The place of lexicography in (computer) science

Laurent Romary

The place of lexicography in (computer) science

Laurent Romary, Inria, team ALMAnaCH

Overview

Understanding the role of lexicographic work in other scholarly fields:

- Dictionaries as primary sources in the humanities
- The CS perspective: from data modelling to data mining
  - Lexicography as a rich playground for data modelling
    - Current developments in international standardisation
  - Automatic analysis of legacy print dictionaries
    - The GROBID-dictionary experience
- Perspectives – automatic data enrichments

Lexicographic works as a primary source

- Dictionaries integrate a wealth of linguistic information, but also represent a mirror of their times
  - The objective of comprehensiveness makes them essential primary sources of further humanities studies
- A wide range of possible reuse possibilities, illustrated through 2 examples:
  - Digitising the *Dictionnaire Universel* from Trévoux-Basnage (ANR project BASNUM)
  - The *Vocabulario en lengva misteca* at the service a language documentation project

An encyclopaedic witness – digitising the 1701 *Dictionnaire Universel*

- The *Dictionnaire Universel (DU)*, the first truly encyclopaedic dictionary
  - Covers general language, but above all terms from arts, crafts and sciences
  - Highly influential throughout Europe both directly and indirectly
- Antoine Furetière (1619-1688)
  - Ex-member of the *Académie française*, at loggerheads over his personal universal dictionary. The DU was published posthumously in 1690 in the Netherlands
  - Created the *Dictionnaire universel* as an encyclopaedic dictionary including all words used in France of his day
- Henri Basnage de Beauval (1657-1710)
  - A Protestant lawyer, son of a leading member of the Parliament of Normandy
  - Forced into exile after the Revocation of the Edict of Nantes that outlawed protestants
  - Succeeds the protestant philosopher Pierre Bayle as as editor of a literary and philosophical journal - *Histoire des ouvrages des savants*
  - Engaged by Leers, publisher of the DU, to compile a revised and enlarged version, published 1701
  - Uses experts to write scientific entries

ANR project BASNUM
Colleagues involved: Geoffrey Williams, Mohamed Khemakhem (Univ. de Grenoble), Ioana Galleron, Clarissa Stincone (Univ. Sorbone Nouvelle), Benoit Sagot, Laurent Romary, Pedro Ortiz (Inria)

The complex editorial history of the *Dictionnaire universel*

1690: first DU version, written and published by Antoine Furetière

- normative approach

1701: second DU version, entirely revised by Basnage de Beauval and much augmented (1/3)

- *descriptive approach*

1702: reprint of Basnage's version in 2 volumes

1708: second reprint (3 vols)

1725: new version (4 vols) revised by Brutel de La Rivière

CARME. ſ. m. Ordre de Religieux, qui eſt l'un des quatre Mendians, qui pretend tirer ſon nom du Mont Carmel en Syrie, qu'on dit avoir été habité par Elie. Ils ont été amenez en France par le Roi Louïs IX. Il eſt celebre par la devotion du Scapulaire, & par la viſion de Simon Stock, à qui il fut donné par la Sainte Vierge. Sur quoy de Launoy a écrit une curieuſe Diſſertation. La Vierge attacha ce privilege au Scapulaire, & à l'habit des *Carmes*, que ceux qui meurent le Samedi chargez de ces pieuſes depouilles, ſont exempts des flâmes du Purgatoire. Les *Carmes* ſe diſent oncles de J. CHRIST, & freres de la Vierge. On dit les *Carmes du grand Couvent*; les *Carmes Mitigez*, qu'on nomme à Paris *Billettes*; & les *Carmes Dechauſſez*, qui ont été reformez des autres. Dans des Theſes ſoutenuës à Beziers mentionnées dans le Journal de Hollande, on dit qu'il eſt fort probable que Pythagore étoit *Carme*, & que les Druides des Gaulois avoient auſſi les obſervances regulieres des *Carmes*.

MONT CARMEL, eſt un Ordre Militaire de Chevaliers Hoſpitaliers, fondé par le Roi Henri IV. ſous le titre, l'habit & la Regle de Nôtre Dame du *Mont Carmel*; & en conſequence des Bulles du 16. Fevrier 1607. il a été uni à l'Ordre des Chevaliers de St. Lazare de Jeruſalem, par acte du dernier Octobre 1608. avec toutes ſes Commenderies, Prieurez & autres biens pour ſa dotation.

CARME, eſt auſſi une eſpece d'acier. Voyez ACIER.

CARME, eſt auſſi un vieux mot qui ſignifioit un *vers*. Il vient du Latin *carmen*; & en ce ſens il eſt tout-à-fait hors d'uſage.

```
<entry xml:lang="fre" xml:id="Carme">
  <form><orth rendition="#uc">carme</orth>
    <gramGrp><pos expand="Substantif">s.m.</pos></gramGrp></form>
  <sense n="1">
    <def> Ordre de Religieux, qui est l'un des quatre Mendians, qui
      pretend tirer son nom du Mont Carmel en Syrie, qu'on dit avoir
      été habité par Elie. </def>
    <note> Ils ont été amenez en France par le Roi Louïs IX. Il est
      celebre par la devotion du Scapulaire, &amp;amp; par la vision de
      Simon Stock, à qui il fut donné par la Sainte Vierge. Sur quoy
      de Launoy a écrit une curieuse Dissertation. La Vierge attacha
      ce privilege au Scapulaire, &amp;amp; à l'habit des <hi
        rendition="#i">Carmes</hi>, que ceux qui meurent le Samedi
      chargez de ces pieuses dépouilles, sont exempts des flâmes du
      Purgatoire. Les <hi rendition="#i">Carmes</hi> se disent oncles
      de <name ref="#Jesus_Christ_ISN0000000120370699"
        >J.Christ</name>, &amp;amp; freres de la Vierge. On dit les <seg
          type="terme_variant" rendition="#i">Carmes du grand
          Couvent</seg> ; les <seg type="terme_variant" rendition="#i"
            >Carmes Mitigez</seg>, qu'on nomme à Paris <hi
              rendition="#i">Billettes</hi> ; &amp;amp; les <seg
                type="terme_variant" rendition="#i">Carmes Dechaussez</seg>,
      qui ont été reformez des autres. Dans des Theses soutenuës à
      Beziers mentionnées dans le Journal de Hollande, on dit qu'il
      est fort probable que Pythagore étoit <hi rendition="#i"
        >Carme</hi>, &amp;amp; que les Druides des Gaulois avoient aussi
      les observances regulieres des <hi rendition="#i"
        >Carmes</hi>.</note></sense>
  <sense n="2"><form><orth rendition="#sc">mont carmel</orth>,</form>
    <def> est un Ordre Militaire de Chevaliers Hospitaliers, fondé par
      le Roi Henri IV, sous le titre, l'habit &amp;amp; la Regle de Nôtre
      Dame du Mont Carmel ; &amp;amp; en consequence des Bulles du 16.
      Fevrier 1607. </def>
    <note>Il a été uni à l'Ordre des Chevaliers de St.Lazare de
      Jerusalem, par acte du dernier octobre 1608. avec toutes ses
```

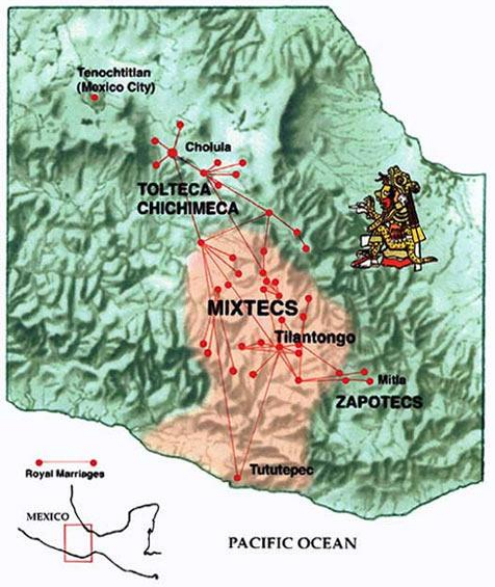
Understanding the dictionary: some humanities research questions

- What source texts? Dictionaries and other texts?
- What links between the DU and *Histoire des ouvrages des savants*
- What lexicographical model between prescription and description?
    - the place of ‘good’ usage and the ‘best ‘ authors
    - the dictionary as a language teaching tool – what users? what means?
    - The characterisation of terms
    - The role of contemporary scientific and literary networks
- To what degree the 1701 DU was theologically a "protestant dictionary”
- What changes were brought in between 1690 and 1701, between 1701 and 1725/27, and between 1701 and the 1704 Trévoux
- Who authored which entries in 1701? - Basnage and his specialist informers.

Linguistic description of Mixtepec-Mixtec

- Sa'an Savi "*rain language"*
- ISO 639-3 code: 'mix'
- San Juan de Mixtepec - Juxtlahuaca district (Oaxaca, MEX)
- "Vigorous" status but highly under-resourced
- Oto-Manguean, Mixtecan, Mixtec-Cuicatec, Mixtepec-Mixtec
- Tonal
- Spoken data mostly collected in sessions working with speakers from a small village called Yucunani in the San Juan Mixtepec municipality
- Estimated (+-9,000 -10,000 speakers)

  Source: (INEGI, 2010)
- Phonology has been studied by Pike and Ibach (1978); Paster and Beam de Azcona (2004-2007);
- Beckman and Nieves-SIL (2005-current) published booklets and are working on developing orthography



The research project

- Language documentation: (Jack Bowers' *PhD)*
- Primary sources of language data:
  - Speaker consultations (recordings, new written material..)
  - +- 40 Children's Booklets (SIL)
  - Public sources (YouTube, *other*)
  - Examples from academic papers
- Goals:
  - TEI Corpus
  - Linguistic descriptions
  - TEI Dictionary *(actually 2 dictionaries, 1 general, 1 inflectional)*
  - *(Etymology)* would like to create data contents and structure that can be copied and integrated into treatment of related languages

Jack Bowers, Laurent Romary. TEI and the Mixtepec-Mixtec corpus: data integration, annotation and normalization of heterogeneous data for an under-resourced language. *6th International Conference on Language Documentation and Conservation (ICLDC)*, Feb 2019, Honolulu, United States. ⟨hal-02075475⟩

Overview of the Source & Output

- ‘Vocabulario en lengva misteca’ published by the Dominican Francisco de Alvarado (1593)
- Variety from Teposcolula Mexico (Mixteca Alta)
  - Classical Mixtec/Colonial Mixtec/Yucu Ndaa
- Entries based on three earlier dictionary sources:
  - Castilian-Nahuatl (Valley of Mexico, 1571)
  - Castilian-Zapotec (Valley of Oaxaca, 1578)
  - Castilian-Latin (1492)
- PDF re-organized, modernized version ‘Voces de Dzaha Dzahui’ (Jansen & Pérez Jiménez, 2009)
- TEI dictionary produced contains roughly 26,600 entries and related entries.

