Creating a German–Basque Electronic Dictionary for German Learners

David Lindemann, *Department of Basque Language and Communications, UPV-EHU University of the Basque Country, Donostia, Spain (david.lindemann@ehu.es)*

Abstract: In this paper, we introduce the new electronic dictionary project *EuDeLex*, which is currently being worked on at UPV-EHU University of the Basque Country.[1] The introduction addresses the need for and functions of a new electronic dictionary for that language pair, as well as general considerations about bilingual lexicography and German as foreign language (GFL).

The language pair German–Basque, which can be called *less-resourced* or *medium-density*, does not have any lexicographical antecedents that could be updated or adapted. Nevertheless, existing monolingual lexicographical databases and a newly created German–Basque parallel corpus support the editing process of the new dictionary. We explain our workflow in macrostructure and microstructure design and editing, and propose a first iteration of the online user interface and publishing process.

Keywords: BILINGUAL LEXICOGRAPHY, ELECTRONIC DICTIONARIES, BASQUE LANGUAGE, GERMAN AS FOREIGN LANGUAGE, PARALLEL CORPORA, USER INTERFACE, WIKTIONARY

Opsomming: Die samestelling van 'n Duits–Baskiese elektroniese woordeboek vir Duitse aanleerders. In hierdie artikel word *EuDeLex*, die nuwe projek vir 'n elektroniese woordeboek wat tans aan die UPV-EHU Universiteit van die baskiese gebied saamgestel word, bespreek. In die inleiding word gewys op die behoefte aan en funksies van 'n nuwe elektroniese woordeboek vir hierdie taalpaar asook algemene aspekte van tweetalige leksikografie met Duits as vreemde taal.

Die taalpaar Duits–Baskies, waarna verwys kan word as 'n taalpaar met minder hulpmiddele en medium digtheid, het geen leksikografiese voorgangers wat hersien of aangepas kan word nie. Desondanks word die samestellingsproses van 'n nuwe woordeboek ondersteun deur bestaande eentalige leksikografiese databasisse en 'n nuwe Duits–Baskiese parallelkorpus. Die werkswyse word bespreek m.b.t. die ontwerp van die makro- en mikrostruktuur en die redigering, en voorstelle word gemaak vir 'n eerste weergawe van 'n aanlyn koppelvlak en die publikasieproses.

Sleutelwoorde: BASKIESE TAAL, DUITS AS VREEMDE TAAL, ELEKTRONIESE WOORDEBOEKE, GEBRUIKERSKOPPELVLAK, PARALLELKORPORA, TWEETALIGE LEKSIKOGRAFIE, WIKTIONARY

1. Introduction

Basque is today, together with Spanish, official language for the Spanish state territories Bizkaia, Gipuzkoa and Araba (*Comunidad Autónoma del País Vasco*) and the northern half of Goi Nafarroa (*Comunidad Foral de Navarra*). In the territories belonging to the French state (Lapurdi, Behe Nafarroa and Zuberoa), Basque lacks official status. In the first three areas mentioned, about 30% of the 1.9 million inhabitants are regarded as active bilinguals (Basque and Spanish), and another 11% are regarded to be passive bilinguals (who understand both, but speak only Spanish). For the entire Basque Country, the figures given for active bilinguals are about 25%, the passive being another 10.5%.[2] In Bizkaia, Gipuzkoa and Araba, today more than a half of secondary and high school students study partly or entirely in Basque language. A third of the nearly 50,000 students at the public University of the Basque Country opts for Basque as teaching language, and the trend is rising.[3]

For German as a foreign language, the first text book written in German and Basque, without resorting to Spanish or French and designed for Basque-L1 GFL learners, was published in 2007 (Reuter and Wolff 2007). In the following years, a series of studies about teaching German to Basque-L1 learners appeared (Braun 2010, Reuter 2010, Wolff 2010). German teachers in the Basque Country have all made the same experience: Regardless of the teacher having recommended the use of monolingual dictionaries, learners stick to bilinguals, and lacking a suitable offer for German–Basque, they use Spanish–German dictionaries, which in terms of macro- and microstructure have a lot more to offer to them than the only available German–Basque pocket-size dictionary (Martínez Rubio 2007).[4/5] In order to fill this gap, it is the author's aim to propose and ultimately to provide an alternative, a German–Basque dictionary for Basque-L1 GFL learners, to be made freely available to all users through a web interface hosted by the University.[6]

1.1 *Editing Software*

By means of modern *Dictionary Writing Systems* (DWS) as, for example, *TshwaneLex* (Joffe, De Schryver and Prinsloo 2003, De Schryver and Joffe 2005), it is possible to create *Multifunctional Lexical Databases* (about the concept, Pajzs 2009) based on XML-coded microstructures. Different monofunctional dictionaries, which respond to different dictionary functions (Tarp 1995, Tarp and Bergenholtz 2005) can be designed as derivatives from that database, i.e. dictionaries that suit to one user profile and communicative situation (as, for example, text reception or text production). The display of contents from the same source database could also be adapted on the fly to the users' needs, by fading in and out parts of the microstructure or changing the language in which lexicographical metadata are presented. In the age of electronic lexicography, limitations of space or specialized output format disappear.[7]

1.2 *Desiderata for a bilingual lexicography for learners of GFL*

In addition to considerations about what is to be expected from electronic dictionaries, and what expectations electronic dictionaries are able to fulfill (see, for instance, Kemmer 2010, Tarp 2012), studies about bilingual pedagogical lexicography (recently in Fuertes-Olivera 2010) have been published, as well as studies about electronic lexicography with the GFL learner as targeted user profile (Meliss 2013).

In this section we will list some features by which a bilingual dictionary could be considered suitable for the needs of GFL learners. Apart from general desiderata for bilingual lexicography, we can point out some features that are linked to students' difficulties in the learning process of GFL.

1.2.1 *Desiderata for Framing Structure and Macrostructure*

A bilingual dictionary should contain introductory and explanatory prefaces in both languages. User interfaces (UI) of electronic dictionaries should be able to toggle between both languages as metalanguage, i.e. all instructions apart from the lemma signs, synonyms and translation equivalents should be available in both languages.

In addition to canonical word forms that figure traditionally as lemma signs in dictionaries, a UI should also provide results if the user enters an inflected word form in the search box: It may provide a morphological analysis of the word form and a link to the corresponding lemma. Apart from single word lemma signs, also multi-word expressions such as light verb constructions or idiomatic phrases should be capable to be found in a dictionary.

1.2.2 *Desiderata for the Microstructure*

A bilingual dictionary article should be furnished not only with insightful instructions for word sense (polysemy) disambiguation and a mapping to suitable Translation Equivalents, but also with instructions related to morphology, syntax, and pragmatics. In the case of a dictionary for GFL learners these are the following:

- Inflection morphology paradigm for the German verb, noun or adjective
- For German verbs: auxiliary selection
- Instructions related to valency (argument structure realisation)
- German synonyms
- Frequency data
- Collocates
- Pragmatics (register)
- Example sentences from bilingual corpora

1.3 *Lexicography and Open Source*

For a *high-density* language like German, many printed and electronic resources are available today. Innumerable dictionaries have been compiled by lexicographers, and phonetic, morphological, syntactic and semantic information about the headwords given in those has been edited many times or updated from former editions. A new dictionary that can not take advantage of relevant antecedents and that is not born in the shade of a publishing house that possesses extensive lexicographical databases, nevertheless may start with data obtained from open sources, i.e. sources that are not only freely accessible from an online UI, but also available as complete source files under non-exclusive, non-proprietary licensing terms. For the case of German, today these are primarily the electronic lexicographical databases *openthesaurus* and the German edition of the crowdsourced *Wiktionary*, that are continuously growing and possibly will reach a professional level (Meyer and Gurevych 2010, 2012). Furthermore, for the definition of a German lemma list, the corpus based lemmatized frequency word lists released by the IDS under a Creative Commons licence (see section 2.1) are a suitable starting point.

