

Hiztegegintza
elebiduna:

Euskara-
Alemana

David Lindemann

eman ta zabal zazu



Universidad
del País Vasco

Euskal Herriko
Unibertsitatea



EUSKAL HIZKUNTZA ETA
KOMUNIKAZIOA SAILA

UPV/EHU | www.ehk.ehu.es

Egitasmoak eta funtzioak (1)

- ▼ 1. Egitasmoa: Hiztegi elebidun elektroniko bat

Hiztegiaren funtzioa:

Alemana ikasten diharduten euskaldunentzako
hiztegi elebiduna

- ▼ Balizko erabiltzaileak:

- ▼ EU-L1 DaF-ikasleak
- ▼ Bigarren Hezkuntzan: 1332/ikasturte
- ▼ EOI/Unibertsitatean: 1025/ikasturte
- ▼ DE-EU itzultzaileak

EuDeLex



euskaldunentzat



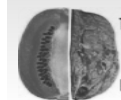
→ Reuter & Wolff 2007

EuDeLex: Ikerketa urratsak

- ✓ DE Lemategia proposatzea
- ✓ DE>EU mikroegitura proposatzea
- ✓ DE lemategiaren %10erako DE>EU artikuluak
- ✓ Erabiltzaile interfazea egituratu eta lehenengo bertsio bat argitaratzea
- ✓ *Bilingual Dictionary Drafting* metodo sorta
 - ✓ Metodoak
 - ✓ Emaizen ebaluazioa (DE A hizkia)
 - x Lortutako datuak datu-basean jaso eta eskuz egiaztatzea
- ✓ EU lemategia proposatzea
 - ✓ Corpusetan oinarritutako maiztasun zerrenda lematizatua
- ✓ EU>DE mikroegitura proposatzea
- x Hiztegi osoa eskuz editatzea

- ✓ Baliabide lexikal eskuragarriak ustiatzea
 - ✓ *Wikipedia* (80.000 DE-EU-links)
 - ✓ *Wiktionary*
 - ✓ *WordNet: GermaNet / PWN / EusWN*
 - ✓ *Elhuyar Pibolex*
- ✓ *Corpus paraleloak*
 - ✓ *DE-EU literatura corpora*
 - ✓ *DE-EU Biblia corpora*
- ✓ *Corpus konparagarriak*
 - ✓ *Wikipedia, "Die Zeit"/"Berria"* (kazetaritza)
 - ▶ Lindemann et al. (2014)

- ✓ EU Maiztasun zerrenda:
 - ✓ Euskarazko bi corpus handietatik erauzitako lemak (ETC, Elh200)
 - ✓ Eskuz landutako sei baliabideetako lemategiak
- ✓ EU mikroegitura-edukiak zirriborratzea:
 - ✓ EDBL, EusWN
 - ▶ Lindemann & San Vicente (2015; in prep.)



Hizkuntzalari konputazionalen elkarlanean



- ▼ I. San Vicente
- ▼ I. Manterola
- ▼ X. Saralegi

*Elhuyar
Fundazioa*



- ▼ R. Nazar

UPF (Barcelona)

U. Valparaíso (Chile)

- ▶ Lindemann & Nazar (2013)
- ▶ Lindemann, Manerola, San Vicente & Saralegi (2014)
- ▶ Lindemann & San Vicente (2015a; 2015b; in prep. 2016)

Egitasmoak eta funtzioak (2)

▼ 2. Egitasmoa: Tesia

Tesiaren funtzioa: „Werkstattbericht“, lan-erreportaia...

- (a) eztabaida metalexikografikorako
- (b) antzeko arazoak lantzen dituzten hiztegegileak (izango) direnentzako ideiak eta argibideak proposatzeko

“Hiztegitza Elebiduna: EU-DE” tesiaren egitura

I. Hurbilketa diakronikoa

- ▼ XIX. mendea
 - ▼ Humboldt 1817
 - ▼ Mahn (ms.)
 - ▼ Charpentier 2002 [1820]
- ▼ XX. mendetik egun arte
 - ▼ Löpeltmann 1968
 - ▼ Kühnel 1999
 - ▼ Euskara-Alemana Hiztegia (Martínez Rubio 2007)

II. Hiztegitza elebiduna gaur egun

- ▼ Oinarrizko zenbait ohar
- ▼ Alemanezko hiztegi elebidun elektronikoak
- ▼ Euskarazko hiztegi elebidun elektronikoak
- ▼ Euskarazko hainbat hiztegiaren konparaketa
- ▼ Hizkuntza ikasleentzako hiztegitza elebiduna
 - ▼ DaF ikasleentzako hiztegien inguruko irizpideak
- ▼ Hiztegien funtzioak eta hiztegien erabilera

III. EuDeLex hiztegiaren egitura

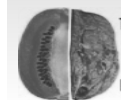
- ▼ Makroegitura: Maiztasunean oinarritutako lematagia
 - ▼ DE: DeReWo-40.000 abiapuntu
 - ▼ EU: “Euskarazko DeReWo bat”
- ▼ Mikroegitura
 - ▼ Hiztegi artikulak: XML
 - ▼ Hiztegi artikulak: argibide motak
 - ▼ DE-EU artikuluen egitura: proposamena
 - ▼ EU-DE artikuluen egitura: proposamena
- ▼ Erabiltzaile-interfazearen egitura

IV. DE-EU itzulpen-ordainen definizioa

- ▼ Glosario elebidunen zirriborrate automatikoa
 - ▼ Metodoak, ebaluazioa, emaitzak
- ▼ Polisemia eta lexikalizazio asimetrikoa
 - ▼ Arazoa, metodo semi-automatikoak
- ▼ Eskuzko lana
- ▼ Adibide-perpaua elebidunak

Eranskinak

- ▼ Mahn 1840 (ms.): Lagin baten transkribaketa
- ▼ Hiztegi laginak (EU)
- ▼ Maiztasun zerrendak (EU)
- ▼ EuDeLex hiztegia
 - ▼ Estatistikak
 - ▼ Eskuz landutako edukiak



Hitzaldi honen egitura

I. Hurbilketa diakronikoa

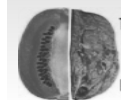
- ▼ XIX. mendea
 - ▼ Humboldt 1817
 - ▼ Mahn (ms.)
 - ▼ Charpentier 2002 [1820]
- ▼ XX. mendetik egun arte
 - ▼ Löpeltmann 1968
 - ▼ Kühnel 1999
 - ▼ Euskara-Alemana Hiztegia (Martínez Rubio 2007)

II. Hiztegi-gintza elebiduna gaur egun

- ▼ Oinarrizko zenbait ohar
- ▼ Alemanezko hiztegi elebidun elektronikoak
- ▼ Euskarazko hiztegi elebidun elektronikoak
- ▼ Euskarazko hainbat hiztegiaren konparaketa
- ▼ Hizkuntza ikasleentzako hiztegi-gintza elebiduna
 - ▼ DaF ikasleentzako hiztegien inguruko irizpideak
- ▼ Hiztegien funtzioak eta hiztegien erabilera

III. EuDeLex hiztegiaren egitura eta edukien zirriborratze (semi-)automatikoa

- ▼ Makroegitura: Maiztasunean oinarritutako lemategia
 - ▼ DE: DeReWo-40.000 (IDS) abiapuntu
 - ▼ EU: "euskarazko DeReWo" bat
- ▼ Mikroegitura
 - ▼ XML
 - ▼ DE-EU hiztegi-artikuluaren egitura
 - ▼ DE-EU hiztegi-artikuluaren egitura
 - ▼ Glosario elebidunen zirriborratze (semi-)automatikoa
 - ▼ Metodoak, ebaluazioa, emaitzak
 - ▼ EU: Entitate sintaktikoen eta adieren zirriborratzea, HAP alorreko baliabideak ustiatuak
 - ▼ Abiatzeko proposamenak
- ▼ Corpus paraleloetatik erauzitako erabilera-adibideak
- ▼ Erabiltzaile-interfazea, 1.0 bertsioa



XIX. Mendea (1): Wilhelm v. Humboldt

Sarea, das Netz.
Sarrasquia, *illotza*, der Leichnam.
Sarri, bald, schnell.
Sar, *sartu*, in etwas hinein gehen, sich in etwas hinein werfen.
Satitu, genießen.
Satorro, *satsuria*, der Maulwurf.
Sei, sechs.
Semea, der Sohn.
Senarra, der Ehemann.
Sendoa, gesund, stark.
Sepa, *seta*, die Hartnäckigkeit, Streitsucht.
Seyala, die Decke.
Sildatu, *zulcatu*, *tincatu*, heften, fest stecken, bestimmen.
Singlea, *ietzia*, *iyetzia*, abgenutzt.
Siñistu, *sinistatu*; *siñetsi*, glauben.
Sisilua, *icellua*, *truesa*, (Labort. *Alkia*), die Bank.
Soca, ein Seil von Binsen.
Soilla, *carsoilla*, kahl. *Soildu*, kahl machen, entkleiden, berauben.
Soina, der Rücken.
Soioa, *seroa*, die Wiese.
Sómatu, vorher sagen.
Soria, erlaubt.
Sorra, fühllos, betäubt.
Sorttea, (Labort.), geboren werden, entstehen. *Sortaldea*, (Seite des Entstehens) Orient. *Sorsaina*, der die Geburt der Menschen lenkende Schutzgeist. S. *Zaina*.
Sost, *sostaguiro*, plötzlich, aus dem Stegreif.
Sua, das Feuer.
Sudurra, *surra*, die Nase.