Versions of the resource

- Original (Printed: 1593) > (facsimile edition 1965)
- Mesolore (Bakewell & Hamman, 2001)
  - Digitized from scanned copy
- Jansen and Pérez Jimenez 2009

A. ANTE B.

PRIMERA letra del a. b. c.
A. prepoſició por cerca,apud. nuu como nuuyàhui.
A. prepoſició por hazia.nuu quaha dzúhua,
A. aduerbio de llamar. dzi.
A. aduerbio para llamar quãdo ſe me ha oluidado algo. dzi. vt. na hadzi. nahacadzi.
A. interiectio. del que ſe quexa de alguna enfermedad. Hmij.
A. interiectio deſiderantis. ha.
A. interiectio ridentis. ha ha ha.
A. interiectio admirantis. a a. a. hi. hi. hi.
A. del que halla a otro en maleficio. a. a.
A alguna parte o en alguna parte. aduerbio. huadza, huadzaca.
Abad. Prelado. dzutu. yõdadzi ſiña ñuhu, dzutu ñohoſiña ñuhu, dzutu yocuvui dzini, yocuvui nuu.
Abad ſer. yondadzindi ſiñañuhu, f. cond. ñohondi ſiñañuhu. f. coho. yocuvuidzinindi. yocuvui nuundi.
Abadia dignidad. ſaſindadzi ſiñañuhu, ſaſiñoho ſiñañuhu, ſaſicuvuinuu.
Abadeſa, prelada de monjas. ñaha yondacañaha yuq dzehe. ñahayocuvuidzini.
Abadejo eſcarauajo. tenuyuq.
Abahar algo conel huelgo. yochidzo yoco yuhundi yotaa yoco yuhundi,
Abahar aſi coſa honda. yotniño yoco yuhundi.
Abahada coſa aſi. ſaſi cohondaa yoco. ſaſinucoho yoco yuhu ſaniſa codzo yoco yuhu.
Abahar algo poniendolo al baho q̃ ſale dela olla. yodzaqnindi. yodzacuidzindi yoco.
Abahar algo dentro de alguna olla para que ſe cõſerue con el baho y no ſe pudra. yodzacaanuuyocondi.
Abahar ſopas. yodzachitundi yoco yotniño yocondi. yodzacutu yocondi.
Abahadas ſopas. dzita ninucoho yoco.
Abahar por hechar el huelgo de la boca. yodzacaindi yoco. yodza canãdi yoco, yodzaqhũdi yoco.
Abahar apoſento. yodzacuidzindi huahi. yotniño yocondi. yondadzaindi. yondadzaindi.
Abahado apoſento. huahi ninucui ñe huidzi. huahi ninucoho yoco. huahi ninacuidzi. huahi ninduvui ini.
Abalançarſe, hechandoſe por los ſuelos. yocoo cavuan dayẽdi, yo ſacavuandayendi. f. qcavua.
Abalançarſe de arriba abaxo. yondiyoninondi.
Abalançarſe metiendoſe entre otros. yoſanudzavuãdi. f. qnanu. yoſanánudzavuandi f. qnánu, yoſivui cuinomañundi, yocai nánu dzavuandi. f. quivui. yoſivuindutnahandi.
Aballar mouer cõ difficultad, duz

folio 1r column 1

PRIMERA letra del a. b. c.
A. preposicion por cerca, apud. nuu como nuuyahui.
A. preposicion por hazia. nuu quaha dzuhua.
A. aduerbio de llamar. dzi.
A. aduerbio para llamar quando se me ha oluidado algo. dzi. vt. nahadzi. nahacadzi
A. interiectio. del que se quexa de alguna enfermedad. Hmii.
A. interiectio desiderantis. ha.
A. interiectio ridentis. ha ha ha.
A. interiectio admirantis. a a. a. hi. hi. hi.
A. del que halla a otro en maleficio. a. a.
A alguna parte o en alguna parte. aduerbio. huadza, huadzaca.
Abad. Prelado. dzutu. yodadzi siña ñuhu, dzutu ñoho siña ñuhu, dzutu yocuvui dzini, yocuvui nuu.
Abad ser. yondadzindi siñañuhu, f. cond. ñohondi siñañuhu. f. coho. yocuvuidzinindi. yocuvui nuundi.
Abadia dignidad. sasindadzi siñañuhu, sasiñoho siñañuhu, sasicuvuinuu.
Abadesa, prelada de monjas. ñaha yondacañaha yuq dzehe. ñaha yocuvuidzini.
Abadejo escarauajo. tenuyuq.
Abahar algo con el huelgo. yochidzo yoco yuhundi yotaa yoco yuhundi.

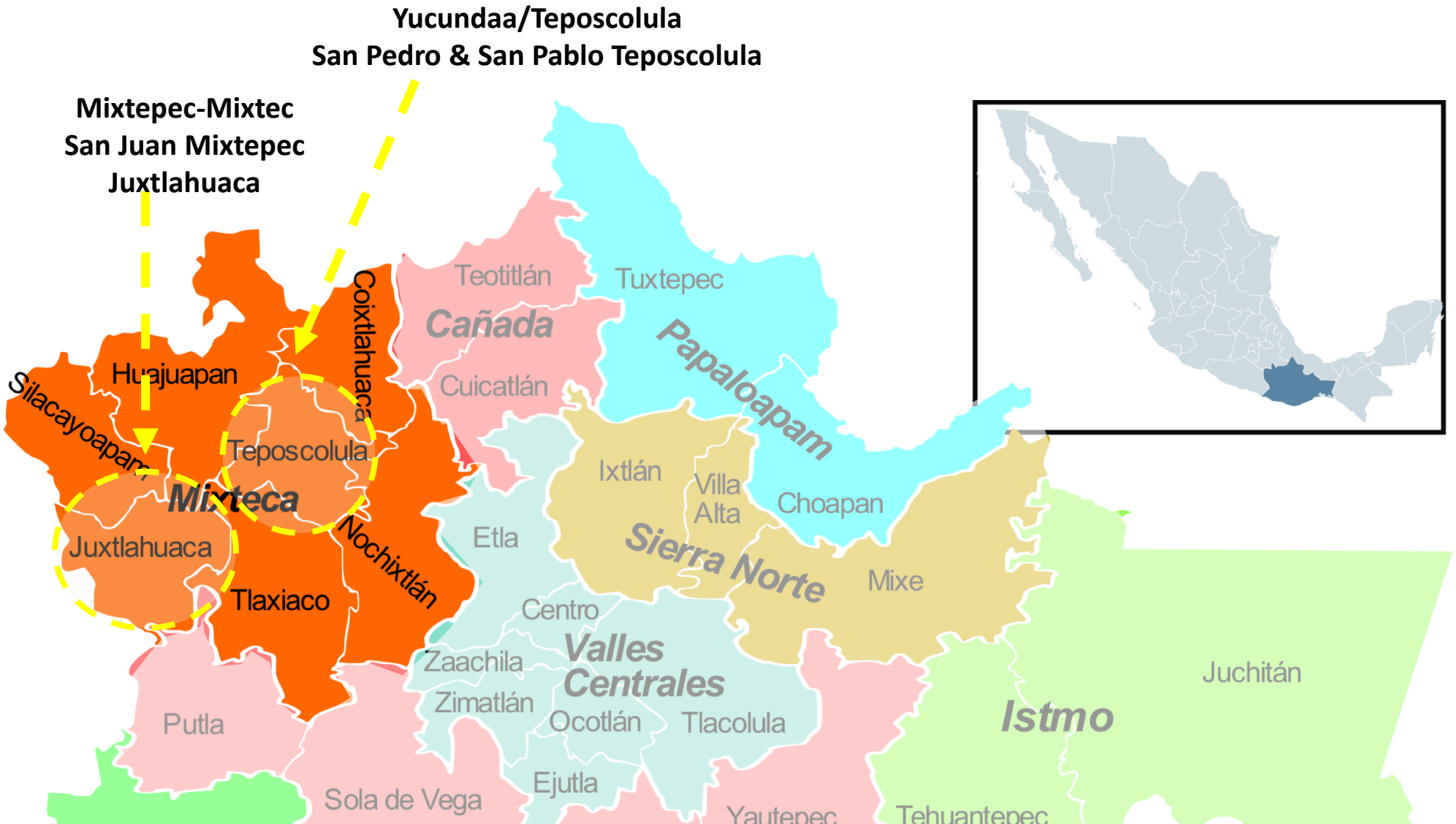
folio 1r column 2

Abahar asi cosa honda. yotniño yoco yuhundi.
Abahada cosa afi. sasi cohondaa yoco. sasinucoho yoco yuhu sanisa codzo yoco yuhu.
Abahar algo poniendolo al baho que sale de la olla. yodzaqnindi. yodzacuidzindi yoco.
Abahar algo dentro de alguna olla para que se conserue con el baho y no se pudra. yodzacaanuuyocondi.
Abahar sopas. yodzachitundi yoco yotniño yocondi. yodzacutu yocondi.
Abahadas sopas. dzita ninucoho yoco.
Abahar por hechar el huelgo de la boca. yodzacaindiyoco. yodza canandiyoco, yodzaqhundi yoco.
Abahar aposento. yodzacuidzindi huahi. yotniño yocondi. yondadzacuidzindi. yondadzainindi.
Abahado aposento. huahi ninucuiñe huidzi. huahi ninucoho yoco. huahi ninacuidzi. huahi ninduvui ini.
Abalançarse, hechandose por los suelos. yocoo cavuandayendi, yosacavuandayendi. f qcavua.
Abalançarse de arriba abaxo. yondiyoninondi.
Abalançarse metiendose entre otros. yosanudzavuandi. f. qnanu. yosananudzavuandi f. qnanu, yosivui cuinomañundi, yocai nanu dzavuandi. f. quivui. yosivuindutnahandi.
Aballar mouer con difficultad, duz

