

Corpusetan oinarritutako hiztegi elebidun berria sortzen

David Lindemann, Iñaki San Vicente

UPV/EHU, Elhuyar Fundazioa

david.lindemann@ehu.eus, i.sanvicente@elhuyar.com

Laburpena

Lan honetan, hiztegi gintza egitasmo berria aurkezten dugu, EHU-n garatzen ari dena: Euskara-alemanerazko *EuDeLex* hiztegi elebidun elektronikoa, hizkuntza-ikasleei zuzentzen dieguna. Sarrera, hizkuntza-bikote honentzako hiztegi berria baten beharrari dagokio.

EuDeLex egitasmoak ez du eguneratu edo egokitu litekeen aurrekaririk. Hala ere, baliabide lagungarriak ditugu hiztegia sortu eta editatzean: alde batetik, datu-base lexikografiko nahiz corpus elebakarrak, eta bestetik, sortu berri diren euskara-alemanerazko corpus paraleloak. Hiztegiaren makroegitura eta mikroegitura osatzeko corpusetan oinarritutako eman ditugun ikerketa-urratsak aurkezten ditugu.

Hitz gakoak: Hizkuntzalaritza, Hiztegi gintza, Corpus Hizkuntzalaritza, Euskara, Alemanera

Abstract

In this paper, we introduce EuDeLex, the new German-Basque bilingual electronic dictionary for language learners, which is currently being worked on at UPV/EHU. The introduction addresses the need for a new electronic dictionary for that language pair.

This new project does not have any lexicographical antecedents that could be updated or adapted. Nevertheless, on the one hand, existing monolingual lexicographical databases and corpora, and newly created German-Basque parallel corpora, on the other, support the creation and editing process of the new dictionary. We explain our workflow in corpus-based macrostructure and microstructure design and editing.

Keywords: Linguistics, Lexicography, Corpus Linguistics, Basque, German

1. Sarrera

Alemanera euskaratik ikasteko lehenengo testu-liburua 2007an argitaratu zen (Reuter eta Wolff 2007). Hurrengo urteetan, alemanera euskaldunei irakasteari buruzko zenbait ikerlan burutu ziren (Braun 2010; Reuter 2010; Wolff 2010). Haien arabera, urtean euskara L1 duten 1330 bat ikaslek hasten dituzte alemaneraren ikasketak bigarren hezkuntzako esparruan, eta beste 1.000 batek unibertsitatean eta Hizkuntza Eskola Ofizialetan. Euskaldunei alemanera irakasten diotenen, ikasleei hiztegi elebakarrak erabiltzea gomendatu arren, ikasleak hiztegi elebidunez baliatzen direla ikusten dute, eta euskara-alemanerazko eskaintzarik ezean, zubi-hizkuntza batetik jo ohi dutela, kasu gehienetan gaztelera-alemanerazko hiztegiak, azken haiek makroegiturari (lemategia) eta mikroegiturari (hiztegi-sarreraren diseinua eta edukia) dagokienez askoz ere gehiago eskaintzen dietela egun euskara-alemanerazko hiztegi bakarra den poltsiko neurriko paperezko hiztegiak baino (Martínez Rubio 2007). Horretaz gain, hiztegien erabilpenaren inguruko ikerketak argi utzi du ikasleek hiztegi elektronikoak erabiltzeko ohitura hartu dutela, paperezko hiztegi hobeak eskuragarri izan gorabehera (Domínguez in press¹). Hutsune hau betetzeko asmoz, hizkuntza-ikasleentzako euskara-alemanerazko hiztegi elektronikoa garatzea dugu helburu, EHU-ko web atariaren bitartez *EuDeLex* izenpean dohain eskuragarri jarriko dena.²

Jarraian, *EuDeLex* hiztegia sortzeko bidean eman ditugun zenbait ikerketa-urrats aurkeztuko ditugu, bi ataletan banaturik, makroegituraren eta mikroegituraren edukia osatzeari dagokionez, hurrenez hurren. Laugarren atalak zenbait ondorio eta gerorako erronkak azaltzen ditu.