Generally, lexicographical data from existing databases can be used for a new project in two ways: By importing data into the new database as draft for manual revision (Dictionary Drafting), or by dynamically including data from external sources in the search-result pages of an online dictionary, as described in section 3.

In the case of Basque, a *medium-density* minority language with less than one million speakers and a co-official legal status in parts of the territory it is spoken and a relatively minuscule web size (see Lindemann 2013), noteworthy open resources are available: The source files of *Hiztegi Batua* (Euskaltzaindia 2008), and the Basque WordNet (EusWN, see Pociello, Agirre and Aldezabal 2011), that are both published under Creative Commons licences.

1.4 *A German–Basque literary parallel corpus*

At the University of the Basque Country, a German–Basque Literary Corpus has been created (Sanz Villar 2013, Zubillaga 2013), using the content of 81 digital or digitized literary German originals and their official, direct translations into Basque. In its current version, the German–Basque translation corpus counts about 2 million tokens per language. The sentence alignment (146.000 sentence pairs) had to be revised manually, starting from an automatic alignment at paragraph level. Today, that corpus is the only parallel resource for German–Basque.[8] The value of parallel corpora for pedagogical lexicography (cf. Bowker 2010) as well as the value of parallel (literary) corpora for lexicography in general are beyond question (cf. Teubert 2002).

This parallel corpus has been imported into the *SketchEngine* (Kilgarriff, Rychly, Smrz and Tugwell 2004), a software system which is capable of dealing with parallel corpora. The German part was lemmatized and POS-tagged with *TreeTagger* (Schmid 1995), a tool that is built in to the *SketchEngine*. The Basque corpus has been lemmatized with *EusTagger* (Aduriz, Aldezabal, Alegria et al. 1996). As a result of the tagging and lemmatization, the corpus can be queried by lemma, in order to consider all appearances of inflected forms of the lemma in the corpus. This is a desirable feature for both dictionary entry drafting process and for display on dictionary search result pages.

In the dictionary editing process, data from this literary corpus has been particularly useful for German lemma signs denoting abstracts: For some abstract nouns and verbs, considered "hard tasks" for a lexicographer, where data from other sources does not lead to satisfying results, the corpus data provides groups of good Translation Equivalent candidates, each one reflecting a translator's choice in a particular context. Table 1 illustrates this by two German nouns, two often cited examples for "hardly translatable" German abstract nouns:

German lemma sign (counts)	*Basque TE from parallel corpus*
Gemütlichkeit (4)	goxotasun, patxada, lasaitasun, konfortea
Schadenfreude (10)	(voll Schadenfreude sein) zoritxarraz poztu (Schadenfreude empfinden) maltzur sentitu bozkario gozatze modu bat poz txiki bat alaitasun maltzur besteak umiliatzeko poza poz gaizto (aus Schadenfreude) besteren gaitzak ninduen pozten kalte poz

Table 1: German abstract nouns and Basque TE from literary parallel corpus

Apart from the described function as documentation for the lexicographer in the lexicographical workflow, a display of parallel corpus concordances as part of the search results in online bilingual dictionaries may be worthwhile (first mention in Atkins 1996, Dickens and Salkie 1996), and it is a desired feature for *EuDeLex*.

2. A new German–Basque Electronic Dictionary

EuDeLex is in its first stage of development. A macrostructure and a microstructure for DE>EU have been proposed, dictionary entries have been edited for around 10% of the planned lemma list (4,500), and a preliminary version of the online user interface is being tested.

This first-stage work is being done on the German–Basque side, but there is also the possibility of including a Basque–German part. There may also be occasion to widen the scope of the dictionary and adopt its structure towards other functions than the described.

The *EuDeLex* dictionary articles for German Letter A, already edited, have served as gold standard in the evaluation of six different corpus based and/or lexical knowledge based Bilingual Dictionary Drafting methods, which has been carried out by the author together with a group of computational linguists (Lindemann, Saralegi, San Vicente, Manterola and Nazar 2014). The (semi-)automatically produced bilingual glossaries are evaluated quantitatively and qualitatively. Results show that the described methods can greatly assist the editing of dictionary entries for the remaining 90% of the German lemma list. Furthermore, the set of Bilingual Dictionary Drafting methods described may serve as reference for lexicographical work on other language pairs that starts from scratch.

2.1 *Macrostructure*

It has been proved that the most frequent words are actually the words most frequently looked up by dictionary users; this is true for the top few thousand (De Schryver, Joffe, Joffe and Hillewaert 2006, Wolfer, Koplenig, Meyer and Müller-Spitzer 2014). At the same time, frequency data is useful information for both dictionary editor and user. Therefore, it makes sense to build a lemma list starting from corpus-based frequency lists and to include frequency data in the published dictionary. For German, lemmatized frequency lists based on large reference corpora are available under public licences (IDS 2009). In our workflow, we compare the DeReWo-40.000 frequency word list with human-revised lemma lists found in three editorial dictionaries, and delete, replace (by adapting to a word form used as lemma in our macrostructure) or add lemmata. DeReWo-40.000 contains the whole lemma list regarded as a GFL learner's basic dictionary (*Wortschatz Zertifikat Deutsch als Fremdsprache,* CEFR B2).

After editing dictionary entries for the German letter A, i.e. the first 3453 entries on the (alphabetically ordered) DeReWo-40,000, our adaptation omitted 132 of those 3453 entries from our lemma list. 101 of the non-imported lemmas were not found in the DUDEN, the first secondary source for reference, mostly so-called [semantically] transparent compound nouns. Another 24 were proper names of people or organizations ("AOK") or topo- and hydronyms ("Alster"), which we generally do not import, and 7 others. Another 37 DeReWo entries we have incorporated in a modified form (mainly led by the form listed in the DUDEN), such as "Abbrucharbeiten" instead of "Abbrucharbeit" or "aufrütteln" instead of "aufrüttelen." In the first edition of EuDeLex, we restrict the published dictionary articles to nouns, verbs, adverbs, and adjectives (see section 2.2.1), as function words in our context bear special difficulties that need further research.

For a definition of homonymy we follow Kempcke's (2001: 67) criterion of morphosyntactic disparity as a necessary condition: If homograph lemma signs follow different inflection morphology patterns or if they are morphological derivatives with different origins, they will be considered as homonyms, which appear in separate dictionary articles. We assume that it is helpful for students to distinguish words with different inflectional morphologies, which in the event they must learn separately.

