),
      sortua(xede_unitatea)
    lortutako helburuak:
      ulertua(sorburu_definizioa),
      lortua(xede_kontzeptua)
```

Ataza-sekuentziadoreak ekiten dio orain ataza-egiturak aztertu eta hurrengo ataza zein izango den erabakitzeari. Xede-kontzeptua egiaztatzen jotzen du eta, helburu horretarako dauden ataza primitiboen artean, definizio-galdeari ematen dio lehentasuna. Hala bada, besterik ezeko leihoan:

Egindako ataza: (ddef '|ferme I 4| 'deduit 'th-lppl 'frantsesa)

Emaitzak: ((and (|ferme I 4| HYPERONYME |domaine I 1|)
(|ferme I 4| ETAT |affermer I 1#1|)),
non |affermer I 1#1|= "afferme")

Itzultzaile zuhurrari zalantzak sortuko zaizkio definizio hori ikusita: benetan baliokideak al dira sorburu-kontzeptua, |etxalde I 1|, eta xede-kontzeptua, |ferme I 4|?

¹²⁸ Hor, noski, joera bat islatzen da. Itzultzaileekin egindako probek erakutsi ziguten baliokidearen bilaketa zuzenekoa dela biderik egokiena itzultzaileak zailtasunak dituenean, bereziki eskarmentu gutxi duenean edota xede-hizkuntza menperatzen ez duenean. Onartze-baldintzek erakuspen horiek jasotzen dituzte.

ANHITZek baliokidetzat hartuko ditu bi kontzeptuok, eta jarraian xede-kontzeptuaren erabilera-adibideak erakustea erabakitzen du, baina, HEBan halakorik aurkitzen ez duenez, ez du ezer pantailaratzen.

Bukatzeko xede-formaren sorkuntzan murgiltzen da. Lan-memorian testuinguruari buruzko informaziorik ez dagoenez, sorkuntza_morfologikoa ("prod") hautatu eta hori burutzen saiatuko da, horretarako lan-memoriako sorburu-morfologia baliatuz. Horrenbestez, besterik ezeko leihoan:

Egindako ataza:

(prod '|ferme I 4| '(izena, inesiboa, mugatua, singularra) 'frantsesa)

Emaita: "dans la ferme"

Horretan amaitzen da, tartean etenik izan ezean, ANHITZen iharduera autonomoa. Azpimarragarria da ANHITZen helburua ez dela nola edo halako itzulpena lortzea, baizik eta pausoz pauso itzultzaileari argibideak ematea, bere neurriko ekintzak burutuz, eta bere esku utziz erabakiak. Gerta liteke garrantzitsuago |etxalde I 1| eta |ferme I 4| kontzeptuen arteko baliokidetzaz zalantzak piztea, "dans la ferme" azken emaitza pantailaratzea baino.

VI. Ondorioak eta etorkizuneko ikerlerroak.

VI. Ondorioak eta etorkizuneko ikerlerroak.

Kapitulu honetan egindako lanari buruzko gogoetak bilduko ditugu. Deskribapenari eta azalpenari baino gehiago, hausnarketari egin nahi diogu tartea honakoan. Begi kritikoz, barru-barruan murgiltzeak sortzen duen itsukeriatik aldendu eta so egingo diogu ANHITZi. Zer ekarri duen, zer utzi duen ekartzeko, horiek hartuko ditugu hizpide, laburki, hemendik aurrerakoan.

VI.1 Ondorioak.

Lan honetan hiztegi-sistema eleanitza diseinatu da. Sistemak maila lexikaleko giza itzulpenean laguntzen du, zeregin horretan portaera pasiboak nahiz aktiboak erakutsiz. Kontsulta-sistema izateaz gain, laguntza-sistema adimendunaren jitea ere hartzen du, horrela, ANHITZek.

Ordenadorez lagundutako giza itzulpena, lexikografia konputazionala eta ezagutzan oinarritutako laguntza-sistemen arloak batera bildu dira saio integratzaile honetan. Gure ustetan aberasgarria da zinez hiru arloon elkarrenganatzea. Aberasgarria eta are beharrezkoa, giza itzultzaileari hiztegi-erabileran laguntzeak hiztegi pasibo klasikoetatik areago joatea eskatzen baitu: ez da aski euskarria aldatzea, hiztegiek ataza zailak —hitzen itzulpenak, eskuarki— egiteko tresna berezitu, aktibo eta erabilterraz behar dute izan.

Hiru ikuspegi horiek batera biltzeko, ANHITZ sistemaren garapenerako, KADS metodologiaren ildo eta ideiak baliatu ditugu. KADS, neurri batean, Softwarearen injinerutzako ideiak Ezagutzaren injinerutzara eramateko saioa da. Hala, prototipatze azkarraren ordeaz, eredu formalen erabilpena eta, horrekin batera, espezifikazio eta inplementazioaren arteko bereizketa nabarmentzen dira. KADSek, gainera, ezagutza-moten bereizketa proposatzen du oinarritzko printzipio gisa ezagutzaren eskuratzeari eta erabilpenari aurre egiteko. Proposamen hori gure eginez, lau mailako bereizketa —domeinu-, inferentzi, ataza- eta estratagi mailak— islatu da gure ihardun eta emaitzetan. Diseinu argi, egituratu eta modulatuak lortzeko bidea erakusten du KADSek gure iritzirako.

Metodologia horren ildotik, atazatan oinarritutako hurbilpena erabili da ANHITZen eraketan. Lehen-lehenik itzulpen lexikalaren eta hiztegi-erabileraren modelizazioa burutu da. Langintza horretarako baliatu den metodologia itzultzaileen behaketa zuzenekoan eta ahozko protokoloetan oinarritu da. Proba desberdinetatik jasotako datuek kontzeptualizazio, espezifikazio eta operazionalizaziorako aukera eman dute. Horren buruan modelizazio-lengoaia batean ("CommonKADS Conceptual Modelling Language") adierazi da itzulpen

lexikalari eta hiztegi-erabilerari buruz eskuratu den ezagutza, sistemaren ataza nagusiari buruzko ezagutza, alegia. Hiztegi-erabilpenaz azterketa askotxo egin den arren, ez da burutu gaurdaino, guk dakigula, datu-jasotzetik hasi eta eredu formal bat osatuz amaitu den saiakerarik. Eta, gure ustez, bazen horren premia, beste hainbat eremutan bezala, honakoan ere tratamendu automatikoak ezagutzaren eskuratze sistematikoa eta formalizazioa eskatzen baitu. § II.2n azaldu ditugun gabeziak gorabehera, gure saioak lan-ildo berria irekitzen duelakoan gaude, hemendik aurrera alde edo molde jarraipena izango duena eta etorkizun hurbilean lexikografiaren mundua astinduko duena.