1 [Artikulu hau argitaratzeko zorian dago; idazlan honen azken bertsiorako, erreferentzia eskuragarria izango da.]

2 Erabiltzaile-interfazearen lehenengo bertsioa <http://www.ehu.eus/eudelex> helbidean eskuragarri dago. Artikulu honetan lekurik ez duten *EuDeLex* hiztegiaren garapenaren beste alderdi batzuez, ikus Lindemann (2014a).

2. Corpusetarik erauzitako datuetan oinarritutako makroegitura

2.1 Lemategiak eraikitzeako irizpideak

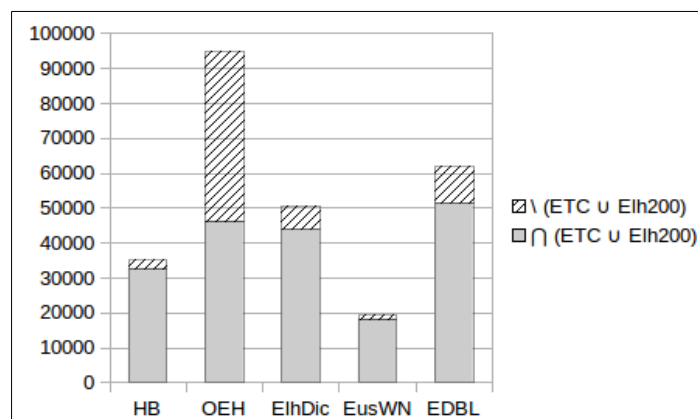
Hiztegi-intzak eta Corpus Hizkuntzalaritzak elkar elikatzen dute hainbat alderdiri begira (ikus, besteak beste, Teubert 2001; Teubert 2002; baita Lindemann 2013 ere). Hala nola, gaurko hiztegi-intzak corpusetarik erauzitako datuetara jotzen du lemategia definitzeko orduan. Hizkuntza Azterketa eta Prozesamendua (HAP) alorrean garatutako teknikei esker, corpus-datu gordinetatik hainbat ebidentzia estatistiko erauzteko gai gara, eta horien artean, lemaren erabilera-maiztasuna. Maiztasuna bi zentzutan da garrantzitsua hiztegi-intzan:

- 1) Batetik, ikasleentzako hiztegi-lemategiak maiztasunean oinarritzea ohiko jokabidetzat har dezakegu (Quasthoff 2009), eta bestetik, erabilera-maiztasuna eta dagokion hiztegi-sarreraren ikustaldi kopurua erlazionatuta daude neurri handi batean (De Schryver et al. 2010; Wolfer et al. 2014);
- 2) Maiztasun-datuak hiztegi-sarreraren osagai gisa eskura jarri ahal zaizkio erabiltzaileari. Besteak beste, hizkuntza-ikaslearen beharrei begira, aspalditik oso zentzuduntzat jotzen da maiztasunari buruzko argibideak hiztegi-intzean ematea; ingelesa ikasten duenak, esaterako, 1990. hamarkadatik aurrera aurkitzen ditu maiztasun-datuak ikasleentzako hiztegi-intzean (Kilgarriff 1997).

2.2 Alemanerazko eta euskarazko lemategiak eraikitzeako metodoak

EuDeLex hiztegiaren alemanezko lemategia zehazteko lehengaita 40.000 sarrera dituen *DeReWo-40.000* maiztasun-zerrenda lematizatua (IDS 2009) izan da, egun alemanerazko corpus handienak diren eta prestigio handiko IDS erakundeak kudeatutako baliabideetan oinarritzen dena. Horren lehenengo % 10a (A hizkia hasten diren lemak) eskuz editatu ondoren, zerrendako lehen % 95a bere horretan aurkitu dugu hiztegiaren lemategian, eta eskuz moldatutakoak edo gehitutakoak gainontzeko % 5a baino ez du osatzen (xehetasun gehiagorako, ikus Lindemann 2014a; Lindemann 2014b). Zenbaki horiek adierazten dute zenbateraino den egokia zerrenda hori hiztegi baten sarrera multzoa zehazteko lanean, ia inolako eskulanik ez baitu behar.