In table 2, we list absolute amounts and percentages of nouns, verb infinitives, verb participles, adjectives, and adverbs in *EuDeLex* German Letter A (DeReWo Letter A edited as described above), DeReWo A-Z tagged with RF-Tagger (Schmid and Laws 2008), and the intersection of DeReWo and the German Wiktionary (wordlist and POS-tags parsed from de.wiktionary).

	EuDeLex DE Letter A		DeReWo A RFTagger		DeReWo A-Z RFTagger		DeReWo ∩ de.wiktionary	
Lemmata	*3,614*		*3,453*		*40,000*		*22,404*	
Nouns	1,963	54.32%	1,813	52.51%	25,529	63.82%	14,017	62.56%
Verb Infinitive	924	25.57%	878	25.43%	5,813	14.53%	3,203	14.30%
Verb Participle	268	7.42%	175	5.07%	819	2.05%	1,003	4.48%
Adjectives	413	11.43%	411	11.90%	5,243	13.11%	3,418	15.26%
Adverbs	150	4.15%	60	1.74%	579	1.45%	585	2.61%
Others	0	0.00%	116	3.36%	2,017	5.04%	461	2.06%
Total	*3,718*	*102.88%*	*3,453*	*100.00%*	*40,000*	*100.00%*	*22,687*	*101.26%*

Table 2: Word classes in the German lemmalist

The above figures show a disproportionately high rate of verbs in German letter A, which is due to the presence of three prepositions, and consequently of verbs formed by a chain of preposition plus verb (such as *abfahren*, *ankommen*, *auffallen*). The disparity between verb participle counts in *EuDeLex A* and *DeReWo A RFTagger* can be explained by the fact that, in *EuDeLex,* participial adjectives are tagged as both adjectives and verb participles (together with a link to the corresponding verb infinitive). The higher number of adverbs in *EuDeLex* is due to the disambiguation between adjectives and adjectives with adverbial use (see section 2.2.4), and a double tagging in cases where both uses are found in German. As *EuDeLex* POS-tags are manually set, these figures also demonstrate the high grade of trustworthiness of the *RFTagger* tool.

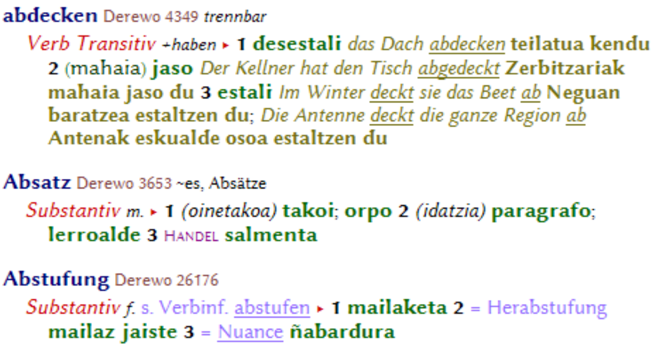
For Basque, the method for lemma list building we are aiming at is the following: The Basque Language Academy *Euskaltzaindia* provides a dictionary with a 55.000 entry corpus-based lemma list under a public licence (Euskaltzaindia

2008), each entry of which has been revised and approved by the Academy's lexicographical board. A comparison of that list with lemmatized frequency word lists extracted from Basque Web Corpora is in course, in order to define a core lemma list as starting point for *EuDeLex* EU>DE (Lindemann and San Vicente in prep., for a survey of Basque corpora see Leturia 2012).

2.2 *EuDeLex Microstructure*

2.2.1 *Structure of the editorial dictionary articles*

The aim to provide a dictionary in the reasonably near future lies in conflict with limited human and other resources in dictionary entry editing. For this reason, we don't hesitate to take pragmatic decisions in prioritizing the lexicographical working agenda. The aim of this dictionary in its first version we define as to provide useful word sense disambiguation (polysemy discrimination) information to Basque-L1 GFL learners and translators. Accordingly, the Basque equivalents provided in the first version of the German–Basque part of *EuDeLex* will be furnished with additional information when that seems useful for polysemy discrimination: Possible information to be included is specification of semantic domain, synonyms and/or usage examples, as shown in Fig. 1:



abdecken Derewo 4349 *trennbar*
Verb Transitiv *+haben* ▸ **1 desestali** *das Dach abdecken* **teilatua kendu** **2** (mahaia) **jaso** *Der Kellner hat den Tisch abgedeckt* **Zerbitzariak mahaia jaso du** **3 estali** *Im Winter deckt sie das Beet ab* **Neguan baratzea estaltzen du**; *Die Antenne deckt die ganze Region ab* **Antenak eskualde osoa estaltzen du**

Absatz Derewo 3653 ~es, Absätze
Substantiv *m.* ▸ **1** *(oinetakoa)* **takoi**; **orpo** **2** *(idatzia)* **paragrafo**; **lerroalde** **3** HANDEL **salmenta**

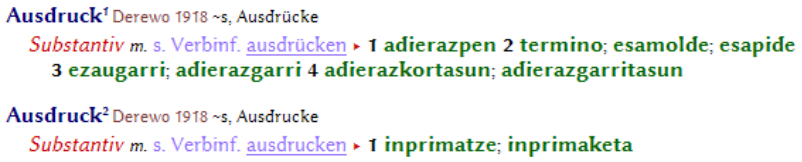
Abstufung Derewo 26176
Substantiv *f.* s. Verbinf. abstufen ▸ **1 mailaketa** **2** = Herabstufung **mailaz jaiste** **3** = Nuance **ñabardura**

Figure 1: Excerpts from *EuDeLex*

After the first version is edited, we plan to provide all dictionary entries with such information. Merely with regard to the domain specification there is a claim for completeness from the first version on, in order to provide domain related glossaries as derivatives of *EuDeLex*.[9]

2.2.2 *German nouns in dictionary articles*

Following the criteria described in section 2.1, we provide, as seen in Fig. 2, the information on the German noun "Ausdruck" in two different dictionary entries and add to both information about inflection morphology (genitive sg. and nominative pl. forms) together with the grammatical gender.[10]

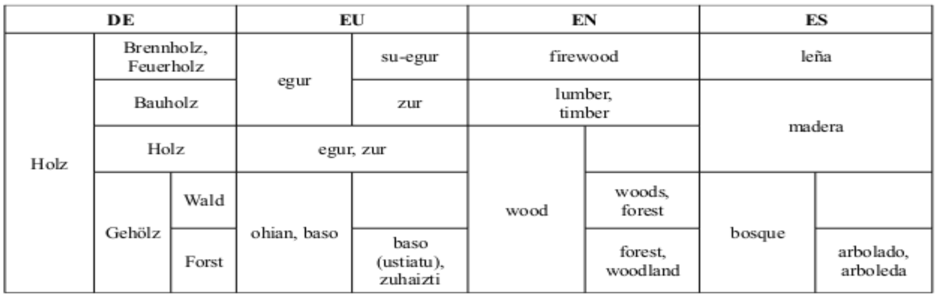


Ausdruck[1] Derewo 1918 ~s, Ausdrücke
Substantiv *m.* s. Verbinf. ausdrücken ▸ **1 adierazpen 2 termino**; **esamolde**; **esapide 3 ezaugarri**; **adierazgarri 4 adierazkortasun**; **adierazgarritasun**