A

a a: a (del que halla a otro en maleficio)
a a a: a (*interiectio admirantis*)
a dzuchica añandaa: poco más o menos
a dzuchica caa cuvui: poco más o menos
a dzuchica coo cuvui: poco más o menos
a hua dzevui: o no
a hua dzevui dzavua: pues no
a huihica añandaa: poco más o menos
a huitnani: ahora poco
a na ndehe cuvui ndatu nicay: o bienaventurado, o dichoso
a ñaha: o no; pues no
a sa dzevui: pues no
a yoo: por ventura alguno
a yoo ee ñahando: por ventura alguno de vosotros
aa: de manera que
aa: ya, acordándoseme lo que se me había olvidado
aa dzuhua huii: así, así (sonriéndose)
aa ndica huii: ya (acordándoseme lo que se me había olvidado)
adzi: o (*disyuntiva*); por ventura; quizás
adzi: suave cosa
adzi cuvui: o (*disyuntiva*); por ventura; quizás
adzi q cuvui: por ventura
adzi yoo: por ventura alguno
adzi yoo ee ñahando: por ventura alguno de vosotros
ahua: ay, quejándose la mujer
ama: así
ama: bien está (otorgando); sí
amana: ¿cuándo? (*adverbio interrogativo*), ¿en qué tiempo?
amana cuiya: en algún tiempo
amana cuvui huatu inindo: cuandoquiera que quisieres
amana na ndita ñumana nuundo: ¿cuándo has de despertar?
amana na ndotondo: ¿cuándo has de despertar?
amana na tahui inindo: ¿cuándo has de volver en ti?; ¿cuándo has de despertar?
amana quevui: en algún tiempo
amana quevui sa cuvui inindo: cuandoquiera que quisieres
amanaca: ¿cuándo, en qué tiempo?
amani: de tarde en tarde o raras veces
andaya: infierno, lugar de dañados
andevui: cielo
andevui isi ndaa tiñoo: cielo estrellado
angel nicoo coo ndaa ndita ña: ángel de mi guarda
angel yondaca ñaha: ángel de mi guarda
aniñe: palacio
anuhu: abismo; centro de la tierra; infierno, lugar de dañados
anuhu: profundo
anuhu maa: profundo
anuhu naa: infierno, lugar de dañados
anuhu ndahui: infierno, lugar de dañados
atana: ojalá
atu: amarga cosa; áspero al gusto
aya: amarga cosa; áspero al gusto

La Mixteca (Mixtec Region)



Utility/Purpose of Endeavour

- Increase coverage of relevant lexical material in Mixtepec-Mixtec documentation (ISO 639-3 [mix])
  - Link and cross-reference in Mixtepec TEI dictionary
- Machine searchable data set for:
  - Study of the Yucu Ndaa variety
  - Historiographical and philological research
- Create a more cohesive body of pan-Mixtecan resources
  - Vocabulary for cross Mixtecan comparison; (81 Varieties of Mixtec)
- TEI format can easily be exported into other formats for non-TEI users

Integration into Mixtepec-Mixtec Project: TEI Structure of Output

- Goal to match the structure used in the Mixtepec-Mixtec TEI dictionary (Bowers & Romary 2018)

```
<entry xml:id="fruit-plaintain">
  <form type="lemma">
    <orth xml:lang="mix">nchika</orth>
    <pron xml:lang="mix" notation="ipa">nʤiká</pron>
  </form>
  <gramGrp>
    <pos>noun</pos>
  </gramGrp>
  <sense corresp="http://dbpedia.org/resource/Plantain">
   <usg type="domain">Fruit</usg>
    <cit type="translation">
      <form>
        <orth xml:lang="en">plantain</orth>
      </form>
    </cit>
    <cit type="translation">
      <form>
        <orth xml:lang="es">plátano</orth>
      </form>
    </cit>
  </sense>
...
</entry>
```

nchika [nʒiká] *(noun)*
[FRUIT] **plantain, plátano**

Mixtepec-Mixtec

```
<entry xml:id="plátano">
  <form type="lemma">
    <orth>chita</orth>
  </form>
  <gramGrp>
    <pos>noun</pos>
  </gramGrp>
  <sense corresp="http://dbpedia.org/resource/Plantain">
    <usg type="domain">Fruit</usg>
    <def xml:lang="es">plátano</def>
    <def xml:lang="en">plantain</def>
  </sense>
</entry>
```

chita *(noun)*
[FRUIT] **plantain, plátano**

Classical Mixtec

Going further: modelling and standardising lexical resources

Lexical resources in their varieties

- A variety of contexts and forms
  - Legacy dictionaries, dialectological studies, NLP lexica
  - Full form, etymology, corpus based research
  - Word document, database, shoebox, proprietary XML…
  - Lexical vs. Editorial views
  - Onomasiological vs. semasiological structures

Lexicography or terminology

- Lexicography
  - Generic view on “words”
  - Attempt to provide a large coverage of a language
  - Semasiological view
    - Word > meaning(s)
- Terminology
  - *Term*: form associated to a specific concept within a given domain
  - Onomasiological view
    - Concept > various possible linguistic forms
- Depends on available data, objectives and user scenarios

Comparing approaches

Semasiological approach

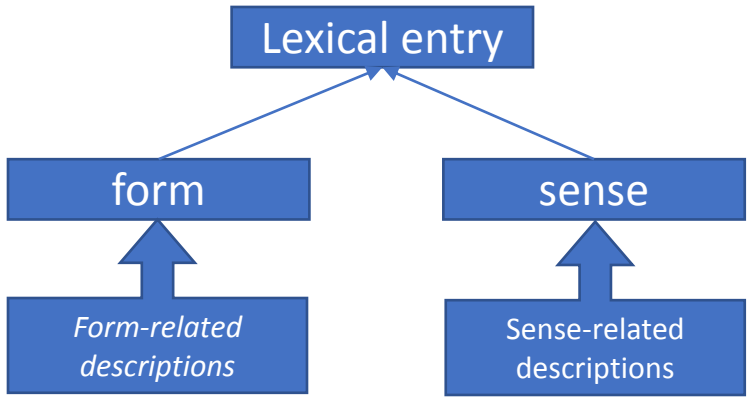
- Large coverage
- All parts of speech
- Build-in polysemy
  - Multiple senses for the same entry
- Referential synonymy

Onomasiological approach

- Domain oriented
- Essentially nouns
  - Extension to verbs, adjectives
- No polysemy (needs to be reconstructed)
- Build-in synonymy
  - Multiple terms for the same concept

Basic modelling of lexical components

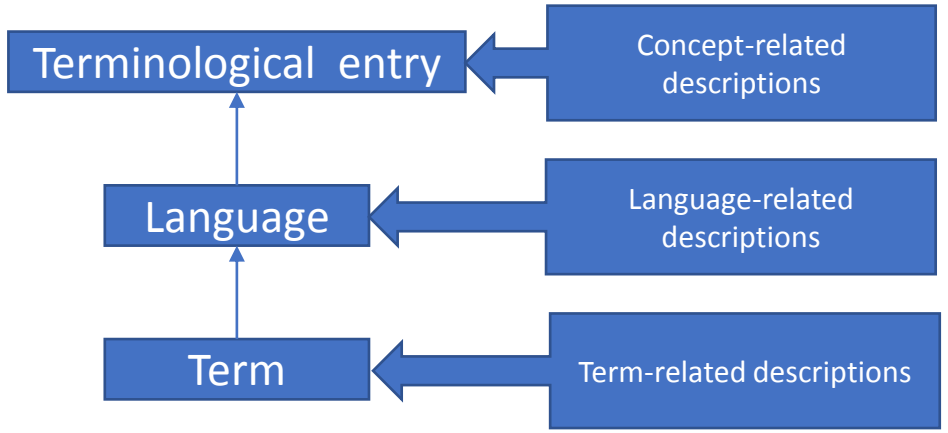
Semasiological models



Issues

- Various levels/sensibilities in entry groupings: homonyms, families (e.g. roots)
- Providing a neat way of representing lexical dependencies: from “see also” to multi-word expressions

Onomasiological models



Issues

- Representing conceptual relations between entries
- Providing fine-grained semantic information at term level (e.g. usage, translation equivalents)

Issues

- Making the appropriate choice of model
- Integrating information between the two types of models

Why standardizing all this?