1 irudia: Hiztegi-lemategiak corpusetan



Euskarazko oinarritzko lemategi bat proposatzeko, eredutzat hartu dugu alemanerazko *DeReWo* zerrendak sortzeko metodologia (xehetasun gehiagorako, ikus Lindemann eta San Vicente 2015). Alde batetik, *ETC* corpus (Sarasola, Landa eta Salaburu 2013) eta *Elhuyar Web-Corpus*-etatik (Elh124 eta Elh200, Leturia 2012; Leturia 2014), hau da, egun euskarazko corpus handienak diren baliabideetatik, erauzi ditugun maiztasun zerrendez baliatu gara.³ Bestetik, aurretik argitaratutako hainbat euskarazko hiztegi eta datu-base lexikaletako lemategiak

3 ETC-ren kasuan, EHUko Euskara Institutuan prozesaturiko maiztasun-zerrenda lematizatua eskuragarri izan dugu.

eskuratu ahal izan ditugu.⁴ Bi multzo haien intersektzioa, hots: corpus batean *eta* hiztegi batean agertzen diren lemen multzoa, abiapuntuzat hartu dugu lemategiaren eskuzko edizioa burutzeko. 75.481 euskarazko lematako multzoa dugu honezkero, bakoitza corpusetatik erauzitako maiztasun-informazioarekin (ikus corpusetako ebakidurak (\cap) hiztegi-lemategi bakoitzarekin eta multzo disjuntuak (\setminus) 1 irudian).

Corpus hizkuntzalaritzako literatura klasikoan proposatutako atalasearen arabera, corpusetan 20 agerpenetatik gora duten 57.919 lema *EuDeLex* hiztegian sarrera-buru izateko hautagai egokitzat hartzen ditugu.⁵ Metodo automatikoen bitartez definituriko lemategi horren eskuzko edizioak argituko digu baieztapen horrek norainoko balioa duen.

3. Mikroegitura osatzeko datuak

3.1 Lemak entitate sintaktiko gisa

Lemategien konparaketa hau burutzeko eskuragarri izan ditugun euskarazko baliabide lexikaletako bat, *EDBL* (Aldezabal et al. 2001) datu-basea izan da, corpus-datu gordinak lematizatzen eta etiketa morfosintaktikoez osatzen dituen *EusTagger* tresnak (Aduriz et al. 1996) ere erabiltzen duen datu-base bera. Datu-base horrek lema-ikur bakoitzari egoki ahal zaizkion kategoria eta azpi-kategoria gramatikalak dakartza, ondorengo adibideak erakusten duen bezala:

1 Taula: *alegia* lema-ikurra *EDBL* datu-basean eta *Elh200* corpusean

| EDBL lema eta etiketa | Kategoria | Agerpenak Elh200 |
|-----------------------|-------------------|------------------|
| alegia_LOT_LOK | lokailua | 41.106 |
| alegia_IZE_ARR | izen arrunta | 4.208 |
| Alegia_IZE_LIB | leku-izen berezia | 921 |

EDBL eta corpusetako ebakidurako lemei dagokienez, ondorengo kopuruak ditugu: 51.157 lemetatik, 46.124 dira 20 agerpenetatik gora dutenak. Ondorengo arrazoiengatik proposatzen dugu *EuDeLex* hiztegiaren abiatzeko euskarazko lemategia multzo horren gainean eraikitzeko:

- 1) *EDBL*-eko etiketa sintaktikoen gainean, hiztegi-sarreraren oinarritzko egitura bat eraiki daiteke. Lema-ikur baten azpian, lema homografoak (ikus adibidea 1 taulan) entitate sintaktiko gisa antolatuko lirarteke.
- 2) Gainera, fintasun maila horretan zehaztutako entitate sintaktiko bakoitzarentzat maiztasun-datuak ditugu; eskuragarri ditugun beste datu-iturrietan, aldiz, lemaren kategoria hutsaz bestelako etiketarik ez dugunez, haiei dagokienez ezin dugu maiztasun daturik azpi-kategoria gramatikalarentzat kalkulatu.
- 3) *EDBL*-ko datuak hiztegi gintzan abiapuntuzat hartu eta beste iturrietatik osatzeko erabakiak eta *EDBL*-ko lemategi hura eskuz editatzeak ahalbidetuko dute datu-base hori bera eguneratu eta aberastea.

3.2 Aleman-euskarazko ordain-bikoteak definitzeko metodoak

Aipatu dugun bezala, *EuDeLex* hiztegia hutsetik lantzen dugu, abiapuntu gisa balia daitekeen glosario elebidunik ez baikenuen eskuragarri. Eskuzko lana murrizteko asmotan, abiatzeko

4 Euskaltzaindiaren *Hiztegi Batua* (HB), *Orotariko Euskal Hiztegia* (OEH), *Elhuyar Euskara-Gaztelania* hiztegiaren euskarazko lemategia (ElhDic), *Euskal WordNet* (EusWN) eta *EDBL*; Erreferentzia zehatzak Lindemann eta San Vicente (2015) lanean.

5 Hiztegi gintza lanetan lema baten esanahia zehazteko ebidentzia nahikoa izateko, 20 agerpenetan finkatzen du Sinclair-rek (2005) gutxieneko agerpen-kopurua. *EuDeLex* hiztegiaren editatze lanetarako giza baliabide murriztuak izatearen ondorioz eskuzko lana murriztu behar izanez gero, gutxieneko kopuru handiagoa finkatu beharko da.

glosario-zirriborroak garatzeko metodo sorta garatu dugu, jarraian laburbiltzen duguna (ikus Lindemann et al. 2014 xehetasunetarako).

Alemanerazko A hizkirako hiztegi-sarrerak editatu ondoren, hau da: proposatutako lemategiaren %10erako (3.400 bat lema), alemanerazko lemaren polisemia-desanbiguate eta euskarazko ordainak barne, baliatu ditugu datu haiek jarraian aipatuko ditugun glosario elebidunak automatikoki sortzeko bost metodo ebaluatzeko. Alemanerazko *DeReWo-40.000* (ikus 2.1 atala) zerrendako 40.000 lementzat saiatu gara metodo hauen bitartez euskarazko ordainak lortzen.

- 1) EHUko Letra Fakultatean garatu berri den euskara-alemanerazko literatura-corpus elebiduneko eskuragarri genituen datuak - 81 alemanerazko liburu eta haien euskarazko itzulpenak; 2 milioi testu-hitz hizkuntza bakoitzean - (Sanz, Zubillaga eta Uribarri 2015) erabili ditugu honako bi tresna hauen bitartez glosarioak sortzeko: *GIZA (Och eta Ney 2000)* eta *Bifid* (Nazar 2012) lerrokatze tresnak, estatistikan oinarrituta testu paraleloetatik itzulpen-ordainak hautatzen dituenak.
- 2) Asmo honetarako eratu dugun Biblia-corpus elebiduna (640.000 testu-hitz hizkuntza bakoitzean) erabili dugu aipatutako lerrokatze tresnek glosarioak sor ditzaten.
- 3) Elhuyar Fundazioak garatutako *Pibolex* hiztegiaren (Saralegi, Manterola eta San Vicente 2012) euskaraz-alemanerazko bertsioak ematen dituen ordain-bikoteak ebaluatu ditugu.
- 4) Existitzen diren kode ireki eta lizentzia libreko datu-base lexikal elebatar biren hizkuntzaz-hizkuntzako loturak erauzi ditugu, EU-DE lotura duten item lexikoak glosario-sarrera gisa batzeko: *Wiktionary* eta *Wikipedia*.
- 5) *WordNet* datu-baseetatik erauzi ditugu datuak, *EusWN* datu baseko item lexikoak, ingelesezko *WordNet*-en bitartez alemanerazko *GermaNet* datu-baseko edukiekin lotuz, alegia.⁶