Ausdruck[2] Derewo 1918 ~s, Ausdrucke
Substantiv *m.* s. Verbinf. ausdrucken ▸ **1 inprimatze**; **inprimaketa**

Figure 2: Excerpts from *EuDeLex*

Below this first syntactic categorization, the polysemy follows in numbered word sense sections. In distinguishing word senses, we try to take into account possible asymmetric lexicalization (about the concept, cf. Hartmann 2007: 33) in German and Basque. Lexical asymmetry of a group of hyper- and hyponyms can be exemplified as illustrated in Figure 3. When there is no suitable bilingual dictionary available for the language pair a dictionary user wants to find information on, asymmetry schemes like this show why using a third language's dictionaries as "pivot" or "bridge" may mislead, why the polysemy of a German word must be mapped lexicographically to a target language's lexical units in a specific way:



<table>
<tr><th colspan="3">DE</th><th colspan="2">EU</th><th colspan="2">EN</th><th colspan="2">ES</th></tr>
<tr><td rowspan="5">Holz</td><td colspan="2">Brennholz,
Feuerholz</td><td rowspan="2">egur</td><td>su-egur</td><td colspan="2">firewood</td><td colspan="2">leña</td></tr>
<tr><td colspan="2">Bauholz</td><td>zur</td><td colspan="2">lumber,
timber</td><td colspan="2" rowspan="2">madera</td></tr>
<tr><td colspan="2">Holz</td><td colspan="2">egur, zur</td><td rowspan="3">wood</td><td></td></tr>
<tr><td rowspan="2">Gehölz</td><td>Wald</td><td rowspan="2">ohian, baso</td><td></td><td>woods,
forest</td><td rowspan="2">bosque</td><td></td></tr>
<tr><td>Forst</td><td>baso
(ustiatu),
zuhaizti</td><td>forest,
woodland</td><td>arbolado,
arboleda</td></tr>
</table>

Figure 3: Asymmetric lexicalization

The above mentioned German terms therefore should be paired with the following equivalents (see table 3):

Holz	1. egur, zur; 2. ohian, baso
Gehölz	baso, zuhaizti
Wald	ohian, baso
Forst	baso, zuhaizti (ustiatu)
Brennholz	su-egur

Feuerholz	su-egur
Bauholz	(eraikuntza) zur

Table 3: German "Holz" and hyponyms, and Basque equivalents

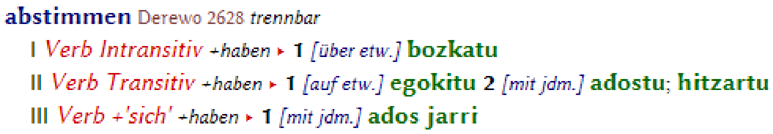
2.2.3 *German Verbs in dictionary articles*

Following microstructures established in some GFL dictionaries (Langenscheidt pocket Dictionaries, PONS dictionary GFL) we organize the verb entry first and foremost syntactically, i.e. below a syntactical entity (transitive, intransitive, reflexive/reciprocal, non-personal use, see Fig. 3 below).[11]

Some arguments in favour of such syntactic element ordering may be summarized as follows (cf. also Marello 2010, Dentschewa 2006: 124f):

- It helps as a first orientation within longer dictionary entries
- In text reception, it is a strategy of advanced GFL learners to identify the verb (the meaning of which they possibly don't know) and its arguments as syntactic entities (subject, accusative object, dative object etc.) and then to proceed to semantics.
- In GFL production, it is often not the meaning but the syntactic properties of a word that dictionary users want to be sure about: Can I use this verb as a transitive? Which auxiliary is selected by that verb in an intransitive sentence?

Criticism of syntactic ordering is based on questioning the ability of dictionary users to deal with basic grammatical concepts such as "transitivity". An alternative could be a labelling as proposed in PONS: "Mit OBJ" (with object), "Ohne OBJ" (without object), "Mit Sich" (with "sich").[12]



abstimmen Derewo 2628 *trennbar*

I *Verb Intransitiv* *+haben* ▸ **1** *[über etw.]* **bozkatu**

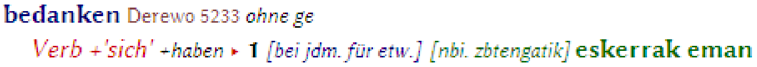
II *Verb Transitiv* *+haben* ▸ **1** *[auf etw.]* **egokitu** **2** *[mit jdm.]* **adostu**; **hitzartu**

III *Verb +'sich'* *+haben* ▸ **1** *[mit jdm.]* **ados jarri**

Figure 4: Excerpt from *EuDeLex*

As support for a usually permanent problem in the learning process of GFL, the verb and its auxiliary selection in perfect tense, the appropriate auxiliary verb is indicated together with the syntactical tag. For a lexicographical illustration of valency, we use the so-called "valence formula" (Wolski and

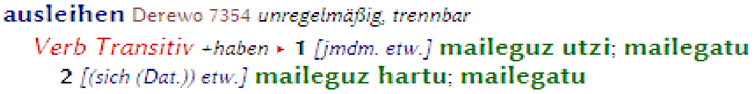
Cyffka 2011: 12). By this formula, the user gets clues about correct connection of actants with prepositions not only for text production, but also in text reception, as in cases like in Fig. 4 the preposition reveals the necessary for identifying the German verb's word sense. Following the above-defined initial target of this dictionary, in the first version we don't provide translations of the German valency formula. A future upgrade of the existing microstructure targeting German-speaking learners of Basque could look like the following:



bedanken Derewo 5233 *ohne ge*
Verb +'sich' +haben ▸ **1** *[bei jdm. für etw.] [nbi. zbtengatik]* **eskerrak eman**

Figure 5: Excerpt from *EuDeLex*

By means of valency formula we also specify syntactic properties of verbs as refinement of the rough classification in the above-described sections, primarily ditransitivity ("somebody something"). Reflexive and reciprocal structures with a reflexive pronoun in dative case, as opposed to those with an accusative pronoun, are not grouped in the "with SICH" section, but are marked as transitive verb with the additional tag *[jmdm. etw.]* ("somebody something"), as shown in Fig. 6:

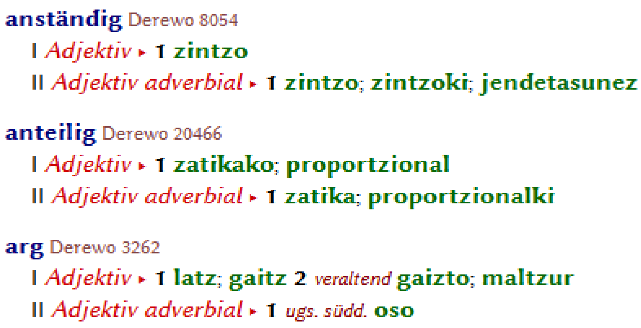


ausleihen Derewo 7354 *unregelmäßig, trennbar*
Verb Transitiv +haben ▸ **1** *[jmdm. etw.]* **maileguz utzi**; **mailegatu**
2 *[(sich (Dat.)) etw.]* **maileguz hartu**; **mailegatu**

Figure 6: Excerpt from *EuDeLex*

2.2.4 *Adjectives and adverbs in dictionary articles*

Regarding adjectives and adverbs, we must keep in mind morphosyntactic symmetries and asymmetries in both German and Basque. As the examples in Fig. 6 show, German adjectives in attributive or adverbial usage require morphologically or lexically different Basque equivalents; the attributive adjectives must be inflected in both German and Basque (being part of a nominal or determinative syntagma), while the adverbial is used without inflection in both languages (which is characteristic for an adverbial syntagma). In Basque, this category change in use of certain lexemes between adjective or adverb requires certain affixes *(e.g. -ki, -to, -ka* to obtain adverbs, and *-ko* to obtain adjectives), which also applies to some German lexemes (cf. allein$_{\text{ADV}}$ vs. alleinig$_{\text{ADJ}}$). The Basque equivalents of German adjectives and adverbs can also appear as postposition syntagma (jendetasun$_{\text{NOM}}$-ez$_{\text{POSTP}}$, "with respectability").