▼ Lehen argitaratutako hiztegia:

in: W. v. Humboldt 1817

▼ Humboldtten euskalaritza:

▼ 1810-15: "Quellen"

▼ 1812: "Ankündigung"

▼ 1817: "Berichtigungen"

▼ 1821: "Prüfung"

▼ Hurch (ed.) 2010

Itzulpenak ES: Garate 1933

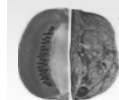
▼ Humboldtten euskalaritzaz:

→ Perlick 1992

→ Gómez 1996

→ Kerejeta 1997

→ Hurch 2002, 2004, 2006



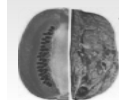
XIX. Mendea (3): Johann v. Charpentier 1820

(45) Verzeichniss einiger in Baigorry üblicher Wörter

Abarastea, sich bereichern.
Abaraxa, abaratcha, reich.
Abaraxtassuna, der Reichthum.
Abendoa, der December.
Abitua, das Kleid.
Aboztua, der August.
Aça, das Kraut (le chou³³).
Açala, die Rinde von einem Baum.
Açança, das Geräusch, der Lärm.
Açaroa, der November.
Açazcala, der Nagel am Finger.
Achala, die Schale von einer Frucht, die Rinde am Brode.
Achalkhentcea, die Rinde wegnehmen, schälen.
Acheria, der Fuchs.
Achola, die Sorge, der Kummer; *ez dut acholeic*, ich habe keine Sorge, es ist mir einerlei.
Achuria, das Lamm.
Acia, der Samen.
Acinda, *Atcinda*, das Vieh, die Herde.
Açuna, trüchtig.
Adarra, das Horn, der Zweig, der Ast.
Adiaraztea, verkündigen.
Adimendia, die Einsicht, Beurtheilungskraft, Erkenntniß.
Adina, das Alter was einer hat; *cer adina baduçu*, wie alt sind Sie.
Adisquidea, *aisquirea*, der Freund.
Adisqui[de]tassuna, die Freundschaft.

Aguinça, das Versprechen, die Verheißung.
Agur, ein unter den Basquen sehr üblicher Gruß.
Ahadia, die Heidelbeere.
Ahal içaitea, können, vermögen.
Ahalguea, die Scham, die Reue.
Ahalgue içaitea, sich schämen.
Aharguella, die Saue.
Aharia, der Widder.
Aharrosi eguitea, gähnen.
Ahatia, die Ente.
Ahaztea, *ahanztea*, vergeßen.
Ahizpa, die Schwester einer Weibsperson.
Ahoa, der Munde.
Ahunia, (eigentlich *ahuntceco* [?] *umia*) die junge Ziege, das Zickl ein.)
(46) *Ahuniamenti* (Ziegenberge) ist der Nahme mit welchem die Basquen die Pÿrenäen bezeichnen.
Ahuntça, die Ziege.
Aicia, *haicia*, der Wind.
Aihena, das Weinlaub.
Ainguiria, der Aal.
Ainguria, der Engel.
Ainharra, die Schwalbe.
Airatcea, in der Luft schweben.
Airia, die Luft.
Airea, das Ansehen, die Mine.

- ▼ Geologo baten hizkuntzalaritza lana
- ▼ Behe-Nafarroako euskalkia
- ▼ Edizioa Hurch & Kerejeta, in: Hurch 2002



XX. mendea (1): F.M.W. Löpelmann 1968

lagun, lagün Gefährte, Begleiter, Kamerad, Gesellschafter, seines Gleichen; Gatte, Gattin (gew. letzteres), Partner, Gehilfe, der Nächste (als Mensch): -arte Gesellschaft; -di Gemeinschaft, Gesellschaft; -du begleiten, geleiten, dienen, helfen, (übertr.) begleiten (musik.), wohlantstehen, zu Hilfe kommen, zusammen einen Weg gehn; -garri geeignet od. fähig zu begleiten od. zu helfen, begleitbar, wem man helfen kann; -go Begleitung, Verzierung, alles was begleitet, vervollständigt, Hilfe, Beistand, Gesellschaft; -kari gesellschaftsliebend, Geselligkeit suchend; -kida Kompanie (Soldaten); -kide Gesellschafter, Partner, Mitarbeiter, der Nächste (relig.); -kin gesellig, sympathisch, kameradschaftlich; -pide Hilfe, Beistand; -tari Beschützer, Kompagnon, Hilfsarbeiter; -tasun Gesellschaft, Vereinigung, Syndikat; -tgu reich an Freunden; -tsa, -tša Beistand, Hilfe, Begleitung, Gesellschaft; -tsabide Hilfe; -tsadun mit Hilfe versehen; -tsagabe hilflos; -tsaile, -tsale, -sale Helfer, Gefährte, Führer; -tse Inf. zu lagundu (s. o.), Hilfe; -skatu sich zusammentun, sich vereinigen (zu gemeinsamer Tat, welche die Kraft eines einzelnen übersteigt), anspannen od. einem Gespann helfen. § Hkft. unsich.; vmtl. arb., zu laġina haften, anhängen, kleben an, vulg. lagina (Afrik.), dazu laġne Ausschuss, Kommission, Komitee, vulg. liġne (Syr.). Zwf. ist Identität mit über. lagun (inschr. l'a'gurn auf dem Bleitafelchen von Ampurias) und Bezieh. zu etr. lacane, denn in beiden Fällen steht die Bed. nicht fest. Fern bleiben ahd. dēgan Diener, mir. laech Held, Laie, lat. laicus Laie.

lagunbat 1. Sonnabend, s. larunbat

lagunbat 2. ein Gefährte, s. lagun, bat

lahapoka Adv. auf vier Pfoten: -tu auf vier Pfoten gehn, sich auf die vier Pfoten niederlassen; s. larrapoka

lahar Dornranke, Dornbusch: -de, -di Dorngestrüpp; -dekatu, -deki, -dekitu die Haut abreißen od. abziehen, abhäuten, schinden, sich die Haut abreißen; -dekasale Gerber; -mahats Stachelbeere; -pe Dornbusch, Rankengewirr; -tegi, -tei Ort mit reichlich viel Dornengerank; -tgu dornig, voller Dornenranken, Sb. Dorngestrüpp; -tu sich mit Dornbüschen bedecken (Land); -s toi = -tegi (s. o.); — lahar r (vor Vok.) Dornranke, Dornbusch, Brombeerstrauch: -aga Dorngestrüpp; -arants Weiss- od. Rotdorn (Crataegus); -esi Dornhecke; -eta Dorngestrüpp; -esko aus Dornenranken gemacht, Dornen-. § Var. zu elhorri. — s. auch laparr, kaparr, naharr.

laharroka(s) Adv. auf vier Pfoten, im Galopp; s. larrapoka

lahatša Krapp (Rubia tinctorum). Nbf. laaš. § wohl kontam. aus *garantsa zu fz. granca, sp. granza Krapp (Et. unsich.) und vlkset. verknüpft mit larrantsa Dorn wegen der hakenförmigen Stacheln des Krapps, bei Ausfall des intervok. r. — s. auch laša 1.

lahiña leichte Schneedecke. Nbf. lahiša. § aus lahi- für *lani- (vgl. bearn. laniu flach) zu sp. llano flach, platt, eben (vgl. die dazu gehörigen dim. Formen: llanito, llanico, llanillo ein wenig flach) + dim. Suff. -ña, bezw. -ša.

lahiša dünne Schneedecke, (übertr.) Schluck (ardolahiša ein Schluck Wein); s. lahiña

lai 1. Hacke, Grabhacke: -an Adv. grabend, mit der Hacke umgrabend; -ari Arbeiter, der mit der Grabhacke arbeitet. § arb., zu laij Biegung, laija Krümmung, Biegung (: lawā drehen, krümmen, biegen). Aus dem Bask.: sp. laya Spaten. Vgl. laia 3.

lai 2. Zweiglein, Sträusschen, Rebe. § arb.; Grdf. *lain zu arb. lāna weich, zart, biegsam sein od. werden, lajjin weich, zart, schmiegsam, biegsam; ebdah. som. lān Zweig, lānto Zweiglein.

▼ Löpelmann 1968 (de Gruyter):

"Ethymologisches WB der baskischen Sprache"

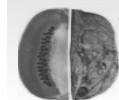
▼ Lhande 1926 (EU-FR)
alemaneratu

▼ Ohar etimologiko ez zientifikoak,
koherentziarik gabeak, iturririk
aipatu gabe

▼ Aho batez kritikatu:

▼ Trask 1997, 1998

▼ Peillen 2008



2.2 XX. Mendea (2-3): Kühnel, Martínez Rubio

hel(du) 1. ankommen, kommen ; **hel niri !** helfen Sie mir !, Hilfe ! 2. reifen

helbide Anschrift *f*, Adresse *f*

helburu Ziel *n*

heldu reif (*auch fig*)

helgaitz unzugänglich

helmuga Ziel *n*

hel(du) ¹ [da] ankommen || eintreffen
▶ etzi helduko dira ▷ sie kommen übermorgen an ▶ bihar helduko zaie erantzuna ▷ die Antwort wird ihnen morgen zugehen || (zoritu) reifen || [zaio] helfen || [dio] (an)packen || fassen ▶ besotik heldu ▷ am Arm fassen

heldu ² iz Erwachsene(r) *m* *f* • izond reif || erwachsen

▼ Kühnel 1999:

"WB des Baskischen"

▼ EU>DE 5.000

DE>EU 8.500

▼ Iturri nagusia "Bi Mila Hiztegia" 90 hamarkadako araugintza baino lehenagokoa

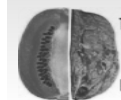
▼ Martínez Rubio 2007:

"Euskara Alemana Hiztegia" (arg. Elkar)

▼ Gaurko euskararen bakarra

▼ Egitura murriztua: 16.000na sarrera, oso artikulu laburrak. Poltsiko neurrikoa.