Ondorengoak dira eraikitako glosarioen ebaluazio kuantitatibo eta kualitatiboaren emaitzak:

- Aipatutako metodo guztiak konbinatuz, alemanerazko lemategiaren %80arentzat lortu dugu euskarazko ordain bat edo gehiago.
- Alemanerazko lemategiaren %60arentzat lortu dugu eskuzko ebaluazioan egokitzen jo dugun ordain bat edo gehiago, gainontzekoak desegokiak izanik. Ordain egokia(k) dituen lema erdiak (%30) ordain egokiez gain ez du ordain desegokirik.
- Eskuzko ebaluazioaren emaitzen arabera, hiru taldetan sailkatu ditugu metodoak. Ondorengo taulan, sailkatutako datuak ikus daitezke:

2 Taula: Glosario elebidunak sortzeko metodoen sailkapena, emaitzaren kalitatearen arabera

| Taldea | Baliabideak | Ordain egoki bat edo gehiago dituen alemanerazko lema |
|--|--|---|
| (1) ia erabateko doitasuna | <i>Wiktionary, Wikipedia, Bifid</i> (Literatura corpusa) | %40,6 |
| (2) doitasun altua (ordain desegoki gutxi) | <i>Bifid</i> (Biblia), <i>WordNet</i> | %20 |
| (3) doitasun baxua (ordain desegoki asko) | <i>Giza, Pibolex</i> | %18,7 |

Lortutako ordain-bikoteak *EuDeLex* hiztegiko datu-basean sarrera emango diogu, aipatu berri dugun sailkapenaren araberrako irizpideak izanik: (1) taldeko datuak, ia datu desegokirik ez egotea espero dugunez, zuzenean sartuko ditugu datu-basean, ondoren eskuz editatzeko; (2) taldeko datuak, berriz, eskuz errebisatuko ditugu datu-basean sartu baino lehen; eta (3) taldeko

6 Erreferentziak Lindemann et al. (2014) lanean.

datuak, datu desegoki franko tartean duten glosarioen taldekoak, ez ditugu datu-basean sartuko.

Eskuzko lana nabarmen erraztuko duen datu-multzoa lortu dugu. Horretaz gain, honako bi baieztapen hau azpimarratu nahi dugu: Proposatutako metodo sorta printzipioz edozein hizkuntza-bikoterako balio du, eta beste hainbat konbinaziotan errepika daiteke, bai hiztegi-zirriborro gisa balia daitekeen glosario elebidunak lortzeko, eta bai hizkuntza-bikote ezberdinen arteko erkaketak egiteko.

4. Ondorioak eta gerorako erronkak

Eguneratu edo egokitu ahalko litzatekeen hiztegitzinta-lan aurrekaririk gabeko hizkuntza-bikote honetako hiztegi elektroniko berri baten oinarri batzuk definitu ahal izan ditugu, bai makroegituraren, eta bai mikroegituraren osaketari begira. Bi kasuetan, hizkuntza corpus elektronikoak ezinbesteko baliabideak direla argi geratu da. Bestalde, ez dago zalantzarik hiztegitzintzan eskuzko lana ere berebiziko baliabidea izaten jarraitzen duela, automatikoki sortutako zirriborro-datuak egiaztatu eta osatzeko.