Sistemaren domeinu mailan hiztegi-ezagutza landu da eta ezagutza lexikal-semanticoei eleanitzaren errepresentaziorako proposamena landu da. Esan behar da eleaniztasunaren trataera jorratu dela bereziki. Hizkuntzen arteko erlazioak errepresentatzeko transferentzi eredua hautatu da, baina erlazio elebidunak informazio osagarri hornitu dira transferentzi hiztegian. Informazio elebidunaren trataera, hizkuntza desberdinetako adieren mapaketari emandako soluzioa eta itzulpenerako informazio espezifiko gehitzeko proposamena dira nabarmentzekoak errepresentazio-ereduan. Gure ustetan, hiztegi-ezagutzaren errepresentazio-proposamena sendoa da, eta euskara eta frantsesa ez diren beste hizkuntzetarako ere erabilgarria. Proposamen horri buruzko iritzia, mugapen eta kritika desberdinak idatzi dira § III.5en. Har bitez haiek ere ondorio gisa. Basearen dimentsioei dagokienez, bi hizkuntza —euskara eta frantsesa—, hiru hiztegi —bi elebatar eta bat elebiduna—, bederetzi azpimodulu, milatik gora kontzeptu eta erlazio jasotzen dira Hiztegi-Ezagutza Basearen egungo prototipoan —datu zehatzak § III.4.6n daude ikusgai—. Ez da, bistan denez, sistema erreal baterako nahikoa. Alabaina, erabileremu zehatz bateko adierak aukeratzeak —udal administrazioarekin zerikusia duten hitzak— kohesioa eman dio hiztegi-ezagutzari eta, hortaz, erabilera-aukerak probatzeko baliagarri suertatu da.

Hiztegi-ezagutzari “portaera” ere erantsi zaio arrazonatze lexikalerako mekanismoen bitartez. Inferentzi mailako ezagutzaren parte dira bai aberasketa mekanismoak, bai dedukzio dinamikorako erregela lexikalak, bai oinarritzko funtzioak. Funtzio hauek hiztegiko unitateei esleitzen zaien portaerari buruzko ezagutza gisa ikus daitezke eta horien bitartez zertzen da inferentzi eta ataza-mailen arteko erlazioa. Hiztegiari dinamismoa ematea da, guretzat, sistemaren ezaugarri nagusienetakoa. Inferentzi maila, Hiztegi-Ezagutza Basean integratua, portaera dinamikoaren zimendarria da.

Hiztegi-kontsultako atazen deskribapenak eta eraketa hierarkikoak maila estrategikoan erabil daitezke, propio xede horretarako diseinatu den arkitekturak eutsita. Gure ustez, sistemaren arkitektura, bipila izanda ere, gai da maila desberdineko ezagutzak oso egoki maneiatzeko. Azken batean, sistemak baduke itzultzaileak hiztegi-kontsultan erabiltzen dituen prozesu kognitiboak islatzeko gaitasuna. Egun, modulu guztien diseinuak burutu direla

eta nagusienak inplementatuta daudela, lehentasun handikoa da sistemaren garapenean arkitekturako modulu guztiak inplementatzea eta integratzea.

Arkitektura hori laguntza-sistema aktiboen moldera egokitu da funtzionamenduaren aldetik. Laguntza-sistemen ezaugarri aproposenak bildu nahiz, interfaze berezitu bati ondo doakion erabilera-estiloa proposatu da. Sistemaren portaera-ezaugarriak bereziak dira oso: pasiboa eta aktiboa da, aurre hartzailea, moldagarria, urguria, adimentsua. Ohikoa ez den jarrera eskatzen du erabiltzailearen aldetik, baina, moldatuz gero, hiztegiaren erabilpenari dimentsio askoz zabalagoa emateko modukoa da.

VI.2 Aurrera begirako lan-lerroak.

VI.2.1 Hiztegien erabilera-azterketak.

Hiztegien erabilera-azterketari dagokionez beste bi saiakerari heldu nahi genieke etorkizunean. Batetik, nahi genuke azterlana datu estatistikoez hornitu: zein maiztasunekin erabiltzen diren ANHITZ ereduko oinarritzko kontsulta-atazak, zein diren arrakastatsuenak, zein diren porroten kausak, etab. Azterbide horretan, protokoloek gain, lagin handiagoak eta euskarri informatikoa ere baliatu beharko genuke.

Heldu nahi diogu, bestetik, hiztegi informatikoen erabilera-azterketari. Gure ustetan, premiazkoa —eta erakargarria— da ikerketa horri serioski heltzea, bai baita arrazoirik pentsatzeko hiztegi elektronikoen erabilerak eta paperezkoenak alde handiak dituztela elkarren artean. Hiztegi arrunten erabilpena ezezik, entziklopediena ere aztergai interesgarria dateke.

VI.2.2 Hiztegi-ezagutza osatu, egokitu eta aberasteko aukerak.

Egin dugu gogoeta orain baino lehen, § III.5en, hiztegi-ezagutzaren errepresentazioaz. Orduko hartan nabarmen geratu dira ANHITZ HEBaren aberasketan landu beharreko aspektu batzuk. Errepikatzeko arriskua gorabehera, hona ekarri ditugu etorkizun hurbilean jorratuko ditugun gaiak:

- Euskarazko hiztegiaren karga automatikoa.

Jakina da aukera urrikoa dela eskuzko langintza hiztegietatik ezagutza-baseak sortzean. Horrela aritzeak "jostailuzko" tresnak eragin ditu behin baino gehiagotan, eta ez genuke guk horretan geratu nahi. Euskarri informatikoan dagoen "Hauta-Lanerako Euskal Hiztegia"tik datu-basea sortu nahi genuke lehenik (Arriola *et al.*,

95). Datu-basea sortzeko esaldien analisi "tipografikoa" gauzatu, informazioa eremutan bereizi eta kargarako prest utziko dugu. Behin datu-basea osatuta, euskarri horretatik ekingo genioke euskarazko ezagutza-basearen eraikuntzari, lehen "Le Plus Petit Larousse" frantsesezko hiztegiarekin erabilitako metodologia (Artola, 93) edo antzekoren bat baliatuz.

- Arrazonatze-mekanismoen hobekuntza. Hitzetik adierarako desanbiguazio automatikoa.

HEBaren arrazonatze-gaitasuna aberasteko aukera bat baino gehiago nabarmendu daiteke, adibidez, erregelak ugaltzea, arrazonamendu ez-zehatzerako teknikak lantzea eta adiera-desanbiguazioa automatizatzea.

Ez dago zertan esan zein onuragarria litzatekeen adiera-desanbiguazioa burutzea, bai hiztegietako erlazio lexikal-semantikoak osatzeko, bai, horren ondorioz, hainbat inferentziari burutzapen egokia emateko. Eragin biderkatzailea duke adiera-desanbiguazioak, domeinuaren mailatik inferentzietara eta, are, itzultzailearekiko elkarrekintzan ere nabarmenduko litzatekeena.

Bide desberdinak daude adiera-desanbiguazioari aurre egiteko. Gurean eleaniztasunak ematen dituen aukerak aprobetxatu nahiko genituzke, behar bada beste baliabideren bat —WordNet, kasu baterako— sostengu gisa hartuz (Rigau & Agirre, 95). Adieren arteko *distantzia kontzeptualaren* erabilera (Agirre *et al.*, 94d; Agirre & Rigau, 95) hurbilpen interesgarria izan daiteke zeregin zail horretan.

- Ezagutza-baseetako klaseen tipifikazio finagoa.