▼ Funtzio murriztua



DE eta EU hiztezigintzari buruzko zenbait ohar

- ▼ Hiztegi elektroniko elebidunak DE
 - ▼ gehien erabilitako hiztegi elektronikoak ez dira paperezko klasikoen mailara iristen
- ▼ Hiztegi elektroniko elebidunak EU
 - ▼ (Salbuespen bat salbuespen) paperezko hiztegien moldapen elektronikoak dira guztiak
- ▼ Hiztegiak EU
 - ▼ Makroegitura eta mikroegitura ezberdinak
 - ▼ Hiztegien erabilerari buruzko ikerketarik oraindik ez

Bi proposamen

▼ Euskararen Hiztegegintza Corpora

▼ Ereduak:

- ▼ Corpus Lexicográfico do Português
- ▼ Corpus lexicográfico da lingua galega “Dicionario de Dicionarios”

▼ Metodoa:

- ▼ OEH *parsing* lanak
 - ▼ Lematizazioa
 - ▼ 1977arteko hiztegiak
- ▼ Hiztegi berriagoak: *Parsing* lanak

▼ Hiztegien erabilerari buruzko zenbait ikerketa-galdera, EU-DE konbinazioa adibide

1. Ikasten ari diren L2 hizkuntzan, zer maila dute ikasleek?
2. Zenbat (eta zein) L1/L2 hizkuntzatan dagoeneko eskuratu dituzte gaitasunak?
3. Zein hiztegi elebidun eskuragarriak dituzte L1-L2 konbinazioan?
4. (3) puntuko hiztegietatik, zeintzuk ezagutzen dituzte?
5. Zeintzuk dira gaztelania L1 duten ikasleek erabiltzen dituzten hiztegi eta bestelako baliabideak?
6. Zeintzuk dira euskara L1 duten ikasleek erabiltzen dituzten hiztegi eta bestelako baliabideak?
7. Euskara L1 duten ikasleek zubi-hizkuntzako hiztegia erabiltzen badute: Zeintzuk dira arrazoiak?

DaF ikasleentzako hiztegegintza elebiduna

▼ Desiderata marko-egitura:

- ▼ Meta-hizkuntza biak aukeran
- ▼ Erabilera-argibideak bi hizkuntzetan

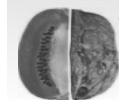
▼ Desiderata mikro-egitura:

- ▼ DE Lema: Morfologia
 - ▼ Flexio-morfologiari buruzko argibideak
- ▼ DE-Lema: Sintaxia
 - ▼ Aditz laguntzailearen hautapena
 - ▼ Balentzia/argumentu egituraren gauzatzeari buruzko argibideak
- ▼ DE-Lemma: Semantika
 - ▼ Polisemia desanbiguatzeko argibideak
- ▼ Adibide-perpausak
 - ▼ asmatutakorik ez
 - ▼ Corpus paraleloetatik erauziak (eskuragarriak diren heinan)

▼ EuDeLex egitasmoan gauza daitezkeenak:

- ▼ Zein argibide motari preferentzia?
 - ▼ 1. argitalpena:
 - ▼ Polisemia desanbiguatzeko ezinbesteko dirudienak
 - ▼ 2. argitalpena:
 - ▼ Argibide mota guztiak artikulu guztietan
- ▼ *Wiktionary*, *Wikipedia* iturrietatik datuak inportatzea
- ▼ Bestelako hiztegi-atarietara lasterbideak
- ▼ Corpus-konkordantzia paraleloak. Tresna egokiak eduki arte: *Glosbe*

- Meliss 2013
- Kemmer 2010
- Tarp 1995, 2008



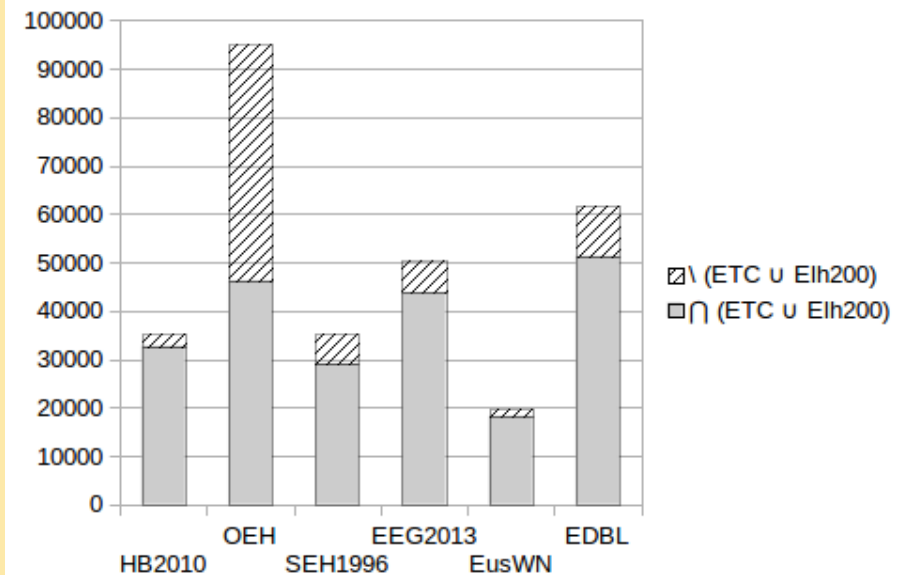
EuDeLex hiztegiaren edukia

▼ Makroegitura:

Maiztasun zerrenda lematizatuak abiapuntu

- ▼ DE: DeReWo-40.000 (IDS 2009)
 - ▼ Konparaketa DaF B1 hiztegiarekin
 - ▼ A hizkia (/^[AaÄä].*/):
 - ▼ 3604 artikulu
 - ▼ 3285 lema-ikur DeReWo-tik, aldeketarik gabe (% 91,15)
 - ▼ 37 lema-ikur DeReWo-tik, egokiturik (% 1,03)
 - ▼ 11 homonimo (% 0,31)
 - ▼ 291 (% 8,07): gehituak
 - ▼ beste hiztegietatik
 - ▼ arrazoi formalengatik

- ▼ EU: Euskarazko “DeReWo bat”
 - ▼ corpusetako lema (ETC, Elh200)
 - ▼ Konparaketa 6 iturriekin



Hiztegia editatzean erabilitako softwarea

Dictionary Writing System:

TshwaneLex (TLEX)

(de Schryver & Joffe 2005, 2010)

DTD, estiloak libreki moldagarri

XML, html, rtf, tsldict

Attribute Lists, Label Sets

Import, Reverse, Compare, Merge

CrossRefs, Hyperlinks

Corpus Tagging eta KWIC:

SketchEngine (SkE)

(Kilgarriff et al. 2004)

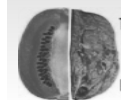
DE eta EU corpus handiak bertan eskuragarri

Wordsketch, GDEX

DE-EU literatura corpora: KWIC paraleloak

The screenshot shows the TshwaneLex (TLEX) software interface. The main window displays a dictionary entry for the lemma 'absetzen'. The entry is structured into sections (Section I and Section II) and senses (Sense 1 through Sense 7). Each sense includes grammatical information (GramInfo) and morphological information (Morph). The interface also shows a list of attributes (F1, F2, F3, F4, F5) and a search bar. The right-hand pane shows the output of the software, displaying the entry in a structured format with various grammatical and morphological information.

▼ Sonstiges: regular expressions, linux command line, perl

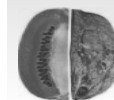


EuDeLex XML: DE aditza

```
<Lemma id="19375" LemmaSign="befinden" Source="Derewo" DerewoRang="569" Frequency="9"
  FlexMorfTag="unregelmäßig, ohne ge">
  <Section id="211053" SectionNumber="1" POS="Verb Transitiv" GramInfo="+haben">
    <Sense id="211054" SenseNumber="1" SyntaxHinweis="jdn./etw. für +adj.">
      <TE id="211055" TE="-tztat eman"/>
      <TE id="289097" TE="iritzi"/>
      <MWE id="289101" MWE="etw. für gut ~" Translation="ontztat eman; -i oniritzi"/>
    </Sense>
  </Section>
  <Section id="289102" SectionNumber="2" POS="Verb Intransitiv" GramInfo="+haben">
    <Sense id="289103" SenseNumber="1">
      <TE id="289104" TE="erabaki"/>
      <TE id="289105" TE="erabakia hartu"/>
    </Sense>
  </Section>
  <Section id="289098" SectionNumber="3" POS="Verb +&apos;sich&apos;" GramInfo="+haben">
    <Sense id="289099" SenseNumber="1">
      <TE id="289106" TE="izan"/>
      <TE id="289100" TE="egon"/>
      <TE id="289107" TE="aurkitu"/>
    </Sense>
  </Section>
</Lemma>
```

befinden Derewo 569 - *unregelmäßig, ohne ge.*

- I *Verb Transitiv* +haben ▶ 1 [jdn./etw. für +adj.] -tztat eman; iritzi
 - etw. für gut befinden ontztat eman; -i oniritzi
- II *Verb Intransitiv* +haben ▶ 1 erabaki; erabakia hartu
- III *Verb + 'sich'* +haben ▶ 1 izan; egon; aurkitu