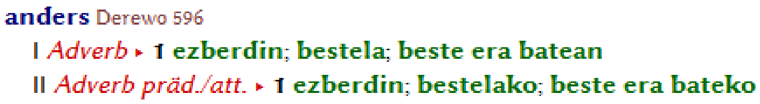
anständig Derewo 8054
I *Adjektiv* ▸ 1 zintzo
II *Adjektiv adverbial* ▸ 1 zintzo; zintzoki; jendetasunez

anteilig Derewo 20466
I *Adjektiv* ▸ 1 zatikako; proportzional
II *Adjektiv adverbial* ▸ 1 zatika; proportzionalki

arg Derewo 3262
I *Adjektiv* ▸ 1 latz; gaitz 2 *veraltend* gaizto; maltzur
II *Adjektiv adverbial* ▸ 1 *ugs. südd.* oso

Figure 7: Excerpts from *EuDeLex*

In various cases, lexemes of both languages only appear in one of the two categories, i.e. they can not be modified by affixes and made use of in the other category. As in this context either German nor Basque has standard rules that would apply universally, GFL learners would ask their dictionary for proper uses of the doubtful lexeme as adjective or adverb (see Fig. 8).[13]



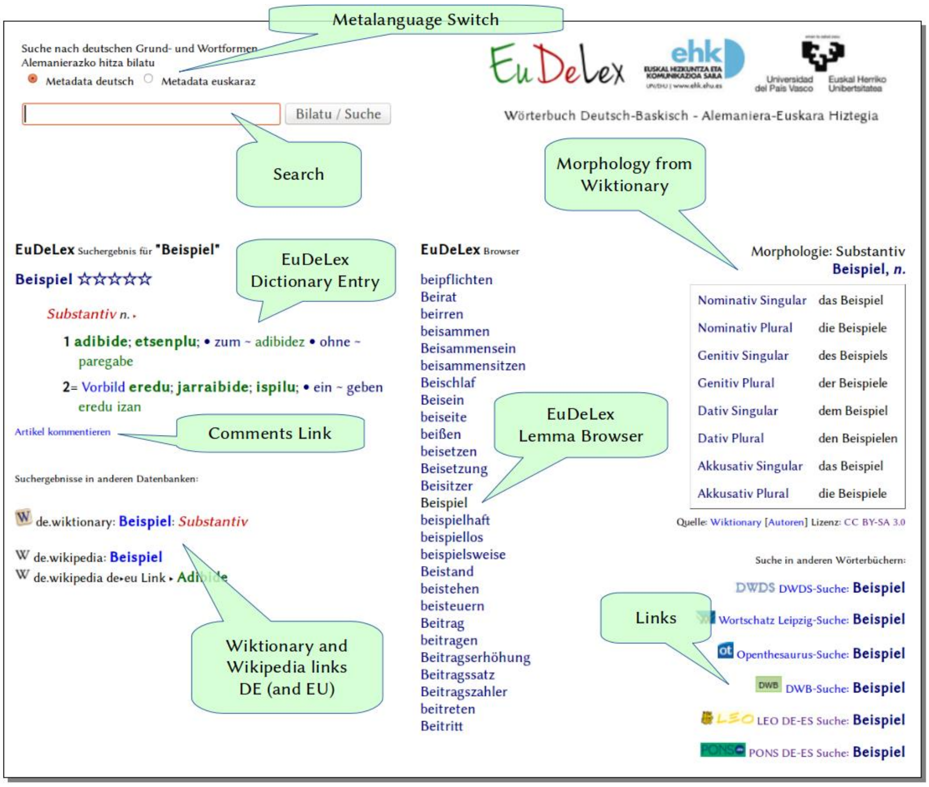
anders Derewo 596
I *Adverb* ▸ 1 ezberdin; bestela; beste era batean
II *Adverb präd./att.* ▸ 1 ezberdin; bestelako; beste era bateko

Figure 8: Excerpts from *EuDeLex*

Consequently, the entry for German adverbs in use as adverbial or as a predicate noun is organized in a similar way: A German adverb in predicative use ("Jazz ist anders$_{\text{ADV}}$" ("Jazz is different")) does not allow inflection; the equivalent use in Basque requires an adjective, i.e. it necessarily requires inflection ("Jazz-a ezberdin$_{\text{ADJ}}$-a$_{\text{DET}}$/bestelako$_{\text{ADJ}}$-a$_{\text{DET}}$ da") The same is true for an attributive use of German Adverbs "die$_{\text{DET}}$ Sitzung$_{\text{NOM}}$ gestern$_{\text{ADV}}$" vs. "[atzo$_{\text{ADV}}$-ko$_{\text{POSP}}$ batzar$_{\text{NOM}}$]-ra$_{\text{DET}}$" ("the meeting yesterday/yesterday's meeting"). In Fig. 8, for German adverbs, on the one side non-inflectable equivalents are given for an adverbial use, and inflectable equivalents for a predicative or attributive use, on the other.

3. *EuDeLex*: Online-Publishing

As of May 2013 *EuDeLex* is *online*.[14] The UI (see Fig. 8 below) is based on a single PHP script. On top of the home page, the user is prompted for a German headword or word form to look for. Starting with the fourth character that is typed, a search routine compares the query string with the *EuDeLex* lemma list, and responds to the user with automatic suggestions to complete a headword. Placed before the search box, switch buttons are provided for selecting the desired metalanguage so the searches performed will lead to result displays with instructions in German or Basque, i.e. the displayed data from *EuDeLex* concerning syntax, pragmatics and domain, as well as section titles and pieces of advice given in all parts of the result page.