Klaseak fintzeak sare kontzeptualetako unitateen ezaugarritze aberatsagoa ekarriko luke eta, noski, unitate-atributu finagoak mesederako lirateke sarearen erabiltzeko. Hori horrela izanda ere, ezin da ahaztu fintzera joz gero ezagutza-baseen sorkuntzak zehatzago izan beharko lukeela, besteak beste, kontzeptuen propietateak atributu edo slot egokiei esleitzeko. Beraz orekaren bila joan beharko litzateke: ustiapena aberastu eta sorkuntza gehiegi zaildu ez.

- Ezagutza-baseak ezagutza gramatikalez aberastea.

Hiztegiak informazio-iturri nagusitzat hartuta ere, itzultzaileari eskaini nahiko geniokeen guztia nekez atera daiteke soilik hiztegietatik. Gramatikari buruzko ezagutza, bereziki azpikategorizazio, hautapen-murritzapen eta kokakidetzei buruzkoa, korpusetatik erauztea litzateke gure desioa.

- Hiztegianiztasuna eta eleaniztasuna.

Egungo ANHITZ prototipoak bi hizkuntza darabiltza eta hiztegi bakarra dauka hizkuntzako —beste hiztegi bat elebiduna da—. Hizkuntzen aniztasunari baino hiztegien ugalketari eman nahi diogu lehentasuna hemendik aurrerakoan. Eta, gainera, mota desberdineko hiztegiak bildu nahi genituzke batera, beraien arteko lankidetzaz aztertu eta ahal den neurrian probetxua ateratzeko. Jakina, bide horretatik abiatzeak errepresentazio-formalismoaren egokitzapena eskatzen du. Laxatu egin beharko dugu gure formalismoa, hiztegi-informazioa errepresentatzeko aukera berriak bilatu, behar bada ez hain aberatsak beti, baina probetxugarriak edonola ere. Hiztegi gehiago erabiltzeak ez luke ANHITZ sistemaren arkitekturan eta funtzionamenduan funtsezko aldaketarik eragingo.

VI.2.3 Laguntza emateko beste erak.

ANHITZek, laguntza ematen duen sistema izaki, baleukake hobetzekorik itzultzaileari laguntza-informazioa aurkeztean.

Esan dugu V. kapituluan ANHITZek ez duela hartutako erabakiei buruzko azalpenik ematen, soilik hiztegi-informazioa ematera mugatzen dela. Zenbaitetan desiragarria litzateke, ordea, hiztegi-erabilpenari buruz dakienaz azalpenak ematea. Besteak beste, hiztegia nola erabiltzen duen justifikatzea interesgarria izan daiteke.

Itzultzailearekiko interfazearen eraketa ere ez genuke albora utzi nahi. Lehen ekimena, noski, V. kapituluko diseinua inplementatzea izango da. Baina, ezin daiteke ahanztu, bestalde, interfazearen diseinuan aipatu arren, ez diegula kabidari eman ez hipertestuari, ez lengoia naturallezko formulazioei. Bi aukera hauek eta, oro har, interfazearen diseinua aberastea etorkizunean egin beharrekoen artean daude.

VI.2.4 Baliozkotasuna eta ebaluazioa.

Sistema giza itzultzaileari zuzendua dagoenez premiazkoa da itzultzailearen parte hartzea sustatzea, nolabait, interfazea eta itzultzailearekiko elkarriketa bideratzea. Hori gabe, uste huts izango dira sistemaren portaerari iritziko dizkiogunak, eta hori baino sendoago aztertu nahi genuke sistema. Beraz, ANHITZen baliozkotasuna itzultzailearekiko elkarlanean egiaztatuko da.

T. Hoppe-ren "Validation of User Intention" (1990) artikuluan ezagutzaren baliozkotasuna erabiltzaileak espero duenarekin lotzen da: "*...knowledge validation means to determine whether the user intention meets the actual system*". Horren arabera, ezagutzaren

eskurapenean erabiltzailearekin izandako harremanaren antzekoa beharko litzateke ezagutzaren baliozkotasuna neurtzean ere.

(Krause *et al.*, 93)n COMFOHELP laguntza-sistemaren prototipo baten ebaluazioa deskribatzen da. COMFOTEX testu-prozesadorearen laguntza-sistema da COMFOHELP. Balidatze-probak test esperimentaletan oinarritzen dira¹²⁹, eta horiek medio, laguntza-sistemaren baliozkotasun enpirikoa neurtzen da. Badira, gainera, test berezituak sistemaren ataza-ezagutzailea eta erabiltzaile-ereduak egiaztatzeko. Test horien ezaugarriak egokiak dira ANHITZ bezalako sistemei aplikatzeko.

ANHITZen Hiztegi-Ezagutza Basearen egiaztapenak ere badu interesa guretzat. Sare kontzeptualaren topologia eta ezagutzaren zehaztasun, sendotasun eta osotasuna aztertu nahiko genituzke. Jakitun gara azterketa horien zailtasunez, baina, hasteko, probarako lagin bat eskuz egiaztatzeari ekingo diogu, etorkizunerako hobekuntzen aztarna har dezagun.

VI.2.5 Sistemaren berrerabilgarritasuna.

ANHITZ sistema, osotara hartuta, testuinguru desberdinetan izan daiteke erabilia, duen arkitekturak eta eskaintzen duen funtzionalitateak badute-eta malgutasun nahikoa horretarako.

Konparazio batera, erakargarria izan daiteke ANHITZ irakaskuntza-sistema adimendunen (Intelligent Tutoring Systems, ITS) esparrura eramatea, hiztegien erabilpenaz eta hitzen mailako itzulpenaz ikasten laguntzeko. Egokitzapen horretan ataza konposatuak egikaritzeko eta azalpenak emateko gaitu beharko litzateke, besteak beste. Ez dugu uste, hala ere, gaur egungo diseinuak aldaketa handirik jasan beharko lukeenik, ITS gehienek bai baitute ANHITZen moduluen parekorik. Aldaketak baino gehiago, eransketak eta aberasketak beharko lirateke ANHITZ irakaskuntza-sistema adimendunen gisara eratzeko. Ezin da ahanzi, gainera, gaur egun ikasteko inguruneek duten garrantzia eta, testuinguru horretan, interesgarria izan daiteke ANHITZ beste tresna batzuekin integratzea, adibidez, analizatzaile morfologiko eta sintaktikoekin. Kontuan izatekoa da, bestalde, KADS metodologian egiten den domeinu- eta ataza-errepresentazioen bereizketa onetsia izan dela irakaskuntza-sistemen garapenean (Kemp &Smith, 94).

Irakaskuntza-sistemetara pasa gabe ere, giza itzulpenerako laguntza-sistema izaten jarraituz, ez da baztertzekoa ANHITZ beste testuinguru zabalagoetan erabiltzeko aukera. Unitate lexikalen itzulpenean izandako esperientzian oin hartuz, planteatzekoa da ANHITZen ereduako sistema baten eraikuntza, oro har, testu-itzulpenerako. Gure ustetan, sistemaren

¹²⁹ Ehun bat erabiltzailearekin egin zituzten probak.

ezagutza-baseak egokitu beharko liratekeen arren, eraikuntza-metodologia, diseinu orokorra eta funtzionamendu-ezaugarriak berrerabilgarriak dira izaera horretako sistema zabalagoetan.