DE izena

Ausdruck¹ Derewo 1918 ~s, Ausdrücke

Substantiv m. s. Verbinf. ausdrücken ▶ 1 adierazpen 2 termino; esamolde; esapide
3 ezaugarri; adierazgarri 4 adierazkortasun; adierazgarritasun

Ausdruck² Derewo 1918 ~s, Ausdrücke

Substantiv m. s. Verbinf. ausdrücken ▶ 1 inprimatze; inprimaketa

DE Adjektiboa/Adberbioa

DWDS-Wörterbuch (2013)

gut – Adjektiv
besser, am besten

Adjektive / Adverbien

▶ good *adj.* - better, best ◻ ▶ gut - besser, am besten

▶ well *adj. adv.* ◻ ▶ gut

Adjektive / Adverbien

▶ bien *adv.* ◻ ▶ gut

▶ bueno, buena *adj.* ◻ ▶ gut

alleinig *adj.* bakar; soil

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

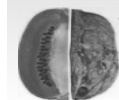
I *Adjektiv* ▶ 1 latz; gaitz 2 *veraltend* gaizto; maltzur

II *Adjektiv adverbial* ▶ 1 *ugs. südd.* oso

anders Derewo 596

I *Adverb* ▶ 1 ezberdin; bestela; beste era batean

II *Adverb präd./att.* ▶ 1 ezberdin; bestelako; beste era bateko



EU-DE artikuluen egitura: proposamena (aditza)

heldu

I *aditz perifrastikoa* ▶

A *ad. intrans. +da* ▶

- 1 = *iritsi* ankommen; kommen; herkommen
- 2 kommen; geschehen
- 3 reifen

B *ad. trans. +dio* ▶

- 1 fassen; greifen; halten; festhalten
- 2 ergreifen; anfangen
- 3 (*Tier*) beißen; stechen

C *ad. trans. +du* ▶

- 1 *ipar. zah.* = *bidali, helarazi* schicken

D *ad. intrans. +zaio* ▶

- 1 *ipar.* helfen; zur Hilfe kommen

II *izlag.* ▶

- 1 reif
- 2 *irud.* erwachsen
- 3 *irud.* vernünftig

III *iz.* ▶

- 1 (*eines Tiers*) Biss; Stich

heldu

I *ad.* ▶

- 1 *+da* = *iritsi* ankommen; kommen; herkommen
- 2 *+da* kommen; geschehen
- 3 *+da* reifen
- 4 *+dio* fassen; greifen; halten; festhalten
- 5 *+dio* ergreifen; anfangen
- 6 *+dio (Tier)* beißen; stechen
- 7 *+du ipar. zah.* = *bidali, helarazi* schicken
- 8 *+zaio ipar.* helfen; zur Hilfe kommen

II *izlag.* ▶

- 1 reif
- 2 *irud.* erwachsen
- 3 *irud.* vernünftig

III *iz.* ▶

- 1 (*eines Tiers*) Biss; Stich

„*adiera sintaktikoak*“ lehen mailan ezberdintzeak
adiera semantiko egokia zehazten laguntzen dio
(1) *hiztegi-erabiltzaileari*; (2) *tresna automatiko bati*

Itzulpen ordainen bikoteak zirriboratzeko metodoak (*Bilingual Dictionary Drafting*)

▼ Corpusetan oinarriturik

▼ 1. DE-EU corpus paraleloak: GIZA++ (Elhuyar)

▼ 2. DE-EU corpus paraleloak: Bifid (R. Nazar)

DE-EU literatura corpora

DE-EU Biblia corpora

▼ 3. Elhuyar Pibot Hiztegiak

EU-EN+EN-DE hiztegi-pibotajea

Kazetaritza/Wikipedia corpus konparagarria

▼ Bestelako baliabideetatik erauzteak

▼ 4. de.wikipedia: EU interlingual links

▼ 5. de.wiktionary: EU itzulpen ordainak

▼ 6. GermaNet / EusWN (PWN pibot gisa)

DE-EU literatura-corpus paraleloa:

- Sanz, Uribarri & Zubillaga, 2012
- 81 alemanezko liburu, itzulpen ofiziala alboan
- 2 milioina inguru testu-hitz

DE-EU Biblia-corpora:

- Saralegi 2014
- 640.000 testu-hitz

Zerrenda elebidunen ebaluazioa

▼ Neurketak kalkulu bitartez

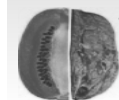
- ▼ Kuantitatiboa: *Recall* (erantzun-kopurua) Goldstandard Lemategiari begira (DeReWo bzw. EuDeLex A)
- ▼ Kuantitatiboa: Zenbat itzulpen-ordain bikote? (zerrendaren tamaina)
- ▼ Kuantitatiboa: Zenbat itzulpen-ordain bikote DE lema batean?

▼ Eskuzko neurketak

- ▼ Kualitatiboa: Ordainen bikoteen egokitasuna (hurbiltasun semantikoa)
- ▼ Kualitatiboa: Ordainen POS egokitasuna (kategoria gramatikala)

▼ (Oraindik) neurtu gabe

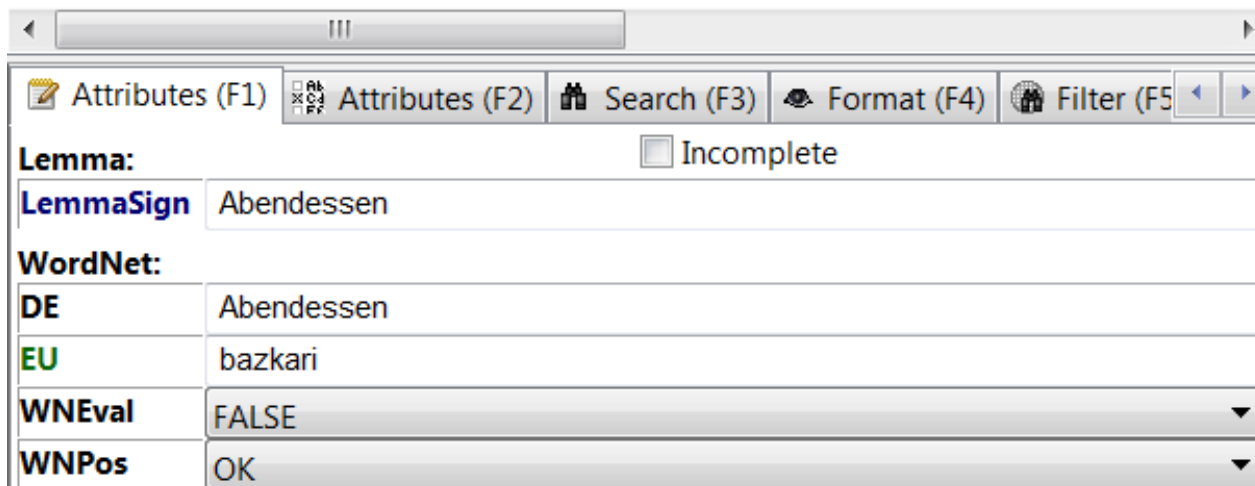
- ▼ DE lemaren homonimia eta polisemiaren estaldura



Ebaluazioa TLex aplikazioan

▼ Zerrenda elebidunak EuDeLex-eko datuetan txertatu

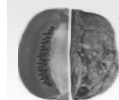
```
[-] Lemma: Abendessen LemmaSign=Abendessen, Source=DeReWo, DerewoRang=7129,
[-] SE: SectionNumber=1, POS=Substantiv, GramInfo=+n.
  [-] Sense: 1 SenseNumber=1
    [-] TE (Translation Equivalent): TE=afari
    MWE: MWE=zu Abend essen, Translation=afaldu
  LitGizaPos: DE=Abendessen_NN, EU=afari_IZE, score=0.3809524, posmap=582, LitC
  BibleGizaPos: DE=Abendessen_NN, EU=afari_IZE, score=0.2000000, posmap=582, E
  Pibolex: DE=Abendessen, EU=afari, type=D, score=0.131871003151754, PiboEval=O
  PiboWiki: DE=Abendessen, EU=afari, type=D, score=0.0399536951758275
  Wikipedia: DE=Abendessen, Redir=Abendessen, EU=Afari, PageID=241696, WikiEval
  Wiktionary: DE=Abendessen, Redir=Abendessen, POS=Substantiv, PageID=15040, E
  WordNet: DE=Abendessen, EU=afari, WNEval=OK, WNPos=OK
  WordNet: DE=Abendessen, EU=bazkari, WNEval=FALSE, WNPos=OK
```



The screenshot shows the TLex application interface. At the top, there is a menu bar with options: Attributes (F1), Attributes (F2), Search (F3), Format (F4), and Filter (F5). Below the menu bar, there is a section for the lemma 'Abendessen'. The 'LemmaSign' field contains 'Abendessen'. Under the 'WordNet' section, there are several fields: 'DE' with the value 'Abendessen', 'EU' with the value 'bazkari', 'WNEval' with the value 'FALSE', and 'WNPos' with the value 'OK'. There is also an 'Incomplete' checkbox which is currently unchecked.