- Defining methods, models and format to facilitate
  - Exchange of lexical data
  - Pooling heterogeneous lexical data
  - Interoperability between software components
    - Search engines, layout, extraction of linguistic properties
  - Comparability of results
    - E.g. Linguistic coverage of lexical databases
  - Exchange of ideas within a community with a common background

Standardization initiatives for lexical/terminological resources

- TEI
  - P5 edition of the guidelines
    - Cf. specification platform (ODD)
    - Dictionary chapter
    - Former terminology chapter (ancestor of TBX)
- ISO
  - ISO/TC 37: Language and terminology
    - ISO/TC 37/SC 3: ISO 16642 (TMF), ISO 30042 (TBX)
    - ISO/TC 37/SC 4: ISO 24613 (LMF)
- W3C
  - SKOS, Ontolex

In the beginning

Text archives
Humanities
Standards
SGML

Not intended (immediately) for individual scholars

1. Novembre 1987: Vassar College, Poughkeepsie

A quick historical overview

- 1960’s — GML (Generalized Markup Language) by IBM
- 1970’s & 1980’s — ANSI initiates project to develop a Standard text-description language based on GML
- 1983 — SGML becomes an industry standard
- 1986 — SGML (Standard Generalized Markup Language) becomes an ISO standard: ISO 8879:1986
- **1987 — TEI (Text Encoding Initiative)**
- 1990 — HTML 1.0 (HyperText Markup Language)
- 1992 — TEI edition P3 (Michael Sperberg-McQueen and Lou Burnard, eds)
- 1997/1998 — XML 1.0 (eXtensible Markup Language) (Tim Bray, Jean Paoli and Michael Sperberg-McQueen, eds)

The TEI Dictionary chapter

- Initially designed within a working group lead by N. Ide and J. Veronis
- Accounts for both presentational and database views
  - Cf. <entry>, <entryFree>, … and <dictScrap>
- Based on a hierarchical abstract model (crystals)
  - <form>: for characterizing the orthographic or phonetic form of the word
    - <orth>, <pron>, etc.
  - <gramGrp>: grammatical features
    - May characterize an entry, a specific form or a specific sense
    - <pos>, <gen>, generic <gram> feature
  - <sense>: iterative and recursive
    - May contains definitions, examples, etymological information, translations, etc.

IRL: Petit Larousse illustré (1906)

```
<entry xml:id="pléthore" n="1906-001_unknown">
    <form type="lemma"><orth>PLÉTHORE</orth></form>
    <gramGrp><pos expand="nom">n.</pos>
        <gen expand="féminin">f.</gen></gramGrp>
    <etym><pc>(</pc>du <lang expand="grec">gr.</lang>
        <mentioned>plêthorê</mentioned><pc>,</pc>
        <gloss>plénitude</gloss><pc>)</pc><pc>.</pc></etym>
    <sense><def>Surabondance de sang, d'humeurs</def><pc>.</pc></sense>
    <sense><usg type="style" rend="italic" expand="figuré">Fig.</usg>
        <def>Surabondance quelconque amenant un état fâcheux</def>
        <pc>:</pc>
        <cit type="example">
         <quote>la pléthore des capitaux cause la diminution du taux de l'intérêt</quote>
        </cit><pc>.</pc></sense>
</entry>
```

Hervé Bohbot, Francesca Frontini, Fahad Khan, Mohamed Khemakhem, Laurent Romary. Nénufar: Modelling a Diachronic Collection of Dictionary Editions as a Computational Lexical Resource. *ELEX 2019: smart lexicography*, Oct 2019, Sintra, Portugal. ⟨hal-02272978⟩

Advantages of being in the TEI framework

- Benefitting from the TEI environment
  - TEI Modelling Language: ODD
    - Customizing the guidelines within a project (e.g. restraining possible values)
  - Availability of a wealth of additional elements (~600)
    - E.g. annotating textual content, reflecting the specificities of the source etc.
- Standardisation reactivity
  - Issuing GitHub tickets for resolving bugs or introducing new features
- A community of experts
  - Support through the mailing list
- Main characteristic (drawback?): +very+ flexible

Going ISO to provide a stable background

- Advantages of going ISO
  - International approval of ISO members
    - And international expert participation by construction
  - Stable background that is easy to reference (and known by third parties, non linguistic geeks, our institutions etc.)
- From a lexical point of view
  - Providing a generic model, independently of any specific implementation/serialisation
    - Stabilizing concepts, constraints and vocabulary
  - With the on-going LMF revision: Introducing the TEI as one possible model

LMF

ISO 24613:2008 Language resource management — Lexical markup framework (LMF)

- Developed within ISO TC37/SC4/WG4
  - TC 37: Language and terminology
  - TC37/SC4/WG4: Lexical resources
- Shortcomings
  - Bulky: a single document with annexes
    - Major hindrance to revision
  - Complex modelling
    - Complex relationships between classes, redundant mapping mechanisms
  - Complex and ad hoc serialisation
  - Does not cover prominent information: Etymology and Diachrony

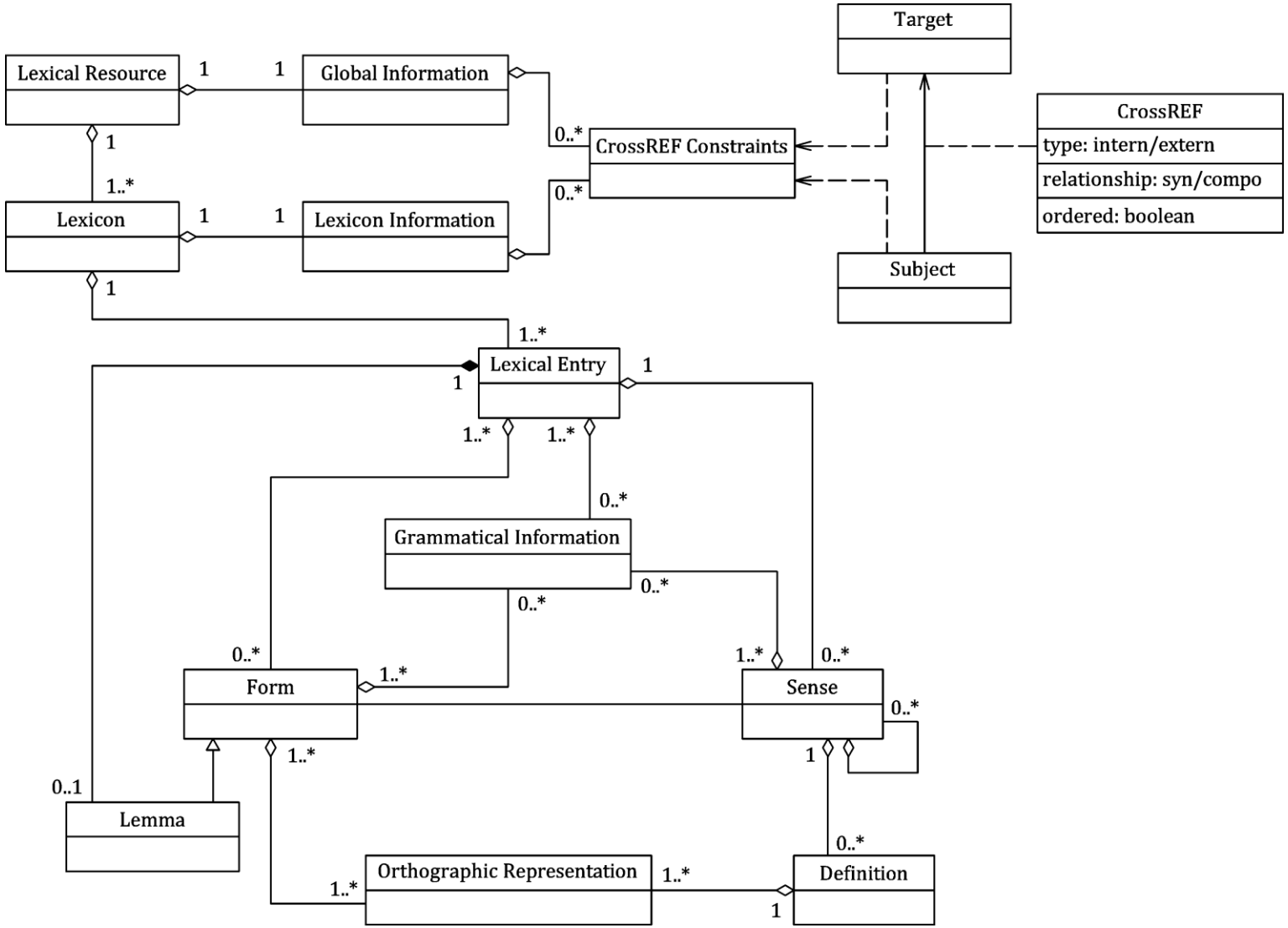
Laurent Romary, Mohamed Khemakhem, Fahad Khan, Jack Bowers, Nicoletta Calzolari, et al.. LMF Reloaded. *AsiaLex 2019: Past, Present and Future*, Jun 2019, Istanbul, Turkey. ⟨hal-02118319⟩