Hirugarren aplikazio-esparrua mundu komertzialean dagoke. ANHITZek, helburu eta irizpide praktikoez garatutako sistema izaki, maila komertzialera jo behar du. Mundu komertzial horretara egokitzeko bideragarria da, hiztegi errealetan dauden erlazio esplizitu nahiz implizituak —taxonomikoak edo sinonimikoak, adibidez— erabiliz, oinarritzko funtzionalitatea eutsiko duen sistema garatzea. Sistemaren garapen-zikloak bertsio komertzial baten inplementazioa du azken urratsa, eta bide horretan, hastapeneko urratsa da txosten honetan azaldutakoa. Beto, bada, hurrengoak.

BIBLIOGRAFIA

- Ageno A., Ribas F., Rigau G., Rodríguez H.. TGE: Un entorno para la generación de enlaces léxicos multilingües, *Boletín SEPLN*, no. 14, 291-312, 1994.
- 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, *Proc. 3rd Conf. on Applied Natural Language Processing (ANLP'92, Trento)*, 119-125, 1992.
- Agirre E., Arregi X., Arriola J.M., Artola X., Insausti J.M.. *Euskararen Datu-Base Lexikala (EDBL)*. Lengoia eta Sistema Informatikoen saila, Euskal Herriko Unib. (EHU). Txosten teknikoa UPV/EHU/LSI/TR 8-94. 1994a.
- Agirre E., Arregi X., Artola X., Díaz de Ilarraza A., Evrard F., Sarasola K.. Analysing word-level translation activity to design a computerised dictionary help system, (Panel Session) *EURALEX'94 Congress*, Amsterdam. 1994b.
- Agirre E., Arregi X., Artola X., Díaz de Ilarraza A., Evrard F., Sarasola K.. IDHS, MLDS: Towards Dictionary Help Systems for Human Users, *Semantics and Pragmatics of Natural Language: Logical and Computational Aspects*, Sara, 167-188, 1993.
- Agirre E., Arregi X., Artola X., Díaz de Ilarraza A., Evrard F., Sarasola K.. Sistema diccionario multilingüe: aproximación funcional, *Boletín SEPLN*, 14, 313-333, 1993.
- Agirre E., Arregi X., Artola X., Díaz de Ilarraza A., Sarasola K.. A methodology for the extraction of semantic knowledge from dictionaries using phrasal patterns, *Proc. IBERAMIA'94* (Caracas), 263-270, 1994c.
- Agirre E., Arregi X., Artola X., Díaz de Ilarraza A., Sarasola K.. Conceptual Distance and Automatic Spelling Correction, *Workshop on Computational Linguistics for Speech and Handwriting Recognition*, Leeds, 1994d.
- Agirre E., Arregi X., Artola X., Díaz de Ilarraza A., Sarasola K.. Lexical Knowledge Representation in an Intelligent Dictionary Help System, *Proc. COLING'94* (Kyoto), 544-550, 1994e.
- Agirre E., Rigau G.. A Proposal for Word Sense Disambiguation using Conceptual Distance, *Int. Conference on Recent Advances in Natural Language Processing-era bidalia*, Bulgaria, 1995.

- Ahlswede T., Evens M., Rossi K., Markowitz J.. Building a Lexical Database by Parsing Webster's Seventh Collegiate Dictionary, *Advances in Lexicology. Proceedings 2nd Annual Conference of the UW Centre for the New Oxford English Dictionary* (Waterloo, Canada), 65-78, 1986.
- Akkermans H., van Harmelen F., Schreiber G., Wielinga B.. A Formalization of Knowledge-Level Models for Knowledge Acquisition, *Int. J. of Intelligent Systems*, vol. 8, 169-208, 1993a.
- Akkermans H., Wielinga B., Schreiber G.. Steps in Constructing Problem Solving Methods, in N. Aussenac et al. eds., 45-65, *Knowledge Acquisition for Knowledge-Based Systems*. Toulouse: Springer-Verlag, 1993b.
- Al B.. Dictionnaire Bilingue et Transfert Automatique de Donnees Lexicales, in L. Fignoni, C. Peters eds., vol. I, 17-32, *Computational Lexicology and Lexicography (Special Issue dedicated to B. Quemada)*. Pisa: Giardani, 1990.
- Al-Kasimi A.. *Linguistics and Bilingual Dictionaries*. Leiden: E.I.Brill, 1977.
- Amsler R.A.. *The Structure of The Merriam-Webster Pocket Dictionary*, PhD. Diss. Computer Science, TR-164. University of Texas, Austin, 1980.
- Ard J.. The use of bilingual dictionaries by ESL student while writing, *ITL Review of Applied Linguistics*, 58, 1-27.1982.
- Arregi X., Artola X., Díaz de Ilarraza A., Evrard F., Sarasola K.. Aproximación funcional a DIAC: Diccionario inteligente de ayuda a la comprensión, *Boletín SEPLN*, 11, 127-138, 1991.
- Arriola J.M., Artola X., Soroa A.. Análisis automático del diccionario Hauta-Lanerako Euskal Hiztegia, *SEPLN-95erako onartua*, Bilbo, Iraila, 1995.
- Artola X.. *HIZTSUA: Hiztegi-sistema urgazle adimendunaren sorkuntza eta eraikuntza / Conception et construction d'un système intelligent d'aide dictionnaire (SIAD)*, Ph.D. Tesia, EHU, 1993.
- Atkins B.T., Knowles F.E.. Interim report on the EURALEX/AILA Research Project into Dictionary use, in T. Magay, J. Zigány eds., 381-392, *Proceedings BudaLEX'88*. Budapest: akadémiai kiadó, 1990.
- Atkins B.T.S., Zampolli A. (eds.). *Computational Approaches to the Lexicon*. New York: Oxford University Press, 1994.
- Bachut D., Verastegui N.. Software Tools for the Environment of a Computer Aided Translation System, *Proc. of COLING'84* (Stanford), 330-333, 1984.
- Barnhart C.L.. Problems in editing commercial monolingual dictionaries, in F. Householder, S. Saporta eds., 161-181, *Problems in Lexicography*. Bloomington: Indiana University, 1962.

- Bartholomew D.A., Schoenhals L.C.. *Bilingual Dictionaries for Indigenous Languages*, Mexico: Summer Institute of Linguistics. 1983.
- Bates M.J.. Where should the Person Stop and the Information Search Interface Start?, *Information Processing & Management*, vol. 26, no. 5, 575-591. 1990.
- Baxter J.. The Dictionary and Vocabulary Behavior: a Single Word or a Handful?, *TESOL Quarterly*, vol. 14, no. 3, 325-336. 1980.
- Beaven J. L.. Shake-and-Bake Machine Translation, *Proc. COLING'92* (Nantes), 602-609, 1992.
- Beckwith R., Fellbaum C., Gross D., Miller G.A.. WordNet: A Lexical Database Organized on Psycholinguistic Principles, in U. Zernik ed., 211-232, *Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon*. Hillsdale/New Jersey/Hove/London: Lawrence Erlbaum Associates Publishers, 1991.