▼ Aurretik definituriko DE Lemategian agertzen dena baino...

- ▼ ...ez dugu ebaluatu
- ▼ ...ez dugu EuDeLex hiztegian jasoko



Ebaluazioaren emaitza orokorra

6 metodoek bateraturik,
[EuDeLex A]
erreferentzia lemategiaren
% 80a
estaltzen dute zerrenda elebidunek!

Corpusetan oinarritutako metodoa: GIZA++

▼ Corpus paralelo elebidunak: *Word Alignment*, **GIZA++**

▼ Emaizak:

▼ %9-35 Recall

▼ %25-45 bikote egoki

▼ DE glosario sarreren %47-63:
ordain egoki 1 edo gehiago

▼ “Tweaks”

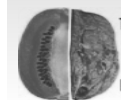
▼ DE izenkien hasierako hizki
LARRIA

▼ Lema_POS hizki-kate

▼ POS-mapping

▼ EU Stoplist (150 maizenak)

abend	<i>gauean</i>	aber	<i>ordea</i>
abend	<i>gau</i>	aber	<i>baina</i>
abend	<i>arrats</i>	abgesehen	<i>utzita</i>
abend	<i>arratsean</i>	abgesehen	<i>aparte</i>
abend	<i>iluntzean</i>	abgrund	<i>amildegira</i>
abend	<i>arratsaldean</i>	abgrund	<i>amildegiaren</i>
abend	<i>gauerako</i>	abgrund	<i>amildegia</i>
abend	<i>arratsalde</i>	abgrund	<i>amildegi</i>
abend	<i>gauetz</i>	abholen	<i>bila</i>
abendessen	<i>afaria</i>	abrupt	<i>-batean</i>
abendessen	<i>afaltzeko</i>	abschied	<i>agur</i>
abends	<i>gauean</i>	abschied	<i>adio</i>
abends	<i>iluntzean</i>	absicht	<i>asmoa</i>
abends	<i>gauetz</i>	absicht	<i>nahita</i>
abends	<i>arratsean</i>	absicht	<i>asmo</i>
abends	<i>gauetan</i>	absicht	<i>asmorik</i>
abends	<i>arratsalde</i>	absichten	<i>asmo</i>
abenteuer	<i>abentura</i>	absichtlich	<i>nahita</i>
abenteuer	<i>abenturak</i>	absichtlich	<i>propio</i>
aber	<i>baina</i>	abstand	<i>tartea</i>



Corpusetan oinarritutako metodoa: Bifid

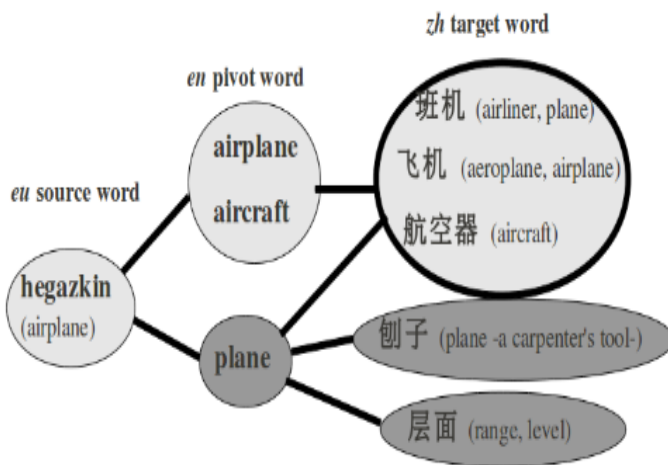
- ▼ Corpus paraleloak: *Word Alignment*, **Bifid** (Nazar 2012)
 - ▼ Emaidzak:
 - ▼ %1-5 Recall
 - ▼ %71-94 ordain egoki
 - ▼ DE glosario sarreren %79-94: ordain egoki 1 edo gehiago

wills gesicht	<i>willen aurpegia</i>
bildhauer	<i>eskultorea</i>
einem tiefen	<i>hasperen sakon</i>
individuum	<i>gizabanako</i>
eines hundes	<i>txakur baten</i>
mensch ärgere	<i>partxisean</i>
rechts links	<i>eskuin ezker</i>
schranktür	<i>armairuko</i>
sagte irene	<i>irenek</i>
wolfgang	<i>amadeo</i>
fünfunddreißig	<i>hogeita hamabost</i>
fußgänger	<i>oinezko</i>
tante	<i>izeba</i>

...*Word Alignment*: “Zarata” eta “Isiltasunaren” erdibide egokienaren bila...
Corpusaren tamaina aldagai ezin garrantzitsuagoa...

Elhuyar Pibot Hiztegiak

- ▼ Pibot hiztegia, corpus konparagarrietako datuetan frogatua
 - ▼ Abian ari den ikerketa Elhuyarren (Saralegi et al., 2012): “Pivot-based bilingual dictionary building”
 - ▼ Itzulpen-ordain izateko hautesleak EU-EN eta EN-DE hiztegi bana kateatzetik
 - ▼ *Inverse Consultation* (IC) eta *Distributional Similarity* (DS) corpus konparagarri elebidunetan
 - ▼ Polisemiaren aurrean, arazoak (Anbigutasunak lema-ikurra baztertzea dakar)
 - ▼ Orain arte euskara eta txinera, alemana, hindi, arabiera eta swahiliren artean
 - ▼ DE-EU: egungo bertsioan, 13.000 DE lema-ikur, 1-25na EU ordainekin
 - ▼ Ebaluazioa: % 15-16 Recall, % 62-64 bikote egokiak, % 79-80 DE zerrenda-sarrera ordain egoki 1+ batekin



de.wiktionary: itzulpen-ordainen erauzketa

- ▼ 4927 bikote (A-Z)
- ▼ 3120 DE lema-ikur (A-Z)
- ▼ 3025 EU ordain ezberdin (A-Z)
- ▼ Emaizak (A):
 - ▼ % 4 Recall
 - ▼ % 97 egokitasuna
 - ▼ DE glosario-sarrerren %100: ordain egoki 1 edo gehiago

Abend	<i>arrats</i>
Abendessen	<i>afari</i>
Abendessen	<i>afarite</i>
Abendessen	<i>afari</i>
aber	<i>baina</i>
Aberrationskonstante	<i>aberrazio iraunkor</i>
abfahren	<i>abiatu</i>
abfahren	<i>irten</i>
abfragen	<i>galdetu</i>
abfragen	<i>kontsultatu</i>
abgeben	<i>entregatu</i>
abgeben	<i>eman</i>
abhandenkommen	<i>galdu</i>
Abhängigkeit	<i>dependentzi</i>
abisolieren	<i>kendu</i>
abisolieren	<i>desisolatu</i>
ablehnen	<i>ezeztatu</i>
ableiten	<i>deribatu</i>
ableiten	<i>atara</i>

Wikipedia: DE-EU lotura interlingualen erauzketa

Emaitzak:

- DE-EU: 80.000 lotura
- % 7,35 Recall
- % 89 egokitasuna
- DE glosario sarreren % 89: ordain egoki 1 edo gehiago

The screenshot shows the German Wikipedia article for 'Arm'. The article title is 'Arm' and the content discusses the human arm as an extremity. The text mentions that the arm is also referred to as the 'freie obere Extremität' and is involved in various movements like 'Greifwerkzeug' and 'Pendelbewegungen'. There are three anatomical illustrations of the arm and shoulder. The article is in German, but a Baskic translation is overlaid on the page. The Baskic translation is circled in red and includes the text 'Beso - Baskisch' and 'Bewegungen im Schultergelenk'. The URL 'eu.wikipedia.org/wiki/Beso' is visible at the bottom of the page.