LMF Reloaded: Abstract Modelling

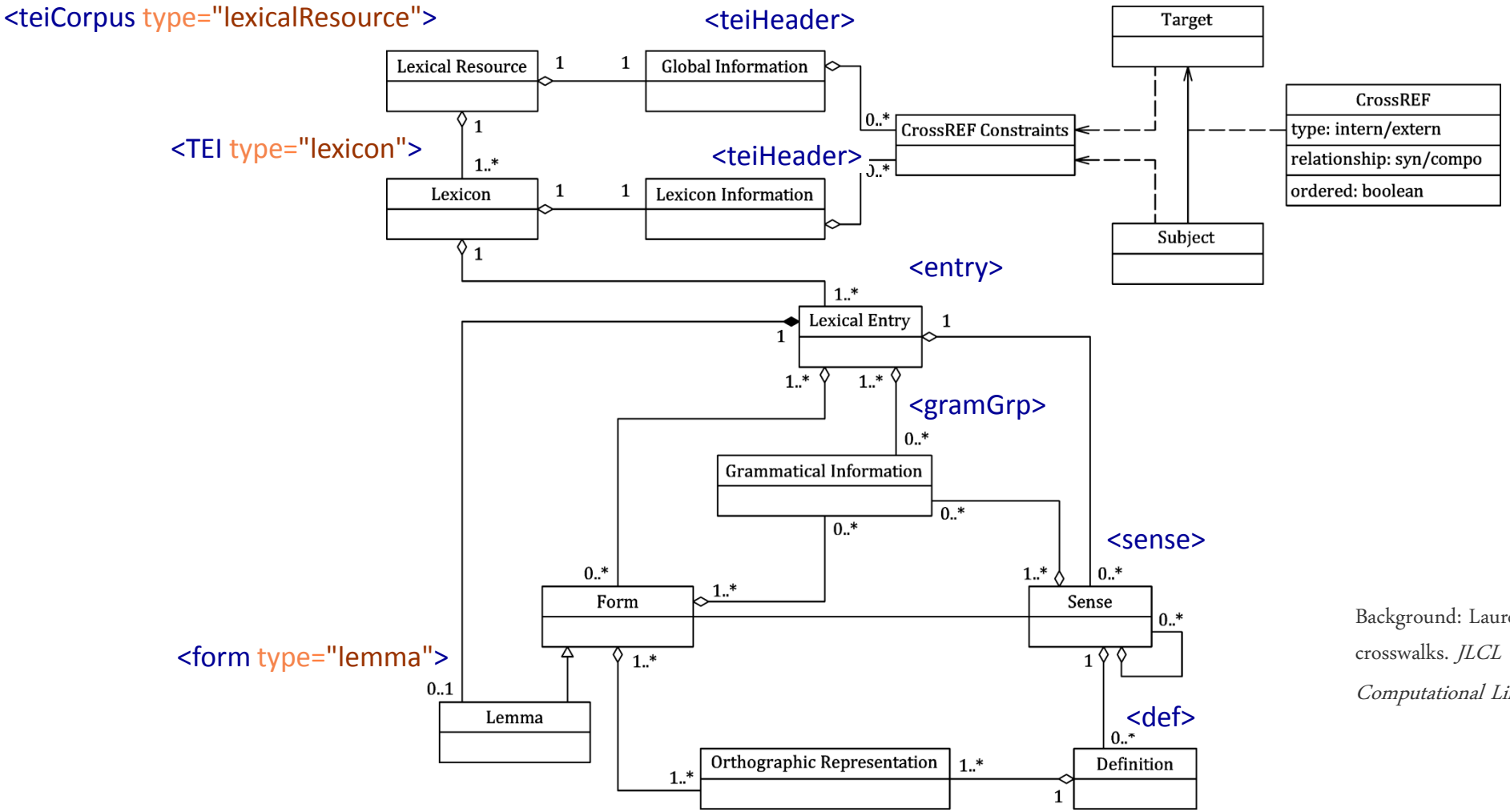
Restructuring: Multi-part standard

- ISO 24613-1 - Core model (published in June 2019)
- ISO 24613-2 - Machine Readable Dictionaries (MRD) model
- ISO 24613-3 - Diachrony-Etymology
- ISO 24613-4 - TEI serialisation
- ISO 24613-5 - LBX serialisation
- ISO 24613-6 - Syntax and Semantics
- ISO 24613-7 - Morphology

Language resource management — Lexical markup framework (LMF) — Part 1: Core model

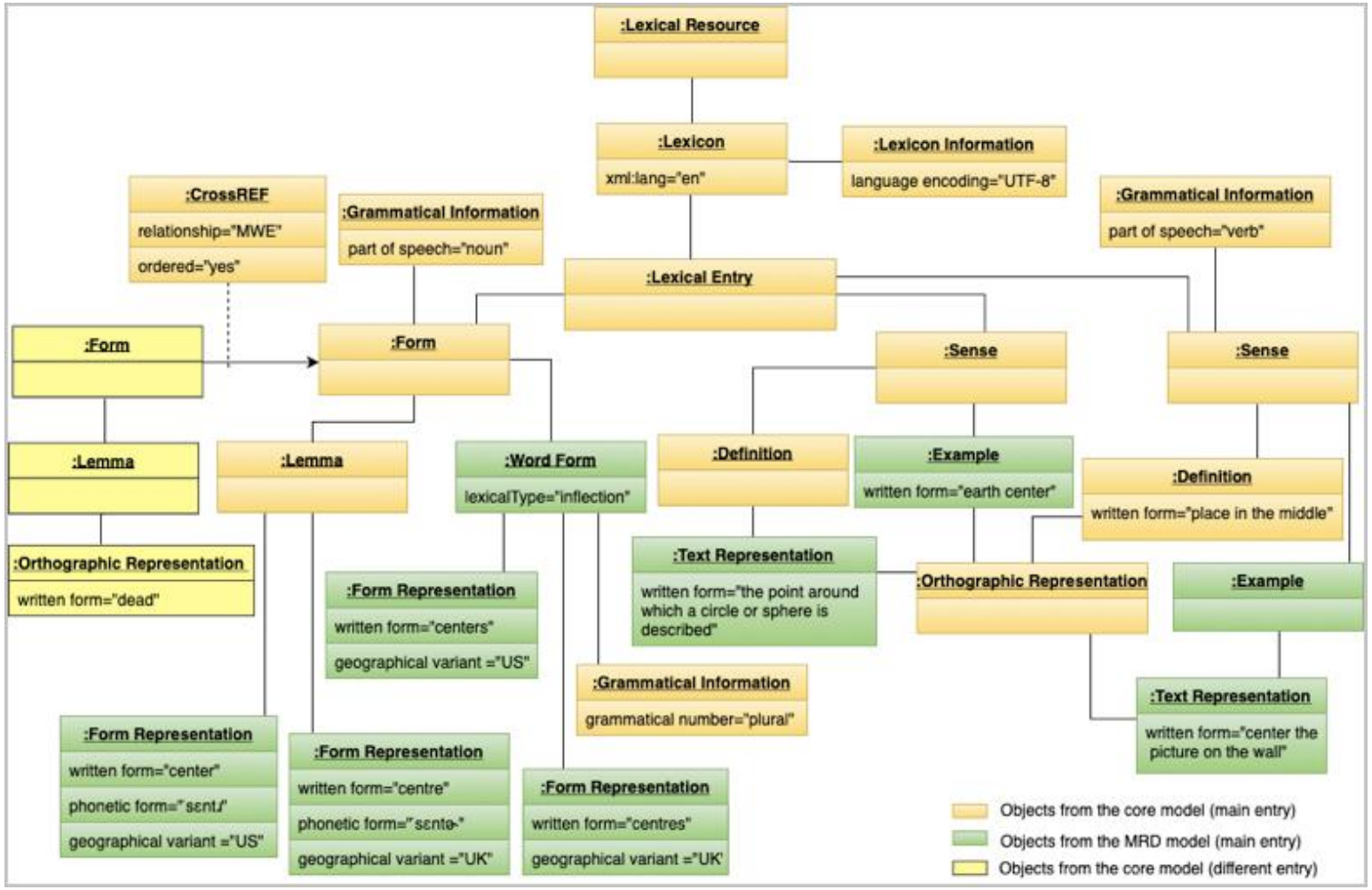


From LMF to TEI – serialising one with the other (Part 4 – TEI serialisation)



Background: Laurent Romary. TEI and LMF crosswalks. *JLCL - Journal for Language Technology and Computational Linguistics*, 2015, 30 (1). ⟨hal-00762664v4⟩

Example: inflectional (full-form) lexicon



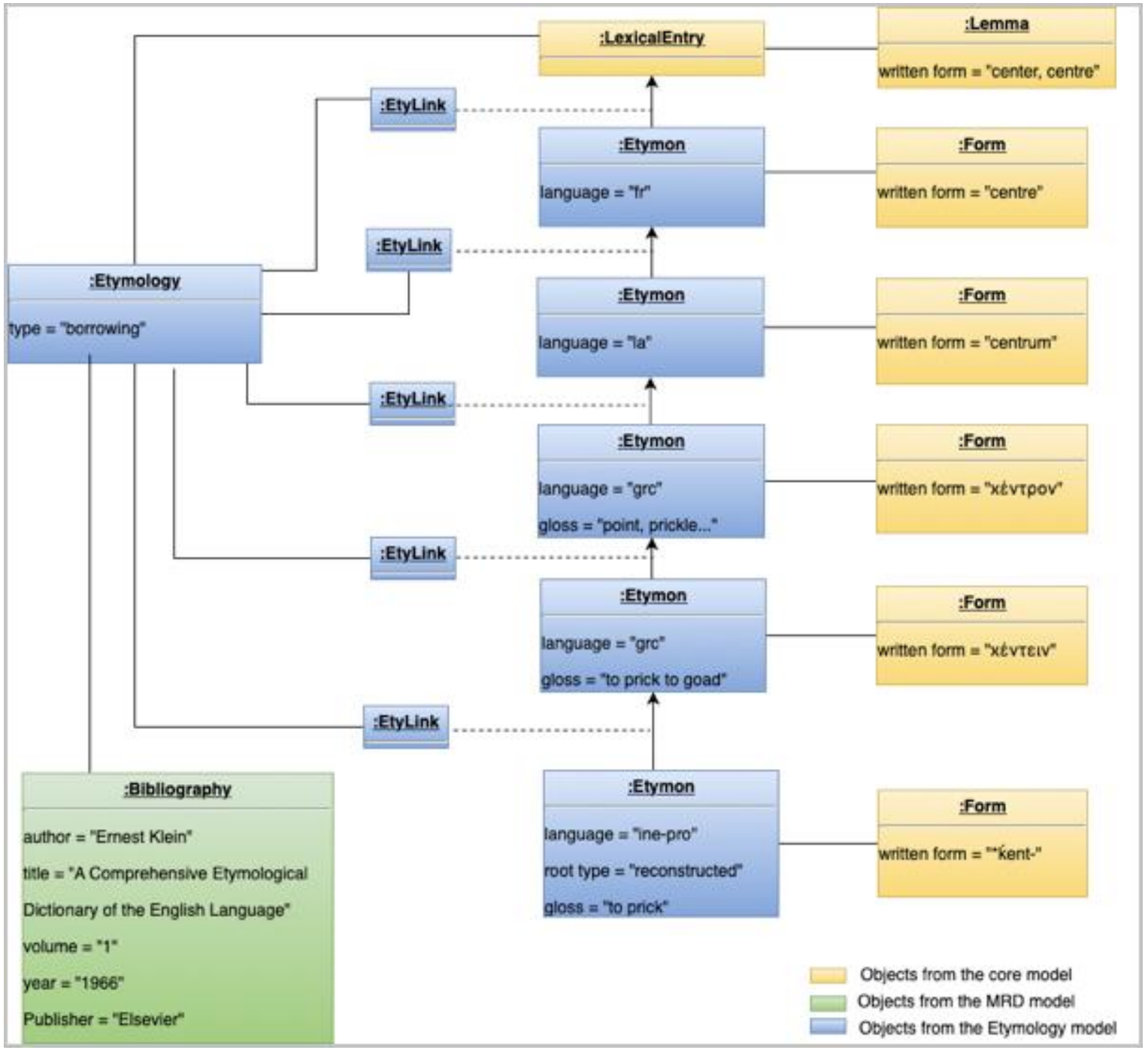

```
<entry>
   <form type="lemma" xml:id="center_form">
      <orth>center</orth>
      <pron>ˈsɛntɹ</pron>
      <gramGrp>
         <pos>noun</pos>
      </gramGrp>
      <usg type="geo">U.S</usg>
      <form type="variant">
         <orth>centre</orth>
         <usg type="geo">U.K</usg>
         <pron>ˈsɛntə</pron>
      </form>
   </form>
   <form type="inflected">
      <orth>centers</orth>
      <usg type="geo">U.S</usg>
      <gramGrp>
         <number>plural</number>
      </gramGrp>
   </form>
   <form type="inflected">
      <orth>centres</orth>
      <usg type="geo">U.K</usg>
      <gram type="number">plural</gram>
   </form>
   <sense>
      <def>the point around which a circle or sphere is described</def>
      <cit type="example">
         <quote>earth center"</quote>
      </cit>
   </sense>
   <sense>
      <gramGrp>
         <pos>verb</pos>
      </gramGrp>
      <def>place in the middle</def>
      <cit type="example">
         <quote>center the picture on the wall</quote>
      </cit>
   </sense>
   <re type="multiWordExpression">
      <form>
         <seg corresp="#dead_form" n="1">dead</seg>
         <seg corresp="#center_form" n="2">center</seg>
      </form>
   </re>
</entry>
```