- Béjoint H.. The Foreign Student's Use of Monolingual English Dictionaries: A Study of Language Needs and Reference Skills, *Applied Linguistics*, vol. 2, no. 3, 207-222. 1981.
- Béjoint H.B. and Moulin. The place of the dictionary in an EFL programme, in A. Cowie ed., *The dictionary and the language learner*. Tübingen. 1987.
- Bensoussan M. Dictionaries and tests of EFL reading comprehension, *ELT Journal*, vol. 37, no. 4, 341-345, 1983.
- Blanchon H.. Perspectives of DBMT for monolingual authors on the basis of LIDIA-1, an implemented mock-up, *Proc. of COLING'94* (Kyoto), 115-119, 1994.
- Bläser B., Schwall U., Storrer A.. A Reusable Lexical Database Tool for Machine Translation, *Proc. COLING'92* (Nantes), 510-516, 1992.
- Boguraev B., Levin B.. Models for Lexical Knowledge Bases, in J. Pustejovsky ed., 325-340, *Semantics and the Lexicon*. Dordrecht/Boston/London: Kluwer Academic Publishers, 1993.
- Boitet C., Blanchon H.. Multilingual Dialogue-Based MT for Monolingual Authors: the LIDIA Project and a First Mockup, *Machine Translation*, 9, 99-132, 1994.
- Boitet C.. Towards Personal MT: general design, dialogue structure, potential role of speech, *Proc. of COLING'90* (Helsinki), 30-35, 1990.
- Brajnik G., Tasso C.. A shell for developing non-monotonic user modeling systems, *Int. J. of Human-Computer Studies*, vol. 40, 31-62, 1994.
- Breuker J. (ed.). *EUROHELP. Developing Intelligent Help Systems*. Report on the P280 ESPRIT Project EUROHELP. Copenhagen, 1990.
- Bujas Z. Testing the performance of a bilingual dictionary on topical current texts, *Studia Romanica et Anglica Zagrabienia*, no. 39, 194-204. 1975.

- Byrd R.J., Calzolari N., Chodorow M.S., Klavans J.L., Neff M.S., Rizk O.A.. Tools and Methods for Computational Lexicology, *Computational Linguistics*, vol. 13, ns. 3-4, 219-240. 1987.
- Calzolari N., Picchi E.. A Project for a Bilingual Lexical Database System, *Proc.2nd Annual Conf. of the UW Centre for the New OED* (Waterloo), 79-92, 1986.
- Calzolari N.. Structure and access in an automated lexicon and related issues, in L. Fignoni, C. Peters eds., vol. I, 139-161, *Computational Lexicology and Lexicography (Special Issue dedicated to B. Quemada)*. Pisa: Giardani, 1990.
- Carberry S.. Tracking User Goals in an Information-Seeking Environment, *Proceedings of the National Conference on Artificial Intelligence* (Washington), 59-63, 1983.
- Card S.K., Moran T.P., Newell A.. *The psychology of human-computer interaction*. Hillsdale, NJ: Erlbaum, 1983.
- Carpenter B.. Typed feature structures: Inheritance (In)equality and Extensionality, *Proceedings of the First Int. Workshop on Inheritance in NLP* (Tilburg, The Netherlands), 9-18. 1990.
- Copestake A., Briscoe T., Vossen P., Ageno A., Castellon I., Ribas F., Rigau G., Rodríguez H., Samiotou A.. *Acquisition of Lexical Translation Relations from MRDs*, Report LSI-94-35-R (Univ. Politècnica de Catalunya), 1994.
- Copestake A., Sanfilippo A.. Multilingual lexical representation, *ACQUILEX II - Esprit Project 7315*, 1993.
- Copestake A.. The ACQUILEX LKB: representation issues in semi-automatic acquisition of large lexicons, *Proceedings 3rd. Conference on Applied Natural Language Processing* (Trento, Italia), 88-95. 1992.
- Cruse D.A.. *Lexical Semantics*. Cambridge: Cambridge University Press, 1986.
- Chandrasekaran B., Johnson T.R.. Generic Tasks and Task Structures: History, Critique and New Directions, in J.M. David, J.P. Krivine and R. Simmons eds., 232-272, *Second Generation Expert Systems*. Springer-Verlag, 1993.
- Chandrasekaran B., Tanner M.C., Josephson J.R.. Explaining Control Strategies in Problem Solving, *IEEE Expert*, vol. 4, no. 1, 9-24, 1989.
- Chandrasekaran B.. Design Problem Solving: A Task Analysis, *AI Magazine*, vol. 11, no. 4, 59-71, 1990.
- Chandrasekaran B.. Towards a Functional Architecture for Intelligence Based on Generic Information Processing Tasks, *Proceedings IJCAI'87* (Milan), 1183-1192, 1987.
- Dobrovolskij D.. Idioms in a Semantic Network: Towards a New Dictionary-Type, *Proceedings EURALEX'94* (Amsterdam), 263-270. 1994.

- Dorr B.. Conceptual Basis of the Lexicon in Machine Translation, in U. Zernik ed., 263-308, *Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon*. Hillsdale/New Jersey/Hove/London: Lawrence Erlbaum Associates Publishers, 1991.
- Dorr B.J.. Machine Translation Divergences: A Formal Description and Proposed Solution, *Computational Linguistics*, vol. 20, no. 4, 597-633, 1994.
- Dorr B.J.. *Machine Translation: A View from the Lexicon*. The MIT Press, 1993.
- EDR *Electronic Dictionary Technical Guide*. Japan Electronic Dictionary Research Institute, Ltd. TR-042, 1993.
- Evens M.W. (ed.). *Relational models of the lexicon. Representing knowledge in semantic networks*. Cambridge: Cambridge Univ. Press, 1988.
- Fass D.. Lexical Semantic Constraints, in J. Pustejovsky ed., 263-289, *Semantics and the Lexicon*. Dordrecht/Boston/London: Kluwer Academic Publishers, 1993.
- Fensel D. and van Harmelen F.. A Comparison of Languages which Operationalize and Formalize KADS Models of Expertise, *The Knowledge Engineering Review*, vol. 9, no. 2, 105-146, 1994.
- Fillmore C.J., Atkins T.S.. Starting where the Dictionaries Stop: The Challenge of Corpus Lexicography, in B.T.S. Atkins, A. Zampolli eds., 349-393, *Computational Approaches to the Lexicon*. New York: Oxford University Press, 1994.
- Fischer G., Lemke A., Schwab T.. Active Help Systems, in T. Green, M. Tauber, G. van der Veer eds., 116-131, *Proceedings of 2nd Conf. on Cognitive Ergonomics*. Heidelberg: Springer Verlag, 1984.
- Fischer G., Lemke A., Schwab T.. Knowledge-based Help Systems, *Proceedings CHI'85* (San Francisco), 161-167, 1985.
- Fontenelle T., Adriaens G., de Braekeleer G.. The Lexical Unit in the Metal[®] MT System, *Machine Translation*, no. 9, 1-19, 1994.
- Fox M.S., Bebel D.J., Parker A.C.. The Automated Dictionary, *IEEE Computer*, 35-48, July, 1980.
- Gale W.A., Church K.W.. A Program for Aligning Sentences in Bilingual Corpora, *Computational Linguistics*, vol. 19, no. 1, 75-102, 1993.