Metalanguage Switch
Search
Morphology from Wiktionary
EuDeLex Dictionary Entry
Comments Link
EuDeLex Lemma Browser
Links
Wiktionary and Wikipedia links DE (and EU)
EuDeLex
ehk EUSKAL HIZKUNTZA ETA KOMUNIKAZIOA SAILA UPV/EHU | www.ehk.ehu.es
Universidad del País Vasco Euskal Herriko Unibertsitatea
Bilatu / Suche

Suche nach deutschen Grund- und Wortformen
Alemanierazko hitza bilatu
● Metadata deutsch ○ Metadata euskaraz

Wörterbuch Deutsch-Baskisch - Alemaniera-Euskara Hiztegia

EuDeLex Suchergebnis für **"Beispiel"**

Beispiel ☆☆☆☆☆

Substantiv *n.* ▸

1 **adibide**; **etsenplu**; • zum ~ adibidez • ohne ~ paregabe

2= Vorbild **eredu**; **jarraibide**; **ispilu**; • ein ~ geben eredu izan

Artikel kommentieren

Suchergebnisse in anderen Datenbanken:

W de.wiktionary: **Beispiel**: *Substantiv*

W de.wikipedia: **Beispiel**

W de.wikipedia de▸eu Link ▸ **Adibide**

EuDeLex Browser

beipflichten
Beirat
beirren
beisammen
Beisammensein
beisammensitzen
Beischlaf
Beisein
beiseite
beißen
beisetzen
Beisetzung
Beisitzer
Beispiel
beispielhaft
beispiellos
beispielsweise
Beistand
beistehen
beisteuern
Beitrag
beitragen
Beitragserhöhung
Beitragssatz
Beitragszahler
beitreten
Beitritt

Morphologie: Substantiv
Beispiel, *n.*

Nominativ Singular	das Beispiel
Nominativ Plural	die Beispiele
Genitiv Singular	des Beispiels
Genitiv Plural	der Beispiele
Dativ Singular	dem Beispiel
Dativ Plural	den Beispielen
Akkusativ Singular	das Beispiel
Akkusativ Plural	die Beispiele

Quelle: Wiktionary [Autoren] Lizenz: CC BY-SA 3.0

Suche in anderen Wörterbüchern:

DWDS DWDS-Suche: **Beispiel**
Wortschatz Leipzig-Suche: **Beispiel**
ot Openthesaurus-Suche: **Beispiel**
DWB DWB-Suche: **Beispiel**
LEO LEO DE-ES Suche: **Beispiel**
PONS PONS DE-ES Suche: **Beispiel**

Figure 9: *EuDeLex* User Interface with annotations

The search result frame is organized in three columns:

1. *EuDeLex* dictionary entry or entries (in case of homographs[15]) and search results from the German editions of *Wiktionary* and *Wikipedia*, from where the following is retrieved on the fly[16] and, if found, shown to the user: Links to matching pages, and Basque translation links found in those sources.

2. *EuDelex* Lemma browser: Direct neighbours on the *EuDeLex* lemma list are listed as links that lead to the corresponding entry.

3. Morphological information about the German lemma sign, retrieved on the fly from *Wiktionary*, and, if found, reproduced as table. Below, auto-

matically generated hyperlinks to four German monolingual dictionary websites, containing the query string.[17]

The queries to the local database (*EuDeLex*) and to *Wikimedia* servers function for their own, i.e. the script also generates results in case the query has not lead to any entry in *EuDeLex*[18]. As all data from *Wiktionary* and *Wikipedia* is retrieved on the fly, it will always reflect the latest version of these sources. As for today, for more than a half of the planned dictionary lemma list there already exist extensive articles in *Wiktionary*, and in the case of verbs, nouns and adjectives we can extract and display Wiktionary's tables of inflectional and morphological information (cf. Table 2).

If in column (1) just one *EuDeLex* headword is listed, the user obtains the possibility to comment on the article. For this purpose, a dialogue is opened when clicking on the comment link.

Basque Translation Equivalents in the *EuDeLex* dictionary article, provided they exist as Basque lemma[19], lead by clicking on them to the launch of the Basque–German search script with the selected equivalent as query string, that will also show redirects to Basque dictionary platforms.

The software tools and scripts involved in the publishing are:

— *EuDeLex* data is exported by the *TshwaneLex Dictionary Writing System* to XML format.

— A *perl* script transforms this XML into a MySQL table to be installed on the webserver.

— The above-described UI for the search direction DE>EU is based on one PHP script, that performs the queries from *EuDeLex*-MySQL-database and retrieves data from the *Wikimedia* API and displays the results.

— Another PHP script is planned for search direction EU>DE. In the current preliminary version this script offers automatically generated hyperlinks to Basque dictionary websites containing the query string.

3. Conclusions

By the lexicographical factory report on hand we also hope to provide ideas to projects on other medium-density language pairs that do not have adaptable predecessors, requiring that a dictionary be started from scratch. It has been pointed out that by using modern specialised software it is possible to create multifunctional lexicographical databases, the derivatives of which may serve different dictionary functions. In the event that monolingual corpora or even parallel translation corpora are available for the language pair, the data these sources provide can be gainfully taken into account, alongside data from other sources like existing monolingual dictionaries and dictionaries of third languages. We have also shown our attempt to provide an online-platform for the dictionary by means of a simple PHP script, as well as to exploit existing and free-licensed lexical data from *Wikimedia* servers, in order to enrich the dictionary search result displays according to the needs of the user profile we target.

4. Future Work

Having defined a Basque lemma list for EU>DE (Lindemann and San Vicente in prep.), and a microstructure for EU>DE dictionary articles, those will be edited, starting from a draft obtained by reversing the DE>EU part of *EuDeLex*. Furthermore, we will implement a parallel corpus concordance line generator in both DE>EU and EU>DE PHP scripts, so that search result displays will be complemented with translated usage examples. An upgrade of the existing German–Basque parallel corpus is also planned.

5. Notes

1. The research leading to these results has received funding from the European Union's Seventh Framework Programme for research, technological development and demonstration under grant agreement no. 613465, and from project IT665-13, funded by the Basque Government. Funding is gratefully acknowledged.
2. Source: EUSTAT statistics agency, Eusko Jaurlaritza 2011.
3. Source: UPV-EHU University of the Basque Country 2012.
4. For a survey of Basque Lexicography, see Azkarate (1991, 2014); For German–Basque lexicographical work from 1817 until today, Lindemann (2014).
5. At the moment, we lack any profound investigation in dictionary use (*Wörterbuchbenutzungsforschung*) in this context.
6. The online UI is available at http://www.ehu.es/eudelex.
7. See De Schryver (2003) for a commented bibliography on lexicography in the electronic age, and Tarp (2012).
8. An upgrade of this corpus as well as the creation of a German–Basque Bible corpus are in development (for the Bible as parallel corpus, see Resnik et al. (1999)).
9. These domain-specific glossaries could be contrasted with other German glossaries of the same domains, in order to detect gaps and to obtain more "term-wordsenses" or lemma candidates for the main database.
10. As seen in Section 3, we also add morphological information imported from Wiktionary to *EuDeLex* search results.
11. A monolingual and bilingual dictionary entry structure often reflects the polysemy of the headword first and foremost. This implies to include the description of syntactical features of the headword like transitivity in "gram-groups" as child-elements of wordsense-level (e.g. TEI-Standard (Burnard and Sperberg-McQueen 2007)).
12. "Object" here refers only a direct object (accusative), and not to an indirect object (dative), which can be misleading. The term "transitive" seems to us a more appropriate, since transi-

tivity as syntactic property in Basque is a well-known phenomenon: A subject in a transitive clause is set in ergative case, while a subject in an intransitive clause stays absolutive.