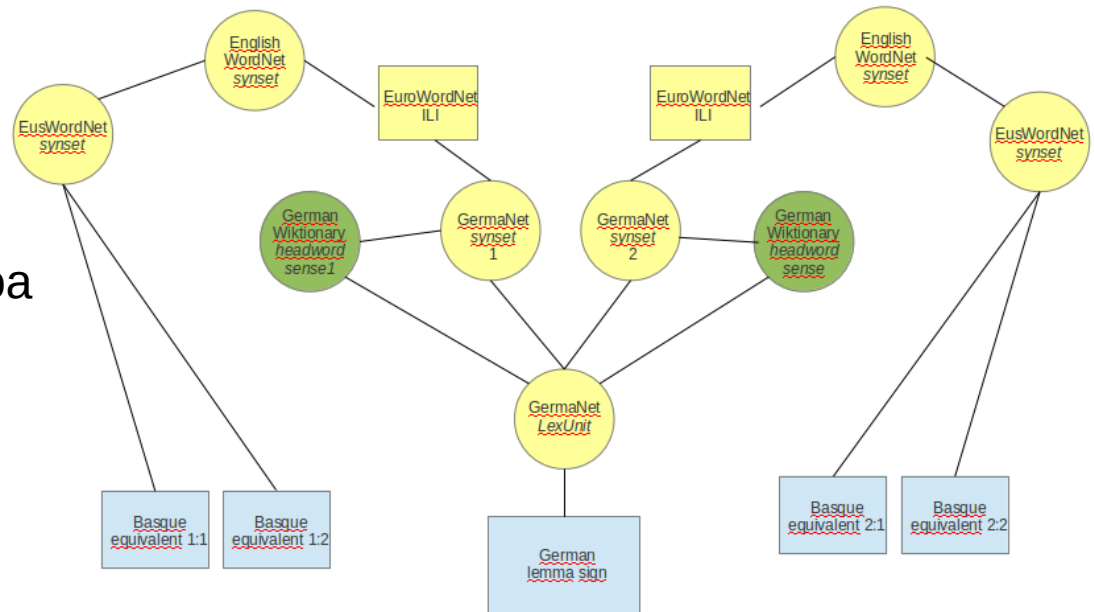
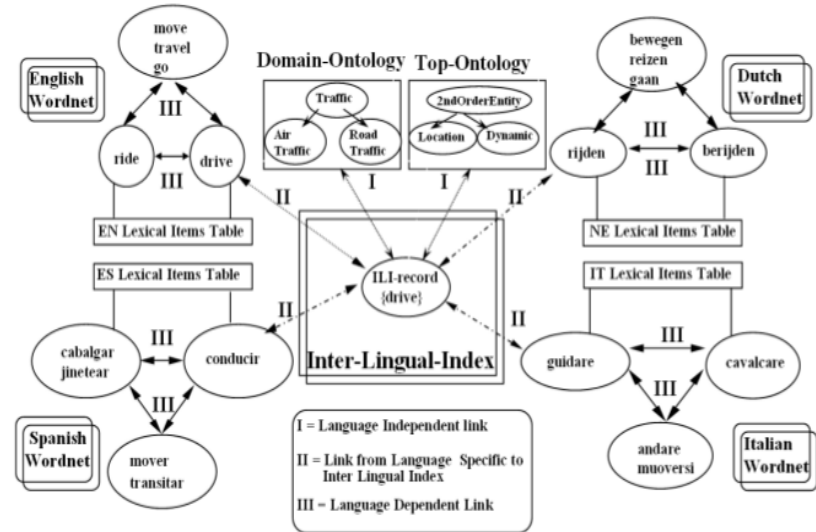
EuroWordNet bidezko erauzketa (1)

Geralekuak WN-ak lotzeko bidean:

- DE lema-ikurra
- GermaNet Lexical Unit
- GermaNet synset(s)
- EuroWordNet ILI record(s)
- Princeton WordNet synset(s)
- EusWN synset(s)
- EU lema-ikurra

WordNet metodoaren abantailak:

- Iturrien tamaina
- Printzipioz, adieren desanbiguazioa ere erauz daiteke



EuroWordNet bidezko erauzketa (2)

- ▼ EN-DE/EU synset lerrotuak: 8.481
- ▼ Gure emaitzak:
 - ▼ % 21,5 Recall
 - ▼ % 83 bikote egoki
 - ▼ DE glosario-sarrereren % 90: ordain egoki 1 edo gehiago

```
<EngSynset id="5945" Synset="eng-30-00102974-v" Source="pwn-30-mcr">
```

```
<PWN30 id="5946" EngLexUnit="drivel_1"/>
```

```
<PWN30 id="5947" EngLexUnit="slabber_1"/>
```

```
<PWN30 id="5948" EngLexUnit="slaver_1"/>
```

```
<PWN30 id="5949" EngLexUnit="slobber_1"/>
```

```
<PWN30 id="5950" EngLexUnit="drool_2"/>
```

```
<PWN30 id="5951" EngLexUnit="dribble_4"/>
```

```
<EusWN30 id="326991" EusLexUnit="adurra_jariatu_1"/>
```

```
<EusWN30 id="359135" EusLexUnit="lerdea_jariatu_1"/>
```

```
<GermaNet id="376901" lexUnitId="176894"/>
```

```
<GermaNet id="376902" lexUnitId="176896"/>
```

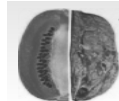
```
<GermaNet id="376903" lexUnitId="176835"/>
```

```
<GermaNet id="376904" lexUnitId="176893">
```

```
</EngSynset>
```

Ebaluazioaren emaitza guztiak

	LitGiza Stop	LitGiza StopPos	BibleGiza Stop	BibleGiza StopPos	LitBifid	BibleBifid	Pibolex News	Pibolex Wiki	Wikipedia	Wiktionary	WordNet	ANY DRAFT
DE Letter A Lemma with EU TE	5,361	3,826	1,349	818	265	230	682	730	5,457	396	1,076	
Intersection with EuDeLex GS	1,276	696	434	323	177	49	559	567	265	147	774	2,928
Recall on 3406 GS lemmata	35,41%	19,31%	12,04%	8,96%	4,91%	1,36%	15,51%	15,73%	7,35%	4,08%	21,48%	81,24%
Ø TE per Lemma	2,27	1,73	2,35	1,46	1,07	1,22	2,64	2,82	1,00	1,33	2,57	2,28
Evaluated Lemma	1,276	696	434	323	177	49	559	567	265	147	774	4,248
Evaluated TE	2,901	1,206	1,019	470	189	60	1,476	1,601	265	195	1,988	9,694
• OK	746	537	256	215	178	43	939	1,007	236	189	1,654	5,248
• FALSE	1,923	519	728	230	3	15	414	513	6	3	188	3,793
• PART	186	122	11	6	6	1	14	7	4		17	246
• NEAR	46	28	24	19	2	1	109	74	19	3	129	407
Evaluated POS	2,901	1,206	1,019	470	189	60	1,476	1,601	265	195	1,988	9,694
• OK	1,071	965	448	422	186	52	1,321	1,286	245	195	1,983	6,787
• FALSE	1,830	241	571	48	3	8	310	315	21		5	3,063
Lemma with 1+ TE OK	609	439	213	189	168	39	443	454	236	147	700	2,081
Lemma with all TE OK	315	294	111	149	168	35	316	322	236	142	601	1,123
Lemma with something usable	444	232	129	62	7	6	165	164	23	5	143	1,056
Lemma with all TE FALSE	517	170	194	112	2	8	78	82	6		30	371
TE OK	25,72%	44,53%	25,12%	45,74%	94,18%	71,67%	63,62%	62,90%	89,06%	96,92%	83,20%	54,14%
POS OK	36,92%	80,02%	43,96%	89,79%	98,41%	86,67%	89,50%	80,32%	92,45%	100,00%	99,75%	70,01%
Lemma with 1+ TE OK	47,73%	63,07%	49,08%	58,51%	94,92%	79,59%	79,25%	80,07%	89,06%	100,00%	90,44%	48,99%
Lemma with all TE OK	24,69%	42,24%	25,58%	46,13%	94,92%	71,43%	56,53%	56,79%	89,06%	96,60%	77,65%	26,44%
Lemma with something usable	34,80%	33,33%	29,72%	19,20%	3,95%	12,24%	29,52%	28,92%	8,68%	3,40%	18,48%	24,86%
Lemma with all TE FALSE	40,52%	24,43%	44,70%	34,67%	1,13%	16,33%	13,95%	14,46%	2,26%	0,00%	3,88%	8,73%



Ebaluazioaren emaitzak eta ondorioak

▼ Metodo guztiak baturik

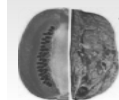
- ▼ % 80 Recall DE erreferentzia lemategiari begira (EuDeLex A)
- ▼ % 45 Egokitasuna (1): Itzulpen bikote egokiak zerrenda elebidunean
- ▼ % 50 Egokitasuna (2): Ordain egoki 1 edo gehiago duten zerrenda-sarrera

▼ Zerrenda elebidunak taldekaturik, egokitasunaren arabera

- ▼ (1) Wiktionary, Wikipedia, Bifid (Literatura): Ia erabateko egokitasuna
- ▼ (2) Bifid (Bibelkorpus), WordNet: Egokitasun-maila altua
- ▼ (3) Giza, Elhuyar Pivot: Egokitasun-maila baxua

▼ Etorkizun hurbileko urratsak

- ▼ (1) [eta (2)] taldeko datuak EuDeLex datu-basean jasotzea
 - ▼ alemanerazko lemategiaren heren bat [erdi bat] euskarazko ordainekin estalia
- ▼ DE-EU corpus paraleloen garapena
 - ▼ Testu gehiago jasotzea
 - ▼ Esaldi mailako lerrotzea hobetzea (automatikoki, *Hunalign* tresnaz)
- ▼ *Word Alignment* tresnen konfigurazioa hobetzea (“little tweaks”)

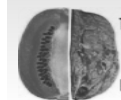


Euskarazko hiztegiaren zirriborratzea:

(1) Homografoaren (lema-ikurraren) maila

```
<homograph homograph="aditu" corpus_counts="42042" />
```

- ▼ Elh200 corpusa, EusTagger (EDBL) tresnaz lematizaturik
 - ▼ 93.081 lema-ikur (20+ agerpen), kategoria gramatikal guztiak
 - ▼ Maiztasun datuak alboan
 - ▼ Horietarik **47.450** agertzen dira 6 erreferentzia-lemategiren batean
 - ▼ Beste erdia: “zarata” guztiaren artean, neologismoak



(1) Homograph (lema-ikurra), (2) Entitate Sintaktikoa

```
<homograph homograph="aditu" corpus_counts="42042">
  <ADI lemma="aditu" pos="ADI_SIN" corpus_counts="18989"/>
  <IZE lemma="aditu" pos="IZE_ARR" corpus_counts="13945"/>
  <ADJ lemma="aditu" pos="ADJ_ARR" corpus_counts="5486"/>
</homograph>
```