- Gates D. et al.. Lexicons, *Machine Translation (Special Issue on Knowledge-Based Machine Translation)*, no. 4, 67-112, 1989.
- Goodman B.A., Litman D.J.. On the Interaction between Plan Recognition and Intelligent Interfaces, *User Modeling and User-Adapted Interaction*, 2, 83-115, 1992.

- van Harmelen F. and Balder J.. (ML)²: A Formal Language for KADS Models of Expertise, in G.Schreiber, B. Wielinga and J. Breuker eds., 169-202, *KADS. A Principled Approach to Knowledge-Based System Development*, Knowledge-Based Systems, vol. 11. Amsterdam: Academic Press, 1993.
- Hartmann R.R.K.. Four perspectives on dictionary use: a critical review of research methods. The Dictionary and the Language Learner, in A.Cowie ed., 11-28, *Papers from the EURALEX Seminar (Leeds)*, *Lexicographica, Series maior*, no. 17, Tübingen; Niemeyer. 1985.
- Hartmann R.R.K.. The bilingual learner's dictionary and its uses, *Multilingua*, vol. 2, no. 4, 195-201, 1983.
- Hartmann R.R.K.. The Use of Parallel Text Corpora in the Generation of Translation Equivalents for Bilingual Lexicography, *Proceedings EURALEX'94* (Amsterdam), 291-297. 1994.
- Hatherall G.. Studying Dictionary Use: Some Findings and Proposals, *Proceedings LEXeter'83*, 183-189, *Lexicographica, Series maior*, no.1, Tübingen; Niemeyer, 1984.
- Hauenschild C.. The Lexicon and the General Design of MT Systems, *Machine Translation and the Lexicon, Workshop of the European Association for Machine Translation*, Heidelberg, 1993.
- Hayes-Roth B.. A Blackboard Architecture for Control, *Artificial Intelligence*, 26, 251-321, 1985.
- Hayes-Roth B.. An architecture for adaptive intelligent systems, *Artificial Intelligence*, 72, 329-365, 1995.
- Heid U., Raab S.. Collocations in Multilingual Generation, *Proc. 4th Conf. of the EACL* (Manchester), 130-136, 1989.
- Heid U.. Bilingual Lexicography and Transfer Dictionaries for Machine Translation — Treating Structural Differences Between French and German, in T. Magay, J. Zigány eds., *Proc. BUDALEX'88*, 107-119. Budapest: Akadémiai Kiadó, 1990.
- Heid U.. On Ways Words Work Together — Topics in Lexical Combinatorics, *Proceedings EURALEX'94* (Amsterdam), 226-257, 1994.
- Herman D.J.. *DSPL++: A High-Level Language for Building Design Expert Systems with Flexible Use of Multiple Methods*, Ph.D. diss., The Ohio State Univ., 1990.
- Heylen D., Maxwell K. and Armstrong-Warwick S.. Collocations, Dictionaries and MT, *AAAI'93 Spring Symposium, Series Building Lexicons for Machine Translation* (Stanford), 1993.
- Heylen D., Maxwell K.G., Verhagen M.. Lexical Functions and Machine Translation, *Proc. COLING'94* (Kyoto), 1240-1244, 1994.

- Hoppe T.. Validation of User Intention, in Wielinga et al. eds., 161-172, *Current Trends in Knowledge Acquisition*. Amsterdam: IOS Press, 1990.
- Houghton R.C.. Online Help Systems: A Conspectus, *Communications of the ACM*, vol. 27, no. 2, 126-133, 1984.
- Hurley W.D.. A Method for Predicting the Impact of User Interface Design on Software Design, *The Journal of Systems and Software*, vol. 22, no. 3, 179-190, 1993.
- Hutchins W.J., Somers H.L.. *An Introduction to Machine Translation*. Cambridge: Academic Press, 1992.
- Ibarzabal A. eta Navarro K.. Euskal lokuzioak/Hiztegi erotikoa, *Senez: Itzulpen eta Terminologiazko Aldizkaria*, vol. 2, no. 1, 91-101, 1985.
- Isabelle P.. Bi-Textual Aids for Translators, *Proc.8th Annual Conf. of the UW Centre for the New OED and Text Research* (Waterloo), 76-89, 1992.
- Johnson P. and Johnson H.. Knowledge analysis of tasks: task analysis and specification for human-computer systems, in A.Downton ed., 119-144, *Engineering the human-computer interface*. London: McGraw-Hill, 1991.
- Johnson R.L., Whitelock P.. Machine translation as an expert task, in S. Nirenburg ed., 136-144, *Machine Translation: Theoretical and methodological issues*. Cambridge: Cambridge Univ. Press, 1987.
- Jonker W., Spee J.W.. Yet another formalisation of KADS Conceptual Models, in T. Wetter et al. eds., 211-229, *Current Developments in Knowledge Acquisition*, Lecture Notes in Artificial Intelligence, no. 599, Springer-Verlag, 1992.
- Karbach W., Linster M., Voß A.. Models of Problem-Solving: One label - One idea?, in Wielinga et al. eds., 173-189, *Current Trends in Knowledge Acquisition*. Amsterdam: IOS Press, 1990.
- Karbach W., Voß A. and Drouven U.. Solving the office allocation task in reflective ASSIGN, *Int. J. Human-Computer Studies*, no. 40, 273-291, 1994.
- Karbach W., Voß A.. MODEL-K for prototyping and strategic reasoning at the knowledge level, in J.M. David, J.P. Krivine and R. Simmons eds., 721-745, *Second Generation Expert Systems*. Springer Verlag, 1993.
- Kawasaki Z., Yamano F., Yamasaki N.. Translator Knowledge Base for Machine Translation Systems, *Machine Translation*, no. 6, 265-278, 1992.
- Kay M.. *The Proper Place of Men and Machines in Translation*, CSL-80-11, Xerox PARC, 1980.
- Kemke C.. *The SINIX Consultant. Requirements, Design, and Implementation of an Intelligent Help System for a UNIX Derivative*, Bericht Nr. 11 (Universität Saarbrücken), 1986.

- Kemp R.H., Smith S.P.. Domain and task representation for tutorial process models, *Int. J. Human-Computer Studies*, 41, 363-383, 1994.
- Klavans J.L.. Building a Computational Lexicon Using Machine Readable Dictionaries, in T. Magay, J. Zigány eds., *Proc. BUDALEX'88*, 265-279. Budapest: Akadémiai Kiadó, 1990.
- Knight K., Luk S.K.. Building a Large-Scale Knowledge Base for Machine Translation, *Proc. AAAI'94*.
- Koizumi *et al.*. Noun Phrasal Entries in the EDR English Word Dictionary, *Proc. COLING'94* (Kyoto), 257-262, 1994.
- Krause J., Mittermaier E., Hirschmann A.. The Intelligent Help System COMFOHELP, *User Modeling and User-Adapted Interaction*, 3, 249-282, 1993.
- Krause J., Mittermaier E., Hirschmann A.. The Intelligent Help System COMFOHELP. Towards a Solution of the Practicality Problem for User Modeling and Adaptive Systems, *User Modeling and User-Adapted Interaction*, no. 3, 249-282, 1993.