Hiztegi elebidun bat hutsetik hasteko zenbait abiapuntu baino ez ditugu deskribatu lan honetan. Hasieratik balio digun asmo bikoitzari jarraituko diogu hemendik aurrera ere: Alde batetik, *EuDeLex* euskara-alemanerazko hiztegi honetako lanak aurreratzea, hizkuntza-ikasleari, batik bat, baliabide egokia eskaintzeko, eta bestetik, euskararekiko hiztegitzinta elebiduneara oraindik landu gabeak diren beste hainbat hizkuntza-bikotetan aplika zitzaketen oinarritzko irizpide eta metodoak landu eta proposatzea.

5. Erreferentziak

- Aduriz, I., I. Aldezabal, I. Alegria, X. Artola, N. Ezeiza eta R. Urizar (1996): EUSLEM: A lemmatiser/tagger for Basque. In *Proceedings of EURALEX 1996*, 17–26. Göteborg: Göteborg University.
- Aldezabal, I., O. Ansa, B. Arrieta, X. Artola, A. Ezeiza, G. Hernandez eta M. Lersundi (2001): EDBL: a General Lexical Basis for the Automatic Processing of Basque. HAL - CCSD.
- Braun, S. (2010): Die Rolle der Muttersprache im Deutschunterricht in zweisprachigen Gebieten am Beispiel des Baskenlandes. In C. Jarrilot (arg.): *Bestandsaufnahme der Germanistik in Spanien: Kulturtransfer und methodologische Erneuerung*, 25–36. Peter Lang.
- De Schryver, G.-M., D. Joffe, P. Joffe eta S. Hillewaert (2010): Do dictionary users really look up frequent words?—on the overestimation of the value of corpus-based lexicography. *Lexikos* 16 (1).
- IDS (2009): Korpusbasierte Wortgrundformenliste DEREWO, v-40000g-2009-12-31-0.1, mit Benutzerdokumentation. Institut für Deutsche Sprache, Programmbereich Korpuslinguistik.
- Kempcke, G. (2001): Polysemie oder Homonymie? Zur Praxis der Bedeutungsgliederung in den Wörterbuchartikeln synchronischer einsprachiger Wörterbücher der Deutschen Sprache. *Lexicographica*, Nr. 17: 61–68.
- Kilgarriff, A. (1997): Putting Frequencies in the Dictionary. *International Journal of Lexicography* 10 (2): 135–55.
- Leturia, I. (2012): Evaluating Different Methods for Automatically Collecting Large General Corpora for Basque from the Web. In *Proceedings of 24th International Conference on Computational Linguistics (COLING 2012)*, 1553–70. Mumbai, India.
- (2014): The Web as a Corpus of Basque. PhD Thesis, Donostia: UPV/EHU Lengoaia eta Sistema Informatikoak Saila.
- Lindemann, D. (2013): Bilingual lexicography and corpus methods. The example of German-Basque as language pair. *Procedia - Social and Behavioral Sciences* 95.
- (2014a): Creating a German-Basque Electronic Dictionary for German Learners. *Lexikos* 24 (1).
- (2014b): Zweisprachige Lexikographie des Sprachenpaares Deutsch-Baskisch. In Domínguez Vázquez, M. J., F. Mollica eta M. Nied (arg.): *Zweisprachige Lexikographie im*