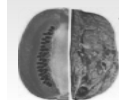
- ▼ Elh200 corpora, EusTagger (EDBL) tresnaz lematizaturik
 - ▼ 93.081 Homografotan:
 - ▼ 103.136 entitate sintaktiko (IZE, ADI, ADJ, ADB), bakoitzak 20+ agerpen
 - ▼ Entitate bakoitzearen inguruko maiztasun datuak
 - ▼ 71.048 IZE (**27.723**ren lema-ikurra agertzen da erreferentzia-lemategietan)
 - ▼ 6.723 ADI (**4.970**ren lema-ikurra agertzen da erreferentzia-lemategietan)
 - ▼ 23.509 ADJ (**7.550**ren lema-ikurra agertzen da erreferentzia-lemategietan)
 - ▼ 1.856 ADB (**1.586**ren lema-ikurra agertzen da erreferentzia-lemategietan)
 - ▼ **Guztira, 41.829** lau kategoria nagusietako corpuseko entitate erreferentzia-lemategietan
 - ▼ Hortaz: Lema-ikur homografo asko euskaraz (% 10 inguru, DE % 0,5)

(1) Homograph, (2) Entitate sintaktikoa, (3) Adiera

```
<homograph homograph="aditu" corpus_counts="42042">
  <ADI lemma="aditu" pos="ADI_SIN" corpus_counts="18989">
    <sense synset="30-00588888-v" equivs="understand"/>
    <sense synset="30-02169702-v" equivs="hear"/>
    <sense synset="30-02571901-v" equivs="heed mind listen"/>
  </ADI>
  <IZE lemma="aditu" pos="IZE_ARR" corpus_counts="13945">
    <sense synset="30-09617867-n" equivs="expert"/>
    <sense synset="30-10557854-n" equivs="scholar scholarly_person bookman"/>
  </IZE>
  <ADJ lemma="aditu" pos="ADJ_ARR" corpus_counts="5486">
    <sense synset="30-02226162-a" equivs="adept expert skillful"/>
  </ADJ>
</homograph>
```

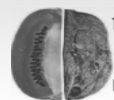
▼ EusWN (Basque WordNet):

- ▼ 49.896 adiera (word senses), 27.118 entitate sintaktikotan, 26.966 lema-ikur banaturik
- ▼ EusWN entitate sintaktikoetatik, zeintzuk agertzen dira corpusean 20+ aldiz?
 - ▼ Laster jakingu dugu entitate sintaktikoaren mailan (Lema-ikur mailan: 26.966tik 14.709)
- ▼ Corpusetako entitate sintaktikoak: EusWN aberasteko hautagaiak
 - ▼ Beste erreferentzia-iturriren batean agertzen direnekin hasiko gara (Lema-ikur mailan: 20.660)



Zirriborroan falta diren adierez jabetzea: Eskuzko lana

EusWN Lexical Unit	Definition EN	EusWN 3.0 synset	EN synset	CAT synset
adar_1	<i>one of the bony outgrowths on the heads of certain ungulates</i>	adar_1	horn_2	banya_1
adar_2	<i>a railway line connected to a trunk line</i>	adar_2	branch_line_1 spur_track_1 spur_5	enforcall_1 forcall_1
adar_3	<i>a warning signal that is a loud wailing sound</i>	adar_3, sirena_2 turuta_5	siren_3	
adar_4	<i>a local branch of some fraternity or association</i>	adar_4	chapter_3	capítol_2
adar_5	<i>a division of a stem, or secondary stem arising from the main stem of a plant</i>	adar_5 abar_2 besanga_1 beso_12	branch_2	branca_1 branc_1
adar_6	<i>an alarm device that makes a loud warning sound</i>	sirena_4 adar_6 turuta_6	horn_9	
adar_7	<i>a device used for easing the foot into a shoe</i>	zapata_sartzeko_1	shoehorn_1	calçador_1



„Begriff“	
kontzeptu	18
ideia	11
hitza	2
adigai	1
burutapena	1
ezagutza	1
gai	1
hitz	1
ikusmolde	1
pentsakera	1
termino	1
ulerbide	1
ulerkera	1

„im Begriff sein“			
-tzera joan	5	-tzekotan izan	1
-tzeko zorian	4	-tzera apoderatzen izan	1
-tzear egon	3	ekarri	1
-tzeko asmoa izan	3	gutxi falta + subjkt.	1
-tzeko asmotan egon	2	hasia izan	1
-tzea pentsatu	1	inf. nahi izan	1
-tzear izan	1	inf. nahian	1
-tzeko duda egin	1	egin behar izan	1

LU und TE Kandidaten im Parallelkorpus (SkE). Beispiel: „Begriff“

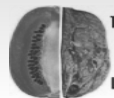
kein Begriff sein	ez ezagutu	1
kein Begriff sein	horren entzuterik ere ez izan	1
keinen andern Begriff haben als baino ez pentsatu	1
nicht allzu schnell von Begriff sein	oso azkarra ez izan	1
schwer von Begriff sein	burugogorra izan	1
sich einen stillen Begriff machen	gutxi gora behera irudika ahal izan	1
sich kaum einen Begriff machen	ozta-ozta ideia bat izan	1
sich keinen Begriff machen	ez jakin	1
sich keinen Begriff machen	ezin imajinatu ere egin	1
über alle Begriffe	ezin esan bezain	1
Der Begriff X	X	1
einen Begriff geben	aditzera eman	1

DE-EU Parallelkorpus

Parallele Wortkonkordanz aus dem lemmatisierten und POS-bezeichneten DE-EU corpus (SkE). Beispiel: „Begriff“

Corpus: DE_1
Hits: 86 (37.3 per million) [EU_1] [EU_2]
Page 1 of 5 [Go] [Next] [Last]

#95733	war frech, davon machen Sie sich keinen Begriff . Denken Sie sich, ich hab meinen Überzieher	EU_2: Zoragarria zen ... Maiatza ... Bero egiten zuen ... Eta ni lotsagabe sentitzen nintzen, ez dakizu ondo.
#110678	Gedächtnis, weil ich gerade zu dieser Zeit im Begriffe war, die Reisebeschreibung des berühmten	EU_2: Oroimenean geratu zitzaidan; izan ere, hain zuzen garai hartan, izen bereko ibiltari ezagunaren bidaiari buruzko kontakizunak irakurtzekotan nintzen.
#136122	sonnverbrannte Wüste geht. Und schon war Andreas im Begriffe , den Helden des Films sympathisch und	EU_2: Eta Andreas hasia zenean heroiart sinpatia eta abikimendua hartzen, filmea uste ez bezala zorian aldera makurtu, handik pasatzen zen zientifiko espedizio batek desertuan galdutako gizona salbatu, eta Europako zibilizazioaren gerizpera ekartzen zuen atzera.
#136185	den Helden des Films. Und schon war er im Begriff , sich zu erheben, als auf der Leinwand	EU_2: Eta jeikitze zorian zegoela, hara non panta-Han agertu zen lehentxeago, barran zegoela, tabemariaren bizkar atzean ikusitako eskola lagunaren irudia.
#139853	Hier stieß er auf einen Mann, der eben im Begriffe war, die Treppe hinaufzusteigen, und	EU_2: Han, eskaileretan gora hastear zuten gizon batekin topo egin zuen. Oso ezaguna egin zitzaion.
#151414	aufgewachsen und hatte von dorthier eine Menge von Begriffen und Schablonen beibehalten. Er hatte theoretisch	EU_2: Gainera hezimolde burges txikian hazia zen, eta ordudunik kontzeptu eta kate andana gordetzen zituen.
#154557	den Dichtungen des alten Indien ist dieser Begriff ganz unbekannt, die Helden der indischen	EU_2: India zaharreko poesietan kontzeptu hori erabat ezezaguna da, Indiako epopeietako heroeak ez dira pertsonak, pertsona matzak, enkamazio sailak baizik.
#155264	Leben. Was die Menschen jeweils unter dem Begriff "Mensch" verstehen, ist stets nur eine	EU_2: Gizonek "gizaki" kontzeptuaren bidez gehienetan ulertzen dutena, akordio burges iragankor bat besterik ez da betiere.
#160170	und erkannte mich erst, als ich schon im Begriff war, an ihm vorüberzugehen. Er stürzte	EU_2: Parez pare zetorkidan jakituna, zurrin eta nahiko begi lauso, eta jaramonik egiteke albotik pasatzea pentsatzen nuen uanean ezagutu ninduen.
#202482	Ihre Arbeit zu Ende zu führen. Wenn der Begriff der Pflicht Ihnen unbekannt ist ... " Er	EU_2: Egitekoaren ideia ezezaguna bazaizu ...
#202528	Warten Sie! " sagte Gustav höflich. " Den Begriff der Pflicht allerdings kenne ich nicht	EU_2: - Egitekoaren ideia egiaz ez dut ezagutzen, ez gehiago.
#202602	Gegenteil getan. Aber wenn ich auch den Begriff der Pflicht nicht mehr kenne, so kenne	EU_2: Baina egin beharrenden ideia gehiago ezagutzen ez badut ere, ongi ezagutzen dut ordea kulparena; agian biak gauza bera izango dira.
#203637	Menschen, einst ein hohes Ideal, ist im Begriff , zu einem Klischee zu werden. Wir Verrückten	EU_2: Gizairudia, garai batean ideal altu bat, azken batean klitxe bihurtzeko zorian dago.
#204784	lückenhafte Seelenlehre der Wissenschaft durch den Begriff , den wir Aufbaukunst nennen. Wir zeigen	EU_2: Guk horregatik, akatsez beteriko jakintza horren arimari buruzko irakaskintza osatu egiten dugu, eraikuntz deritzogu n ideien bidez.
#392066	der Herr Vorsteher schwieg. Ich war im Begriff , ihm irgendeine grobe Beleidigung ins	EU_2: Irain satsuren bat aurpegira botatzear egon nintzen.