LMF reloaded - Etymology

ISO 24613-3, prepared for DIS ballot (as of November 2019)
new classes: Etymology, Etymon, Cognate and EtyLink

center, centre, n. — F. *centre*, fr. L. *centrum*, fr. Gk. κέντρον, 'point, prickle, spike, ox goad, point round which a circle is described', from the stem of κέντειν, 'to prick, goad', whence also κέντωρ, 'a goader, driver', κεστός (for *κεντ-τός), 'embroidered', κέστρᾱ, 'pickax', κοντός, 'pole', fr. I.-E. base ***ḱent-**, 'to prick', whence also Bret. *kentr*, OIr. *cinteir*, 'a spur', OHG. *hantag*, 'sharp, pointed', Lett. *sīts*, 'hunter's spear', *situ*, *sist*, 'to strike', W. *cethr*, 'nail'. Cp. **centrifugal, centripetal, concentrate, eccentric, Dicentra, paracentesis.** Cp. also **cestrum, cestus,** 'girdle', **kent,** 'a pole', **quant,** 'a pole'.
Derivatives: *center*, *centre*, intr. and tr. v., *center-ing*, *centr-ing*, *centre-ing*, n.

Source: Klein's *Comprehensive dictionary of the English Language*



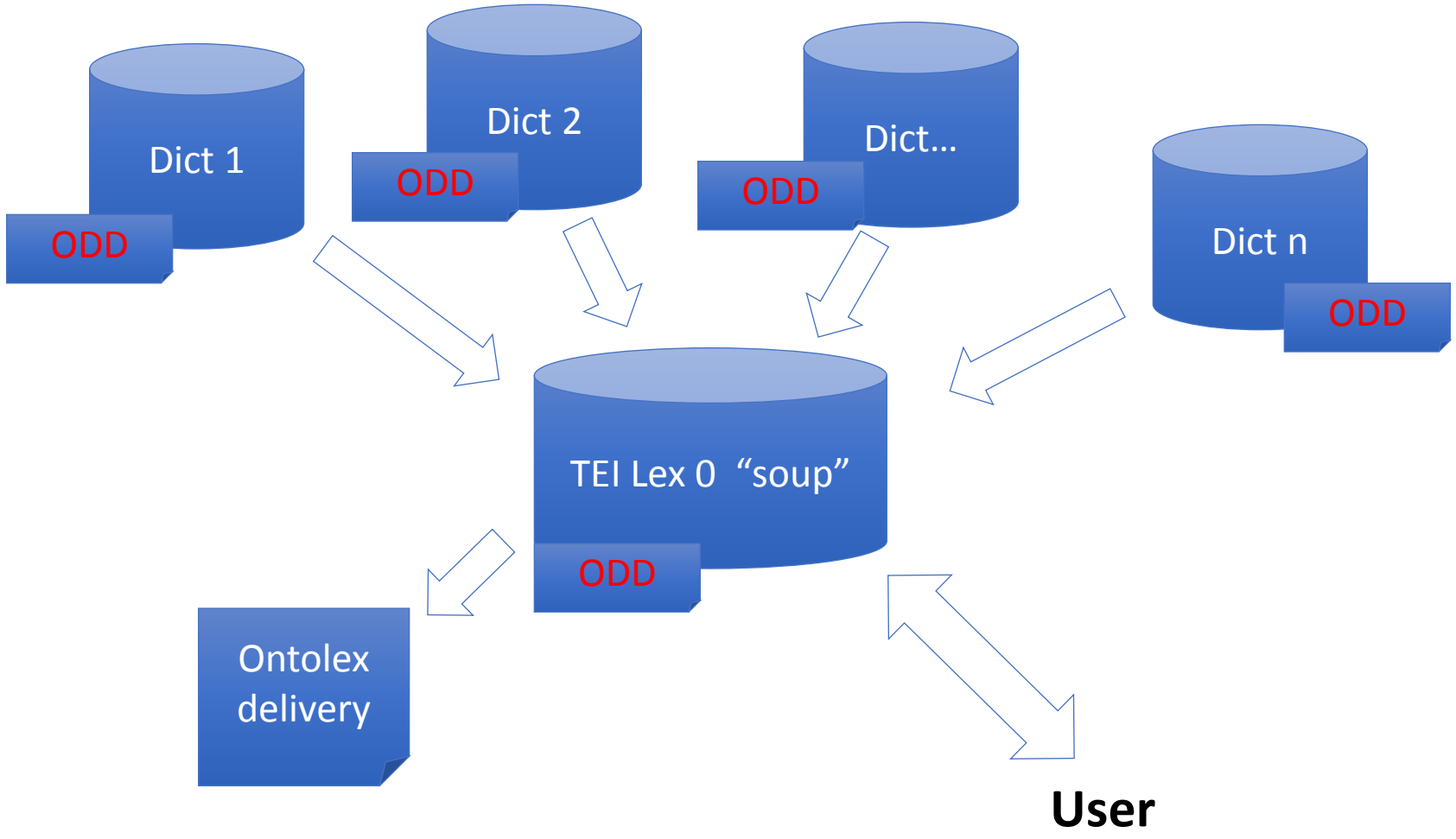
TEI lex 0: tightening the guidelines

- Initiative set up in the context of the DARIAH working group on lexical resources, supported by the EU project Elexis
- Objectives
  - Designing a target format for heterogeneous lexical data integration
- Trade-off
  - Compliance with the standard
  - Fine-tuning to the needs of a specific context/scenario
- Back to what a standard is: a common reference for a transaction
  - Perfect to say: “I am compliant to standard X except for…”
  - The TEI guidelines provides the means to carry out such customizations

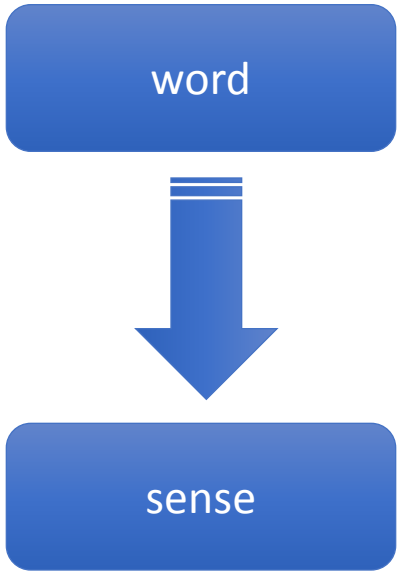
Application – the ELEXIS project

- European Lexicographic Infrastructure
  - 2018-02-01 – 2022-01-31
  - “integrate, extend and harmonise national and regional efforts in the field of lexicography”
    - Focuses on “efficient access”
  - Cooperation with CLARIN and DARIAH for long-term sustainability
- Lexical formats and standards in ELEXIS
  - Double sided approach TEI – Ontolex
  - TEI Lex 0 specification at the core of the data hub

The ELEXIS centralized hub



Enforcing the semasiological model




```
<entry>
        <form type="lemma">
                ...
        </form>
        <sense>
                <def>...</def>
        </sense>
</entry>
```

Simplifying the dictionary micro-structure

- Current situation
  - Containing vs. contained entries
    - <superEntry> – <entry> – <re>
  - Structured vs. unstructured entries
    - <entry> – <entryFree>
- The TEI Lex-0 vision
  - Representing all entry-like objects as <entry>
    - Making <entry> recursive
    - Making more use of <dictScrap>