- Kugler M., Ahmad K., Thurmair G. (eds.). *Translator's Workbench. Tools and Terminology for Translation and Text Processing*, Research Reports, ESPRIT Project 2315, TWB, vol. 1. Luxembourg: Springer-Verlag, 1995.
- Laird J.E., Newell A., Rosenbloom P.S.. SOAR: An Architecture for General Intelligence, *Artificial Intelligence*, 33, 1-64, 1987.
- Large A. and Beheshti J.. A Comparison of Information Retrieval from Print and CD-ROM versions of an Encyclopedia by Elementary School Students, *Information Processing & Management*, vol. 30, no. 4, 499-513, 1994.
- Lee J. and Yen J.. Enhancing the Software Life Cycle of Knowledge-Based Systems using a Task-Based Specification Methodology, *Int. J. of Software Engineering and Knowledge Engineering*, vol. 3, no. 1, 3-15, 1993.
- Levin L., Nirenburg S.. The Correct Place of Lexical Semantics in Interlingual MT, *Proc. COLING'94* (Kyoto), 349-355, 1994.
- Linster M.. Using OMOS to Represent KADS Conceptual Models, in G. Schreiber, B. Wielinga and J. Breuker eds., 221-246, *KADS. A Principled Approach to Knowledge-Based System Development*, Knowledge-Based Systems, vol. 11. Amsterdam: Academic Press, 1993.
- Macklovitch E.. An off-the-shelf workstation for translators, *Proc. of the 30th American Translators Conference* (Washington), 491-498, 1989.
- Marchionini G.. Making the transition from print to electronic encyclopaedias: adaptation of mental models, *Int. J. of Man-Machine Studies*, no. 30, 591-618. 1989.
- Martí M.A., Soler C.. Dealing with Lexical Mismatches, *Proceedings EURALEX'94* (Amsterdam), 360-367. 1994.

- Martin W. and Al B.P.F.. User-orientation in dictionaries: 9 propositions, in T. Magay, J. Zigány eds., 393-399, *Proceedings BudaLEX'88*. Budapest: akadémiai kiadó, 1990.
- Martin W.. On the organization of semantic data in passive bilingual dictionaries, *Actas del IV congreso Internacional.EURALEX'90* (Benalmádena), 193-201, Biblograph, 1992.
- Melby A.. On human-machine interaction in translation, in S. Nirenburg ed., 145-154, *Machine Translation: Theoretical and methodological issues*. Cambridge: Cambridge Univ. Press, 1987a.
- Melby A.K.. Creating an Environment for the Translator, in M. King ed., 124-132, *Machine Translation Today*. Edinburgh: Edinburgh Univ. Press, 1987b.
- Mel'cuk I.A., Zholkovsky A.. The explanatory combinatorial dictionary, in M.W. Evens ed., c. 2, 41-74, *Relational Models of the Lexicon*. Cambridge: Cambridge Univ. Press, 1988.
- Mel'cuk I.A.. Lexical functions in lexicographic descriptions, *Proc. VIIIth Annual Meeting of the Berkeley Linguistic Society*, 427-444, 1982.
- Minaeva L.. Dictionary examples: friends or foes?, *Proceedings EURALEX'92* (Tampere), 77-80, 1992.
- Mitchell E.. *Search-Do Reading: Difficulties in Using a Dictionary*, Formative Assessment of Reading Strategies in Secondary Schools - Working Paper 21 (Aberdeen College of Education), 1983.
- Moss, G.. Cognate Recognition: Its Importance in the Teaching of ESP Reading Courses to Spanish Speakers, *English for Specific Purposes*, no. 11, 141-158, 1992.
- Nesi H.. The Effect of Language Background and Culture on Productive Dictionary Use, *Proceedings EURALEX'94* (Amsterdam), 577-585, 1994.
- Neubach A. and Cohen A.D.. Processing Strategies and Problems Encountered in the use of Dictionaries, *Dictionaries. Journal of the Dictionary Society of North America*, 10, 1-20, 1988.
- Neubert A.. Fact and Fiction of the Bilingual Dictionary, *Actas del IV congreso Internacional.EURALEX'90* (Benalmádena), 29-42, Biblograph, 1992.
- Nirenburg S., Defrise C.. Lexical and Conceptual Structure for Knowledge-Based Machine Translation, in J. Pustejovsky ed., 291-323, *Semantics and the Lexicon*. Dordrecht/Boston/London: Kluwer Academic Publishers, 1993.
- Nirenburg S., Shell P., Cohen A., Cousseau P., Grannes D., McNeilly C.. Multi-Purpose Development and Operation Environments for Natural Language Applications, *Proceedings 3rd Conference on Applied Natural Language Processing* (Trento, Italia), 255-256, 1992.
- Nuccorini S.. On Dictionary Misuse, *Proceedings EURALEX'94* (Amsterdam), 586-597. 1994.

- Pazienza M.T., Velardi P.. A Structured Representation of Word-senses for Semantic Analysis, *Proc. 3rd Conf. of the EACL* (Copenhagen), 249-257, 1987.
- Picchi E., Peters C., Calzolari N.. Implementing a Bilingual Lexical Database System, in T. Magay, J. Zigány eds., *Proc. BUDALEX'88*, 317-329. Budapest: Akadémiai Kiadó, 1990.
- Picchi, E., Peters C., Marinai E.. A Translator's Workstation, *Proc. of COLING'92* (Nantes), 972-976, 1992.
- Punch W.. *A Diagnosis System Using a Task-Integrated Problem Solver Architecture (TIPS), Including Causal Reasoning*, Ph.D. diss., The Ohio State Univ., 1989.
- Pustejovsky J. (ed.). *Semantics and the Lexicon*. Dordrecht/Boston/London: Kluwer Academic Publishers, 1993.
- Pustejovsky J., Boguraev B.. A Richer Characterization of Dictionary Entries: The Role of Knowledge Representation, in B.T.S. Atkins, A. Zampolli eds., 295-311, *Computational Approaches to the Lexicon*. New York: Oxford University Press, 1994.
- Pustejovsky J., Boguraev B.. Lexical knowledge representation and natural language processing, *Artificial Intelligence*, 63, 193-223. 1993.
- Pustejovsky J.. The Generative Lexicon, *Computational Linguistics* 17, 4, 409-441. 1991.
- Quillian M.R.. Semantic Memory, in M. Minsky ed., 227-270, *Semantic Information Processing*. Cambridge (Mass.): MIT Press, 1968.
- Quirk R.. The social impact of dictionaries in the U.K., in R.I. McDavid, A.R. Duckert eds., 76-88, *Lexicography in English*. New York: New York Academy of Sciences, 1973.
- Radermacher F.J.. Decision support systems: Scope and potential, *Decision Support Systems*, 12, 257-265, 1994.
- Rappaport M., Laughren M., Levin B.. Levels of Lexical Representation, in J. Pustejovsky ed., 37-54, *Semantics and the Lexicon*. Dordrecht/Boston/London: Kluwer Academic Publishers, 1993.
- Raymond D.R., Tompa F.W.. Hypertext and the Oxford English Dictionary, *Communications of the ACM*, vol. 31, no. 7, 871-879, 1988.
- Reinders et al.. A Conceptual Modelling Framework for Knowledge-level Reflection, *AI Communications*, vol. 4, ns. 2-3, 74-87, 1991.