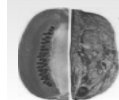
DE-EU literatura-corpus paraleloa

▼ „Gemütlichkeit“: 4 agerpen

- ▼ goxotasun
- ▼ patxada
- ▼ lasaitasun
- ▼ konfortea

▼ „Schadenfreude“: 10 agerpen

- ▼ (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



EuDeLex Online-Publishing

Suche nach deutschen Grundformen

Alemanierazko hitza bilatu

Metadate Deutsch Metadate euskaraz

Bilatu / Suche

Ampel
Ampelanlage
Ampelkarte
Ampelkoalition

EuDeLex

ehk
EUSKAL HEZKUNTZA ETA
KOMUNIKAZIOA SAIALA
UNIBERTSITATEA | www.ehk.ehu.es

Universidad
del Pais Vasco
Euskal Herriko
Unibertsitatea

Wörterbuch Deutsch-Baskisch - Alemaniera-Euskara Hiztegia

EuDeLex Suchergebnis für "Ampel"


Ampel ☆☆☆☆ • Flexion: ~, ~n

Substantiv f. •

1 (zirkulazio) **argi**; **semaforo**;
2 **argiontzi**; **lanpara**;

[Artikel kommentieren](#)

Suchergebnisse in anderen Datenbanken:

 [de.wiktionary-Artikel Ampel](#)

 [de.wikipedia-Artikel Ampel](#)

 [de.wikipedia de-eu Link zu Ampel • Semaforo](#)

EuDeLex Browser

Aminosäure
Ammoniak
Amnestie
Amok
Amoklauf
Amokläufer
amortisieren
Amour
amourös
Ampel
Ampelanlage
Ampelkarte
Ampelkoalition
Amphetamin
Amphibie
Amphibium
Amphitheater
Ampulle
Amputation

Morphologie: Substantiv

Ampel, f.

Nominativ Singular	die Ampel
Nominativ Plural	die Ampeln
Genitiv Singular	der Ampel
Genitiv Plural	der Ampeln
Dativ Singular	der Ampel
Dativ Plural	den Ampeln
Akkusativ Singular	die Ampel
Akkusativ Plural	die Ampeln


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

Suche in anderen Wörterbüchern:

 [DWDS DWDS-Suche: Ampel](#)

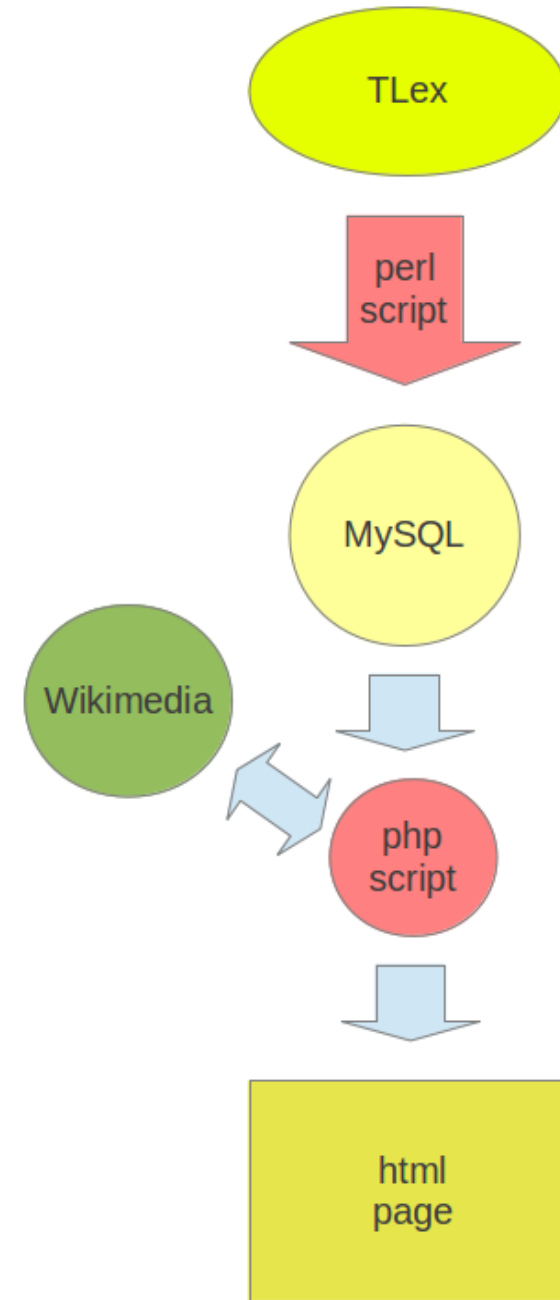
 [Wortschatz Leipzig-Suche: Ampel](#)

 [OpenThesaurus-Suche: Ampel](#)

 [DWB-Suche: Ampel](#)

Argitaratze prozesua:

- ▼ *TshwaneLex Dictionary Writing System* aplikazioak hiztegiaren edukiak XML fitxategi batera esportatzen ditu
- ▼ Guk egindako *perl*-script batek XML datuak MySQL taula batean jasotzen ditu. Taula Unibertsitateko zerbitzarian instalatzen da.
- ▼ Erabiltze interfazea DE-EU norabidean *php*-script bakarrean oinarritzen da. *php* horrek EuDeLex MySQL datu basean eta Wikimedia-API-interfazeen bitartez galdeketak egiten ditu, erabiltzailearen eskaerari erantzunez. Datu haiek *html* formatuan ematen ditu.
- ▼ EU>DE norabiderako beste *php* bat sortu dugu. Oraingo bertsioan, hemen euskarazko bestelako hiztegi-atarietara lasterbideak baino ez du eskaintzen.



EuDeLex Suchergebnis für "Aufichtigkeit"

Aufrichtigkeit ☆☆☆

Substantiv f.:

1 tolesgabetasun; **xalotasun;**

EU-ordaina: Lotura euskarazko hiztegieta

(aurretik definituriko zerrendan agertuz gero)

1 emaitza bilaketarentzat

xalotasun

xalotasun.

♦ Sencillez; ingenuidad, candidez. * *Neskatilleak, xalotasun eta apaltasun berbera dan neskatilla orrek, etzattuzala gura esan dagixun bilddur ete-zara, ba?* Otx 78. *Badakizue, izkeraren arintasuna ta xalotasuna inon ageri baldin bada, izenondoetan ageri dela.* Or LEItz 35. *Zerura nai badugu, aurren idurikoak izan bear dugu apaltasunean, xalotasunean.* Ir YKBiz 253. *Urtebeteen Usoaren aurzai izana naiz ezkongai mintzalarik, eta baten biotz ontasun eta xalotasuna ta bestearen zitalkeri ta apetak ezagutzeko beta izan det.* Etxde AlosT 45. *Neskatx aren edertasunak, xalotasunak eta erregin-antzeko jarduerak, maite-miña erne zioten bat-batean.* EG 1950 (11-12), 21.

Weiterleitung zu baskischen Wörterbuchportalen
Loturak euskarazko hiztegieta

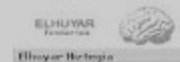
Zurück / Atzera: **Aufrichtigkeit**



Bilatu Hiztegi Batu Oinarridunear: **xalotasun**



Bilatu Orotariko Euskal Hiztegi: **xalotasun**



Bilatu Elhuyar Hiztegi: **xalotasun**



Bilatu Adorez Sinonimoen Hiztegi: **xalotasun**



Bilatu Adorez EU-ES Hiztegi: **xalotasun**



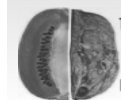
Bilatu Harluxet Enziklopedian: **xalotasun**



Bilatu euskarazko Wikiztegi: **xalotasun**



Bilatu euskarazko Wikipedia: **xalotasun**



The Bilingual Mind
Gogo Elebiduna
La Monte Bilingüe

euskara
institutua

Universidad
del País Vasco



Euskal Herriko
Unibertsitatea

2015-12-02

Hiztegi gintza elebiduna: Euskara-Alemana

Egindakoa eta proposatutakoa: Laburpena

▼ EuDeLex Hiztegia

- ▼ Lemategiak DE, EU
- ▼ Mikroegiturak
- ▼ 5.100 bat hiztegi-artikulu eskuz landuak
- ▼ Zirriborroak
 - ▼ Ordain-bikoteak
 - ▼ Euskaraz: Entitate Sintaktikoak, Adierak
- ▼ Erabiltzaile-interfazea

▼ Proposamenak

- ▼ EuDeLex eskuz landu ahala
 - ▼ EDBL eta EusWN aberastea
- ▼ Bestelakoak
 - ▼ Mahnen eskuizkribua
 - ▼ Euskararen Corpus Lexikografikoa
 - ▼ Hiztegien erabilerari buruzko ikerketa
 - ▼ Hiztegi elebidun berriak zirriborratzea: Euskara eta WordNet duten hizkuntzak



Vielen Dank für Ihre Aufmerksamkeit

Eskerrik asko

David Lindemann

david.lindemann@ehu.eus