13. Once the lexicographer establishes this distinction when editing the entry, a correct mapping to Basque equivalents is also provided, although a Basque-L1 dictionary user might have no need for it. In this sense it is an anticipation that applies to the needs of Basque-L2 users, not in the focus of *EuDeLex* version 1.
14. http://www.ehu.es/eudelex.
15. For result display, case sensitivity is not taken into account. If the search result contains more than one homograph headword, the user is asked in column (3) to choose one by clicking on the headword, in order to get morphological information.
16. *Wikipedia* and *Wiktionary* API are queried for the existence of a homograph page title or redirect page title. No disambiguation of homographs or homonyms is attempted.
17. If those URL do lead to any existing dictionary entry on those platforms is not verified.
18. Consequently, Basque GFL learners might start to use this UI, although *EuDeLex* still does not cover a significant part of the German lemmalist.
19. Until the Basque–German part of the dictionary is published, this *script* verifies if the Basque TE exists on a Basque lemma list (HB, Euskaltzaindia 2008) and, in the positive case, provides a link.

6. Bibliography

Aduriz, I., I. Aldezabal, I. Alegria, X. Artola, N. Ezeiza and R. Urizar. 1996. EUSLEM: A Lemmatiser/Tagger for Basque. Gellerstam, M., J. Järborg, S.-G. Malmgren, K. Norén, L. Rogström and C.R. Papmehl (Eds.). 1996. *Euralex '96 Proceedings I–II, Papers Submitted to the Seventh EURALEX International Congress on Lexicography in Göteborg, Sweden*: 17-26. Gothenburg: Department of Swedish, Göteborg University.

Atkins, B.T.S. 1996. Bilingual Dictionaries: Past, Present and Future. Gellerstam, M., J. Järborg, S.-G. Malmgren, K. Norén, L. Rogström and C.R. Papmehl (Eds.). 1996. *Euralex '96 Proceedings I–II, Papers Submitted to the Seventh EURALEX International Congress on Lexicography in Göteborg, Sweden*: 515-546. Gothenburg: Department of Swedish, Göteborg University.

Azkarate, M. 1991. Basque Lexicography. Hausmann, F.J., O. Reichmann, H.E. Wiegand and L. Zgusta (Eds.). 1991. *Wörterbücher: Ein internationales Handbuch zur Lexikographie*. Vol. 3: 2371-2375. Berlin: De Gruyter.

Azkarate, M. 2014. On-Line Dictionaries of a Minority Language in a Multilingual Society. *XVIth Forum for Iberian studies Cultural and Linguistic Diversity in the Iberian Peninsula*. University of Oxford, Oxford.

Bowker, L. 2010. The Contribution of Corpus Linguistics to the Development of Specialised Dictionaries for Learners. Fuertes-Olivera, P. (Ed.). 2010: 155-168.

Braun, S. 2010. Die Rolle der Muttersprache im Deutschunterricht in zweisprachigen Gebieten am Beispiel des Baskenlandes. Jarillot, C. (Ed.). 2010. *Bestandsaufnahme der Germanistik in Spanien: Kulturtransfer und methodologische Erneuerung*: 25-36. Bern: Peter Lang.

Burnard, L. and C.M. Sperberg-McQueen. 2007. TEI P5: Guidelines for Electronic Text Encoding and Interchange. *TEI Text Encoding Initiative*. Available: http://www.tei-c.org/release/doc/tei-p5-doc/en/html/index.html (Accessed 22 July 2014).

Dentschewa, E. 2006. DaF-Wörterbücher im Vergleich: Ein Plädoyer für 'Strukturformeln'. Dimova, A., V. Jesenšek and P. Petkov (Eds.). 2006. *Zweisprachige Lexikographie und Deutsch als Fremdsprache: drittes Internationales Kolloquium zur Lexikographie und Wörterbuchforschung, Konstantin Preslavski-Universität Schumen, 23.–24. Oktober 2005*: 49-58. Germanistische Linguistik 113–128. Hildesheim/New York: G. Olms.

De Schryver, G.-M. 2003. Lexicographers' Dreams in the Electronic-Dictionary Age. *International Journal of Lexicography* 16(2): 143-199.

De Schryver, G.-M. and D. Joffe. 2005. One Database, Many Dictionaries–Varying Co(n)text with the Dictionary Application TshwaneLex. Ooi, V.B.Y., A. Pakir, I. Talib, L. Tan, P.K.W. Tan and Y.Y. Tan (Eds.). 2005. *Words in Asian Cultural Contexts, Proceedings of the Fourth Asialex Conference, 1–3 June 2005, M Hotel, Singapore*: 54-59. Singapore: Department of English Language and Literature & Asia Research Institute, National University of Singapore.

De Schryver, G.-M., D. Joffe, P. Joffe and S. Hillewaert. 2006. Do Dictionary Users Really Look Up Frequent Words? — On the Overestimation of the Value of Corpus-based Lexicography. *Lexikos* 16: 67-83.

Dickens, A. and R. Salkie. 1996. Comparing Bilingual Dictionaries with a Parallel Corpus. Gellerstam, M., J. Järborg, S.G. Malmgren, K. Norén, L. Rogström and C.R. Papmehl (Eds.). 1996. *Euralex '96 Proceedings I–II, Papers Submitted to the Seventh EURALEX International Congress on Lexicography in Göteborg, Sweden*: 551-559. Gothenburg: Department of Swedish, Göteborg University.

Euskaltzaindia. 2008. *Hiztegi Batua*. Donostia: Elkar.

Fuertes-Olivera, P. (Ed). 2010. *Specialised Dictionaries for Learners*. Lexicographica Series Maior 136. Berlin: De Gruyter.

Hartmann, R.R.K. 2007. The Not so Harmless Drudgery of Finding Translation Equivalents. Hartmann, R.R.K. (Ed.). 2007. *Interlingual Lexicography*: 30-37. Lexicographica Series Maior 133. Berlin: De Gruyter.

IDS. 2009. Korpusbasierte Wortgrundformenliste DEREWO, v-40000g-2009-12-31-0.1, mit Benutzerdokumentation. Institut für Deutsche Sprache, Programmbereich Korpuslinguistik. Available: http://www.ids-mannheim.de/kl/projekte/methoden/derewo.html (Accessed 22 July 2014).

Joffe, D., G.-M. de Schryver and D.J. Prinsloo. 2003. Computational Features of the Dictionary Application 'TshwaneLex'. *Southern African Linguistics and Applied Language Studies* 21(4) [Special issue on 'Human Language Technology in South Africa: Resources and Applications']: 239-250.

Kemmer, K. 2010. *Onlinewörterbücher in Der Wörterbuchkritik. OPAL* 2.

Kempcke, G. 2001. Polysemie oder Homonymie? Zur Praxis der Bedeutungsgliederung in den Wörterbuchartikeln synchronischer einsprachiger Wörterbücher der deutschen Sprache. *Lexicographica* 17: 61-68.

Kilgarriff, A., P. Rychly, P. Smrz and D. Tugwell. 2004. The Sketch Engine. Williams, J. and S. Vessier (Eds.). 2004. *Proceedings of the Eleventh EURALEX International Congress, EURALEX 2004, Lorient, France, July 6–10, 2004*: 105-116. Lorient: Faculté des Lettres et des Sciences Humaines, Université de Bretagne Sud.

Leturia, I. 2012. Evaluating Different Methods for Automatically Collecting Large General Corpora for Basque from the Web. Kay, M. and C. Boitet (Eds). 2012. *Proceedings of the 24th International Conference on Computational Linguistics, 8–15 December 2012, Mumbai, India* (*COLING 2012: Technical Papers)*: 1553-1570. Mumbay, India: The COLING 2012 Organizing Committee.