- Spannungsfeld zwischen Translation und Didaktik*, 263–86. Lexicographica Series Maior. De Gruyter.
- Lindemann, D. eta I. San Vicente (2015): Euskarazko maiztasun lematagia gaurko teknologien ikuspuntutik. In B. Fernandez eta B. Salaburu (arg.): *Ibon Sarasolari Gorazarre*. Bilbao: UPV/EHU.
- Lindemann, D., X. Saralegi, I. San Vicente, I. Manterola eta R. Nazar (2014): Bilingual Dictionary Drafting. The example of German-Basque, a medium-density language pair. In *Proceedings of the XVI Euralex International Congress*, 563–76. Bolzano: EURAC.
- Martínez Rubio, E. (2007): *Euskara alemana hiztegia*. Donostia: Elkar.
- Nazar, R. (2012): Bifid: un alineador de corpus paralelo a nivel de documento, oración y vocabulario. *Linguamática* 4 (2): 45–56.
- Och, F. J. eta H. Ney (2000): Improved statistical alignment models. In *Proceedings of the 38th annual meeting of the Association for Computational Linguistics on Computational Linguistics*, 440–47. Hongkong: Association for Computational Linguistics.
- Quasthoff, U. (2009): Korpusbasierte Wörterbucharbeit mit den Daten des Projekts Deutscher Wortschatz. *Linguistik online* 39: 3.
- Reuter, D. (2010): Interkulturelles Lernen am Beispiel baskischer Muttersprachler im Deutschunterricht. In C. Jarrilot (arg.): *Bestandsaufnahme der Germanistik in Spanien: Kulturtransfer und methodologische Erneuerung*, 261–66. Peter Lang.
- Reuter, D. eta J. Wolff (2007): *Deutsch - Euskaldunentzat*. Donostia: Erein.
- Sanz, Z., N. Zubillaga eta I. Uribarri (2015): Estudio basado en corpus de las traducciones del alemán al vasco. In M. T. Sánchez Nieto (arg.): *Corpus-based Translation and Interpreting Studies: from description to application*, 211–35. Berlin: Frank & Timme.
- Saralegi, X., I. Manterola eta I. San Vicente (2012): Building a Basque-Chinese Dictionary by Using English as Pivot. In *Proceedings of the Eighth International Conference on Language Resources and Evaluation*. Istanbul.
- Sarasola, I., J. Landa eta P. Salaburu (2013): Egungo Testuen Corpora. UPV/EHU.
- Sinclair, J. (2005): Corpus and text-basic principles. In M. Wynne (arg.): *Developing linguistic corpora: A guide to good practice*, 1–16. Oxford: Oxbow Books.
- Teubert, W. (2001): Corpus Linguistics and Lexicography. *International Journal of Corpus Linguistics*, Nr. 6: 125–53.
- (2002): The Role of Parallel Corpora in Translation and Multilingual Lexicography. In B. Altenberg eta S. Granger (arg.): *Lexis in Contrast: Corpus-Based Approaches*, 189–214. John Benjamins Publishing.
- Wolfer, S., A. Koplénig, P. Meyer eta C. Müller-Spitzer (2014): Dictionary Users do Look up Frequent and Socially Relevant Words. Two Log File Analyses. In *Proceedings of the XVI Euralex International Congress*, 281–90. Bolzano/Bozen: Eurac.
- Wolff, Jürgen (2010): Sprachvergleich Baskisch/Deutsch und seine Auswirkungen für den Unterricht. In C. Jarrilot (arg.): *Bestandsaufnahme der Germanistik in Spanien: Kulturtransfer und methodologische Erneuerung*, 37–42. Peter Lang.

6. Eskerrak eta oharrak

Emaizta hauek izan dituen ikerketak diru-laguntza jaso du (Europar Batasunaren *Seventh Framework Programme for research, technological development and demonstration*, hitzarmen zbk. 613465; IT665-13 (Eusko Jaurlaritza)).

Ikerketan zehar aholku eta laguntza eman eta dagozkion argitalpenetan koautore gisa jardun duten ondorengo hizkuntzalari konputazionalari esker onak ematen diegu: Iker Manterola, Rogelio Nazar, Xabier Saralegi.

Lan honek Miren Azkarate irakaslearen zuzendaritzapean EHuko Euskal Hizkuntza eta Komunikazioa sailean David Lindemann burutzen ari den “Hiztegitintza Elebiduna: Euskara-Alemanera” izenburuko doktore tesiaren bi atalen edukia aurreratzen ditu.