Recursive entry - example

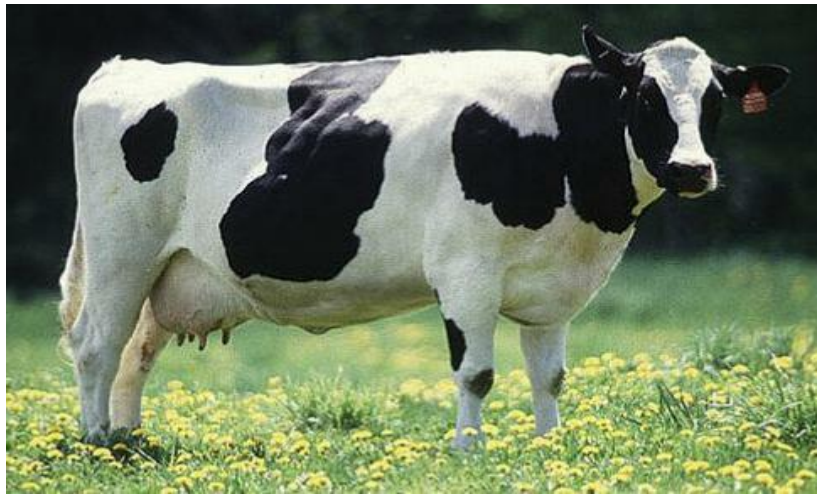
```
<entry type="wordFamily">
  <form type="base">
    <orth>Haus-</orth>
  </form>
  <pc>,</pc>
  <form type="base">
    <orth>haus-</orth>
  </form>
  <pc>:</pc>

  <!-- possibly some shared usg information -->
  <entry type="wordForm">
    <form type="lemma">
      <orth expand="Hausaltar">-altar</orth>
      <pc>,</pc>
      <gramGrp>
        <gen value="masculine">der</gen>
      </gramGrp>
    </form>
    <sense>...</sense>
  </entry>
  <entry type="wordForm">
    <form type="lemma">
      <orth expand="Hausandacht">-andacht</orth>
      <pc>,</pc>

    </form>

          <!-- ... -->
   </entry>
   <!-- ... -->
</entry>
```

Do we really have to encode all this manually?

Considering machine learning techniques

- Strong layout regularities within a given dictionary
- Similarities within a family of dictionaries
- Supervised – unsupervised ⇔ Less data – more data
- Features to be considered:
  - Layout
  - Lexical
- Sequencing task: well adapted for so-called graphical models

Why GROBID?



Cow (structured data)

publishing

text-mining

Hamburger (unstructured data)

"Converting PDF to XML is a bit like converting hamburgers into cows. You may be best off printing it and then scanning the result through a decent OCR package." Michael Kay (http://lists.xml.org/archives/xml-dev/200607/msg00509.html)

Inspired from: Duncan Hull

GROBID-Dictionaries

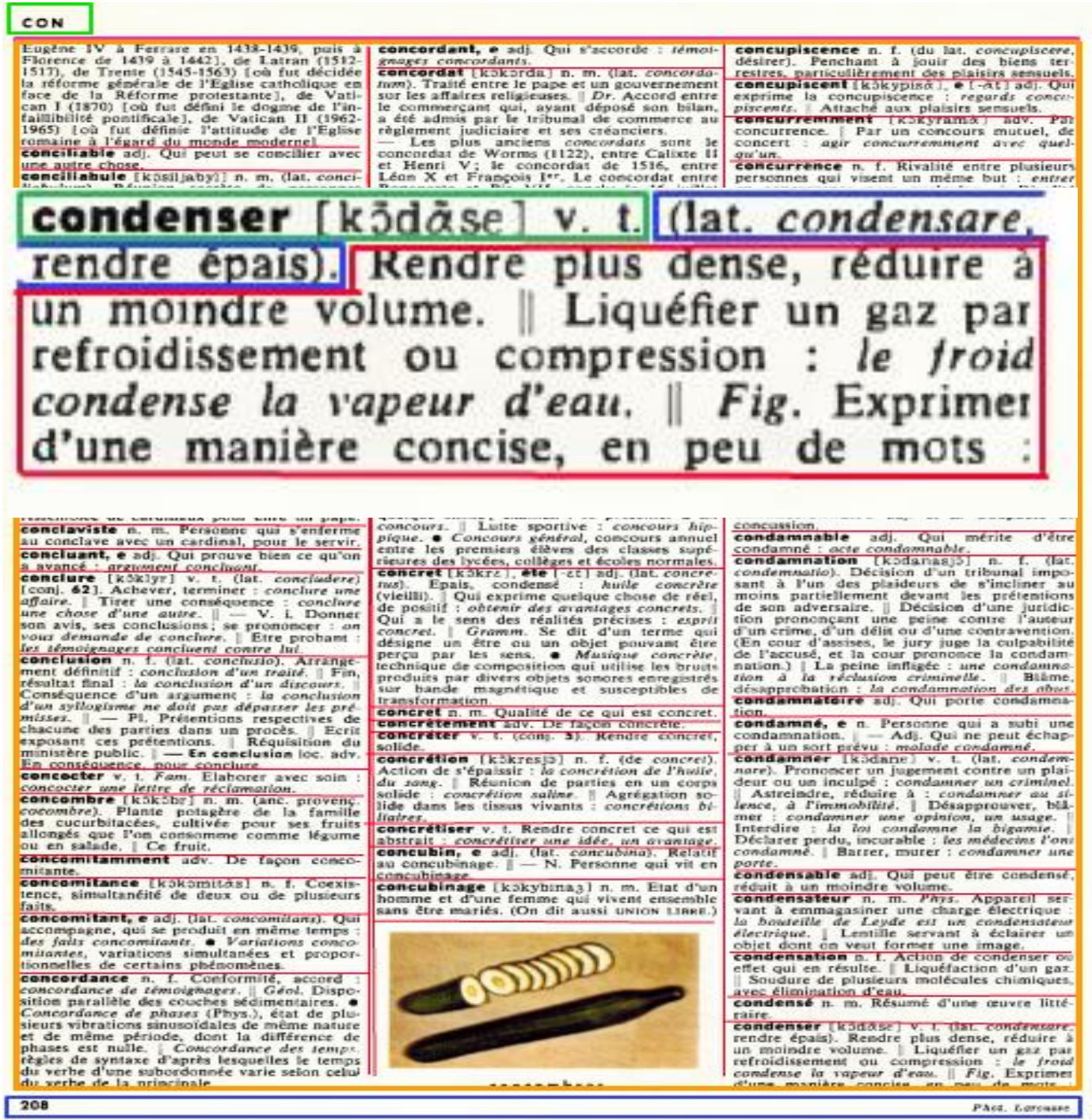
- Automatic extraction of TEI structures from digitised dictionaries (Khemakhem et al. 2017)
  - Input: PDF (soon ALTO)
  - Output: TEI compliant lexical resource
- Spin-off from GROBID (Romary and Lopez 2015)
  - Initiated in 2007
  - Automatic extraction of structural data from scholarly papers
    - Metadata (author, title, affiliations, keywords, abstract), bibliography, … full text
  - And open source…
- Uses Conditional Random Fields (CRF) (Lavergne et al. 2010)
  - Probabilistic models for sequence labelling tasks

State Of the Art

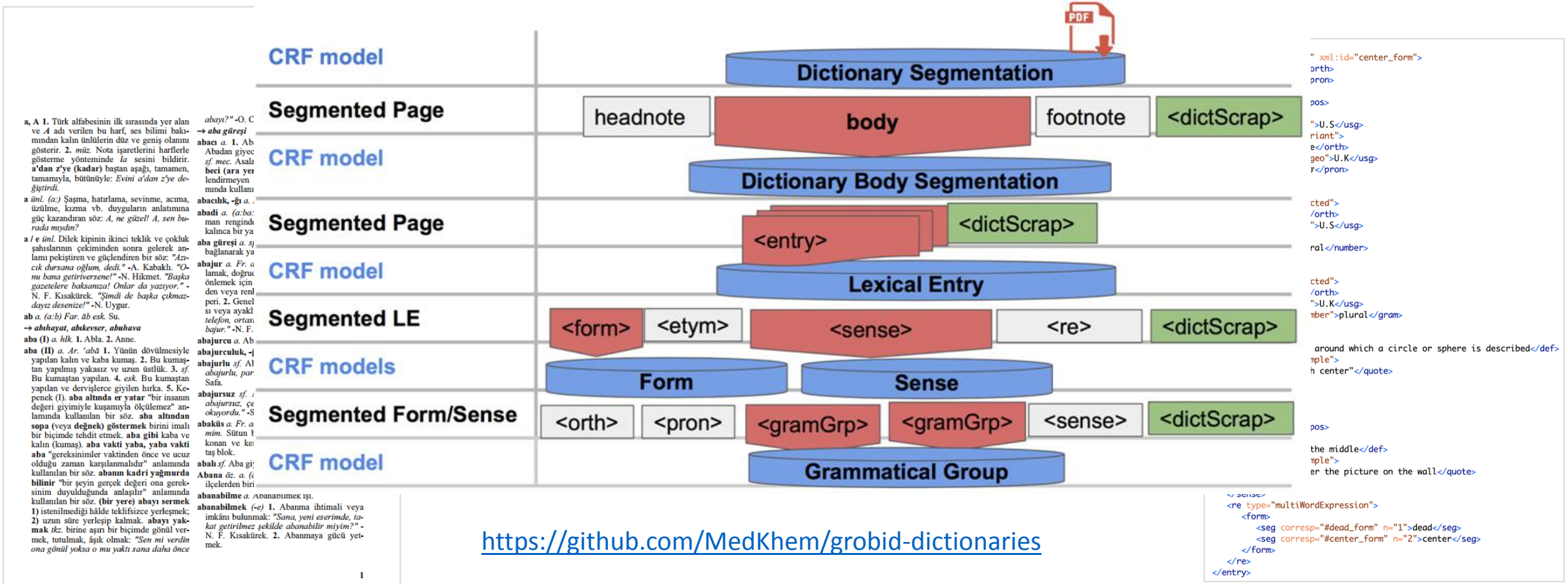
- Rule based approaches dominate
  - *(Khemakhem et al. 2009, Mykowiecka et al. 2012, Fayed et al. 2014)*