Lindemann, D. 2013. Bilingual Lexicography and Corpus Methods. The Example of German–Basque as Language Pair. *Procedia — Social and Behavioral Sciences* 95: 249-257.

Lindemann, D. 2014. Zweisprachige Lexikographie des Sprachenpaares Deutsch-Baskisch. Domínguez Vázquez, M.J., F. Mollica and M. Nied (Eds.). 2014. *Zweisprachige Lexikographie zwischen Translation und Didaktik*: 259-281. Lexicographica Series Maior 147. Berlin/Boston: De Gruyter.

Lindemann, D., X. Saralegi, I. San Vicente, I. Manterola and R. Nazar. 2014. Bilingual Dictionary Drafting. The Example of German–Basque, a Medium-Density Language Pair. Abel, A., C. Vettori and N. Ralli. (Eds.). 2014. *Proceedings of the XVI EURALEX International Congress: The User in Focus, EURALEX 2014, Bolzano/Bozen, Italy, July 15–19, 2014*: 563-576. Bolzano/Bozen: Institute for Specialised Communication and Multilingualism.

Lindemann, D. and I. San Vicente. In preparation. *Euskarazko Maiztasun Lemategia Gaurko Teknologien Ikuspuntutik.* Bilbao: UPV-EHU.

Marello, C. 2010. Verbos con construcciones tanto transitivas como intransitivas y/o pronominales en los diccionarios monolingües y bilingües italianos y españoles. Castillo Carballo, M.A. and J.M. García Platero (Eds.). 2010. *La Lexicografía en su dimensión teórica*: 411-434. Malaga: University of Malaga.

Martínez Rubio, E. 2007. *Euskara alemana hiztegia*. Donostia: Elkar.

Meliss, M. 2013. Online-Lexikographie im DaF-Bereich: Eine erste kritische Annäherung; Bestandsaufnahme — Nutzen — Perspektiven. *Revista de estudos alemães* 4: 176-199.

Meyer, C.M. and I. Gurevych. 2010. Worth its Weight in Gold or yet Another Resource — A Comparative Study of Wiktionary, OpenThesaurus and GermaNet. Gelbukh, A. (Ed.). 2010. *Computational Linguistics and Intelligent Text Processing. Proceedings of the 11th International Conference on Intelligent Text and Computational Linguistics, CICLing 2010, Iasi, Romania, March 21–27*: 38-49. Lecture Notes in Computer Science 6008. Berlin/Heidelberg: Springer.

Meyer, C.M. and I. Gurevych. 2012. Wiktionary: A New Rival for Expert-built Lexicons? Exploring the Possibilities of Collaborative Lexicography. Granger, S. and M. Paquot (Eds.). 2012. *Electronic Lexicography*: 259-291. Oxford: Oxford University Press.

Pajzs, J. 2009. On the Possibility of Creating Multifunctional Lexicographical Databases. Bergenholtz, H., S. Nielsen and S. Tarp (Eds.). 2009. *Lexicography at a Crossroads. Dictionaries and Encyclopedias Today, Lexicographical Tools Tomorrow*: 327-354. Bern: Peter Lang.

Pociello, E., E. Agirre and I. Aldezabal. 2011. Methodology and Construction of the Basque WordNet. *Language Resources and Evaluation* 45(2): 121-142.

Resnik, P., M.B. Olsen and M. Diab. 1999. The Bible as a Parallel Corpus: Annotating the 'Book of 2000 Tongues'. *Computers and the Humanities* 33(1-2): 129-153.

Reuter, D. 2010. Interkulturelles Lernen am Beispiel baskischer Muttersprachler im Deutschunterricht. Jarillot, C. (Ed.). 2010. *Bestandsaufnahme der Germanistik in Spanien: Kulturtransfer und methodologische Erneuerung*: 261-266. Bern: Peter Lang.

Reuter, D. and J. Wolff. 2007. *Deutsch–Euskaldunentzat*. Donostia: Erein.

Sanz Villar, Z. 2013. Hacia la creación de un corpus digitalizado, paralelo, trilingüe (alemán–español–uuskera). Sinner, C. and D. Van Raemdonck (Eds.). 2013. *Fraseología contrastiva del alemán y el español. Traducción y lexicografía*: 43-58. Études Linguistiques Linguistische Studien 11. München: Peniope.

Schmid, H. 1995. Improvements in Part-of-Speech Tagging with an Application to German. *Proceedings of the ACL SIGDAT-Workshop, Dublin, Ireland*: 47-50.

Schmid, H. and F. Laws. 2008. Estimation of Conditional Probabilities with Decision Trees and an

Application to Fine-Grained POS Tagging. Scott, D. and H. Uszkoreit (Eds.). 2008. *COLING-08, Proceedings of the 22nd International Conference on Computational Linguistics, 8–22 August 2008, Manchester, UK. Vol. 1*: 777-784. Manchester: COLING.

Tarp, S. 1995. Wörterbuchfunktionen: Utopische und realistische Vorschläge für die bilinguale Lexikographie. Wiegand, H.E. (Ed.). 1995. *Studien zur zweisprachigen Lexikographie mit Deutsch II.* Germanistische Linguistik 127-128: 17-61. Hildesheim/New York: Olms.

Tarp, S. 2012. Online Dictionaries: Today and Tomorrow. *Lexicographica* 28(1): 253-267.

Tarp, S. and H. Bergenholtz. 2005. Wörterbuchfunktionen. Barz, I., H. Bergenholtz and J. Korhonen (Eds.). 2005. *Schreiben, Verstehen, Übersetzen, Lernen. Zu ein- und zweisprachigen Wörterbüchern mit Deutsch*: 11-26. Frankfurt: Peter Lang

Teubert, W. 2002. The Role of Parallel Corpora in Translation and Multilingual Lexicography. Altenberg, B. and S. Granger (Eds.). 2002 *Lexis in Contrast: Corpus-Based Approaches*: 189-214. Amsterdam: John Benjamins.

Wolfer, S., A. Koplenig, P. Meyer and C. Müller-Spitzer. 2014. Dictionary Users Do Look up Frequent and Socially Relevant Words. Two Log File Analyses. Abel, A., C. Vettori and N. Ralli. (Eds.). 2014. *Proceedings of the XVI EURALEX International Congress: The User in Focus, EURALEX 2014, Bolzano/Bozen, Italy, July 15–19, 2014*: 281-290. Bolzano/Bozen: Institute for Specialised Communication and Multilingualism.

Wolff, J. 2010. Sprachvergleich Baskisch/Deutsch und seine Auswirkungen für den Unterricht. Jarillot, C. (Ed.). 2010. *Bestandsaufnahme der Germanistik in Spanien: Kulturtransfer und methodologische Erneuerung*: 37-42. Bern: Peter Lang.

Wolski, W. and A. Cyffka. 2011. *PONS Großwörterbuch Deutsch als Fremdsprache*. Stuttgart: PONS.

Zubillaga, N. 2013. Alemanetik euskaratutako haur- eta gazte-literatura: zuzeneko nahiz zeharkako itzulpenen azterketa corpus baten bidez. PhD Thesis. Vitoria-Gasteiz: UPV-EHU.