- Rigau G., Agirre E.. Disambiguating bilingual nominal entries against WordNet, *Workshop on the Computational Lexicon-era bidalia*, Bartzelona, 1995.

- Rouge A., Brossier F., Loziquez Y.. The VITAL enterprise to open KADS to knowledge & software engineering methodologies and tools, *Proceedings AVIGNON'93* (Avignon), 437-446, 1993.
- Rundell M. and Ham N.. A New Conceptual Map of English:, *Proceedings EURALEX'94* (Amsterdam), 172-180. 1994.
- Samiotou A., Castellón I., Ribas F., Rigau G.. *Translation Equivalence via Lexicon: A Study on Tlinks*, Report LSI-93-48-R (Univ. Politècnica de Catalunya), 1993.
- Santos D.. Lexical gaps and idioms in machine translation, *Proc. COLING'90* (Helsinki), 330-335, 1990.
- Schefe P.. Natürlichsprachlicher Zugang zu Datenbanken?, *Angewandte Informatik*, no. 10, 419-423, 1983.
- Scholfield P. Using the English Dictionary for Comparison, *TESOL Quarterly*, 16, 185-194, 1982.
- Schreiber G., Wielinga B., Akkermans H., Van de Velde W.. CML: The CommonKADS Conceptual Modelling Language, *European Knowledge Acquisition Workshop (EKAW'94)*.
- Schreiber G., Wielinga B., Breuker J. (eds.). *KADS: A Principled Approach to Knowledge-Based System Development*, Knowledge-Based Systems, vol. 11. London: Academic Press, 1993.
- Sérasset G.. Interlingual Lexical Organisation for Multilingual Lexical Databases in NADIA, *Proc. COLING'94* (Kyoto), 278-282, 1994.
- Sinclair J.M.(ed.). *Looking up. An Account of the COBUILD Project in Lexical Computing*. London: Collins, 1987.
- Smadja F. and McKeown K.. Using Collocations for Language Generation, *Computational Intelligence*, vol. 7, no. 4, 229-239, 1991.
- Smadja F.. Macrocoding the Lexicon with Co-occurrence Knowledge, in U. Zernik ed., 165-189, *Lexical Acquisition: Exploiting On-Line Resources to Build a Lexicon*. Hillsdale/New Jersey/Hove/London: Lawrence Erlbaum Associates Publishers, 1991.
- Solomon P.. Children's information Retrieval Behavior: A Case Analysis of an OPAC, *Journal of the American Society for Information Science*, vol. 44, no. 5, 245-264, 1993.
- Sowa J.F.. *Conceptual Structures. Information processing in mind and machine*. New York: Addison-Wesley, 1984.
- Sowa J.F.. Lexical Structures and Conceptual Structures, in J. Pustejovsky ed., 223-262, *Semantics and the Lexicon*. Dordrecht/Boston/London: Kluwer Academic Publishers, 1993.

- Sowa J.F.. Using a lexicon of canonical graphs in a semantic interpreter, in M.W. Evens ed., c. 5, 113-137, *Relational Models of the Lexicon*. Cambridge: Cambridge Univ. Press, 1988.
- Starren P. and Thelen M.. General dictionaries and students of translation: A report on the use of dictionaries in the translation process, in T. Magay, J. Zigány eds., 447-458, *Proceedings BudaLEX'88*. Budapest: akadémiai kiadó, 1990.
- Stein G.. From the Bilingual to the Monolingual Dictionary, in T. Magay, J. Zigány eds., 447-458, *Proceedings BudaLEX'88*. Budapest: akadémiai kiadó, 1990.
- Steiner E.. Producers — Users — Customers: Towards a Differentiated Evaluation of Research in Machine Translation, *Machine Translation*, 7, 281-284, 1993.
- Strossa P.. Machine Assisted Translation from English to a Slavic Language: What Linguistics and Programming Methodologies Can Do for It, *Machine Translation*, 9, 61-80, 1994.
- Tomaszczyk J.. Dictionaries: users and uses, *Glottodidactica*, 12, 103-119, 1979.
- Tsujii J., Fujita.. Lexical Transfer based on bilingual signs: Towards interaction during transfer, *Proc. 4th Conf. of the EACL* (Manchester), 275-280, 1989.
- Varantola K.. The Dictionary User as Decision Maker, *Proceedings EURALEX'94* (Amsterdam), 606-611. 1994.
- Velardi P., Pazienza M.T., De' Giovanetti.. Conceptual graphs for the analysis and generation of sentences, *IBM J. Res. Develop.*, vol. 32, no. 2, 251-267, 1988.
- Villegas L., Eberts R.E.. A neural network tool for identifying text-editing goals, *Int. J. of Human-Computer Studies*, vol. 40, no. 5, 813-833, 1994.
- van der Vliet H.. Conceptual Semantics for Nouns, *Proceedings EURALEX'94* (Amsterdam), 216-225, 1994.
- Voß A. and Karbach W.. MODEL-K: Making KADS Run, in G. Schreiber, B. Wielinga and J. Breuker eds., 203-220, *KADS. A Principled Approach to Knowledge-Based System Development*, Knowledge-Based Systems, vol. 11. London: Academic Press, 1993.
- Vossen P., Meijs W., den Broeder M.. Meaning and structure in dictionary definitions, in B. Boguraev, T. Briscoe eds., 171-192, *Computational Lexicography for Natural Language Processing*. New York: Longman, 1989.
- Wetter T.. First Order Logic Foundation of the KADS Coceptual Model, in Wielinga et al. eds., 356-375, *Current Trends in Knowledge Acquisition*. Amsterdam: IOS Press, 1990.
- White J.S.. Determination of lexical-semantic relations for multi-lingual terminology structures, in M.W. Evens ed., c. 8, 183-198, *Relational Models of the Lexicon*. Cambridge: Cambridge Univ. Press, 1988.
- Whitelock P.. Shake-and-Bake Translation, *Proc. COLING'92* (Nantes), 784-791, 1992.

Wilensky R., Chin D.N., Luria M., Martin J., Mayfield J., Wu D.. The Berkeley UNIX Consultant Project, *Computational Linguistics*, vol. 14, no. 4, 35-84, 1988.

Winston M.E., Chaffin R., Herrmann D.. A Taxonomy of Part-Whole Relations, *Cognitive Science*, no. 11, 417-444, 1987.

Yen J. and Lee J.. A Task-Based Methodology for Specifying Expert Systems, *IEEE Expert*, vol. 8, no. 1, 8-15, 1993.

Hiztegiak

Le Plus Petit Larousse. Paris: Librairie Larousse, 1980.

Aurrekoetxea G., Badiola R., Baraiazarra L., Goikoetxea J.L., Solabarrieta T., Uribarren P. *Euskararako Hiztegia*. Bilbo: Adorez, 1986.

Lhande P.. *Dictionnaire Basque-Français*. Paris: Gabriel Beauchesne ed., 1926.

Sarasola, I. *Hauta-lanerako Euskal Hiztegia*. Donostia: KUTXA, 1991.

Mel'cûk I.A.. *Dictionnaire explicatif et combinatoire du français contemporain*. Montréal: Les Presses de l'Université de Montréal, 1984.