- Few machine learning attempts
  - Promote CRF for sequence labelling in dictionaries (*Crist 2011*)
  - Reduce the annotation time for labels (*Bago et al. 2015*)

Cascading CRF models

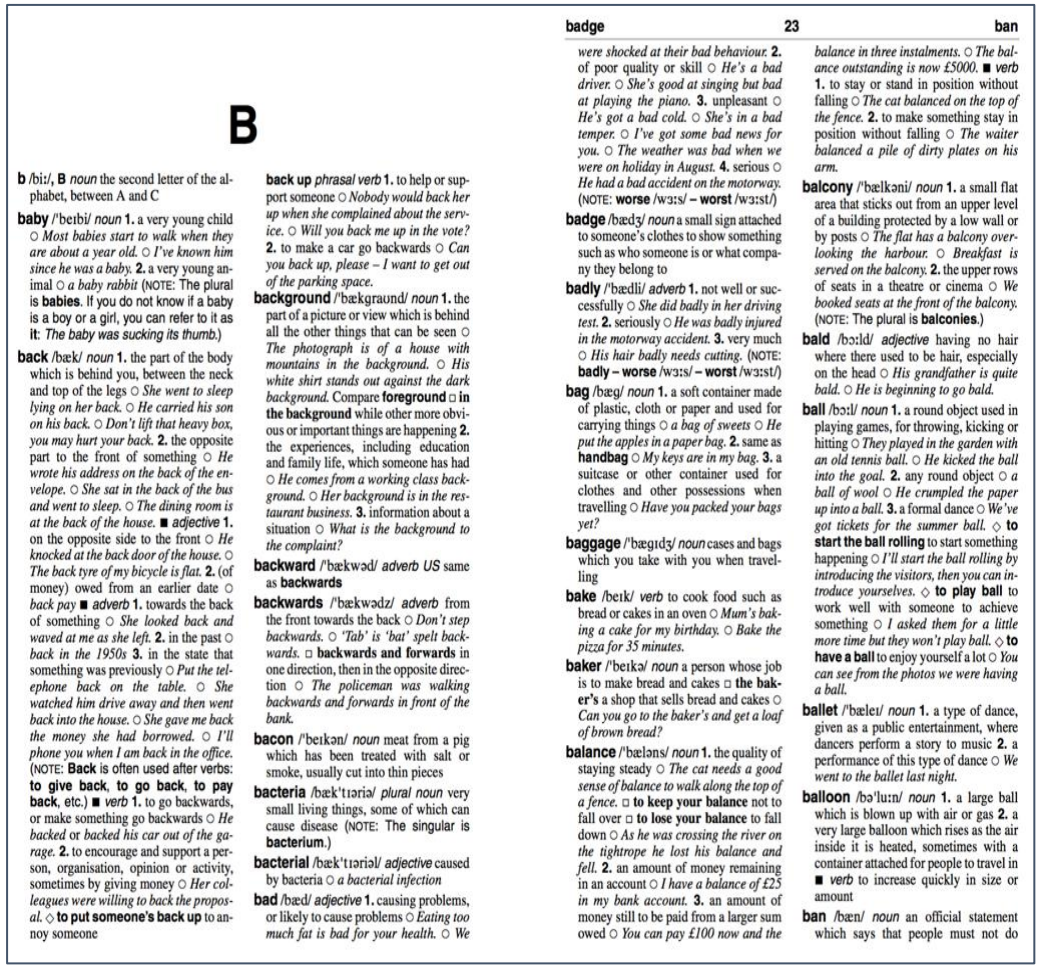


Cascading CRF models



https://github.com/MedKhem/grobid-dictionaries

First experiments



B

b /biː/, **B** *noun* the second letter of the alphabet, between A and C

baby /'beɪbi/ *noun* **1.** a very young child ○ *Most babies start to walk when they are about a year old.* ○ *I've known him since he was a baby.* **2.** a very young animal ○ *a baby rabbit* (NOTE: The plural is **babies**. If you do not know if a baby is a boy or a girl, you can refer to it as **it**: *The baby was sucking its thumb.*)

back /bæk/ *noun* **1.** the part of the body which is behind you, between the neck and top of the legs ○ *She went to sleep lying on her back.* ○ *He carried his son on his back.* ○ *Don't lift that heavy box, you may hurt your back.* **2.** the opposite part to the front of something ○ *He wrote his address on the back of the envelope.* ○ *She sat in the back of the bus and went to sleep.* ○ *The dining room is at the back of the house.* ■ *adjective* **1.** on the opposite side to the front ○ *He knocked at the back door of the house.* ○ *The back tyre of my bicycle is flat.* **2.** (of money) owed from an earlier date ○ *back pay* ■ *adverb* **1.** towards the back of something ○ *She looked back and waved at me as she left.* **2.** in the past ○ *back in the 1950s* **3.** in the state that something was previously ○ *Put the telephone back on the table.* ○ *She watched him drive away and then went back into the house.* ○ *She gave me back the money she had borrowed.* ○ *I'll phone you when I am back in the office.* (NOTE: **Back** is often used after verbs: **to give back**, **to go back**, **to pay back**, etc.) ■ *verb* **1.** to go backwards, or make something go backwards ○ *He backed or backed his car out of the garage.* **2.** to encourage and support a person, organisation, opinion or activity, sometimes by giving money ○ *Her colleagues were willing to back the proposal.* ◇ **to put someone's back up** to annoy someone

back up *phrasal verb* **1.** to help or support someone ○ *Nobody would back her up when she complained about the service.* ○ *Will you back me up in the vote?* **2.** to make a car go backwards ○ *Can you back up, please – I want to get out of the parking space.*

background /'bækgraʊnd/ *noun* **1.** the part of a picture or view which is behind all the other things that can be seen ○ *The photograph is of a house with mountains in the background.* ○ *His white shirt stands out against the dark background.* Compare **foreground** □ **in the background** while other more obvious or important things are happening **2.** the experiences, including education and family life, which someone has had ○ *He comes from a working class background.* ○ *Her background is in the restaurant business.* **3.** information about a situation ○ *What is the background to the complaint?*

backward /'bækwəd/ *adverb US* same as **backwards**

backwards /'bækwədz/ *adverb* from the front towards the back ○ *Don't step backwards.* ○ *'Tab' is 'bat' spelt backwards.* □ **backwards and forwards** in one direction, then in the opposite direction ○ *The policeman was walking backwards and forwards in front of the bank.*

bacon /'beɪkən/ *noun* meat from a pig which has been treated with salt or smoke, usually cut into thin pieces

bacteria /bæk'tɪəriə/ *plural noun* very small living things, some of which can cause disease (NOTE: The singular is **bacterium**.)

bacterial /bæk'tɪəriəl/ *adjective* caused by bacteria ○ *a bacterial infection*

bad /bæd/ *adjective* **1.** causing problems, or likely to cause problems ○ *Eating too much fat is bad for your health.* ○ *We were shocked at their bad behaviour.* **2.** of poor quality or skill ○ *He's a bad driver.* ○ *She's good at singing but bad at playing the piano.* **3.** unpleasant ○ *He's got a bad cold.* ○ *She's in a bad temper.* ○ *I've got some bad news for you.* ○ *The weather was bad when we were on holiday in August.* **4.** serious ○ *He had a bad accident on the motorway.* (NOTE: **worse** /wɜːs/ – **worst** /wɜːst/)

badge /bædʒ/ *noun* a small sign attached to someone's clothes to show something such as who someone is or what company they belong to

badly /'bædli/ *adverb* **1.** not well or successfully ○ *She did badly in her driving test.* **2.** seriously ○ *He was badly injured in the motorway accident.* **3.** very much ○ *His hair badly needs cutting.* (NOTE: **badly – worse** /wɜːs/ **– worst** /wɜːst/)

bag /bæg/ *noun* **1.** a soft container made of plastic, cloth or paper and used for carrying things ○ *a bag of sweets* ○ *He put the apples in a paper bag.* **2.** same as **handbag** ○ *My keys are in my bag.* **3.** a suitcase or other container used for clothes and other possessions when travelling ○ *Have you packed your bags yet?*

baggage /'bægɪdʒ/ *noun* cases and bags which you take with you when travelling

bake /beɪk/ *verb* to cook food such as bread or cakes in an oven ○ *Mum's baking a cake for my birthday.* ○ *Bake the pizza for 35 minutes.*

baker /'beɪkə/ *noun* a person whose job is to make bread and cakes □ **the baker's** a shop that sells bread and cakes ○ *Can you go to the baker's and get a loaf of brown bread?*

balance /'bæləns/ *noun* **1.** the quality of staying steady ○ *The cat needs a good sense of balance to walk along the top of a fence.* □ **to keep your balance** not to fall over □ **to lose your balance** to fall down ○ *As he was crossing the river on the tightrope he lost his balance and fell.* **2.** an amount of money remaining in an account ○ *I have a balance of £25 in my bank account.* **3.** an amount of money still to be paid from a larger sum owed ○ *You can pay £100 now and the balance in three instalments.* ○ *The balance outstanding is now £5000.* ■ *verb* **1.** to stay or stand in position without falling ○ *The cat balanced on the top of the fence.* **2.** to make something stay in position without falling ○ *The waiter balanced a pile of dirty plates on his arm.*

balcony /'bælkəni/ *noun* **1.** a small flat area that sticks out from an upper level of a building protected by a low wall or by posts ○ *The flat has a balcony overlooking the harbour.* ○ *Breakfast is served on the balcony.* **2.** the upper rows of seats in a theatre or cinema ○ *We booked seats at the front of the balcony.* (NOTE: The plural is **balconies**.)

bald /bɔːld/ *adjective* having no hair where there used to be hair, especially on the head ○ *His grandfather is quite bald.* ○ *He is beginning to go bald.*

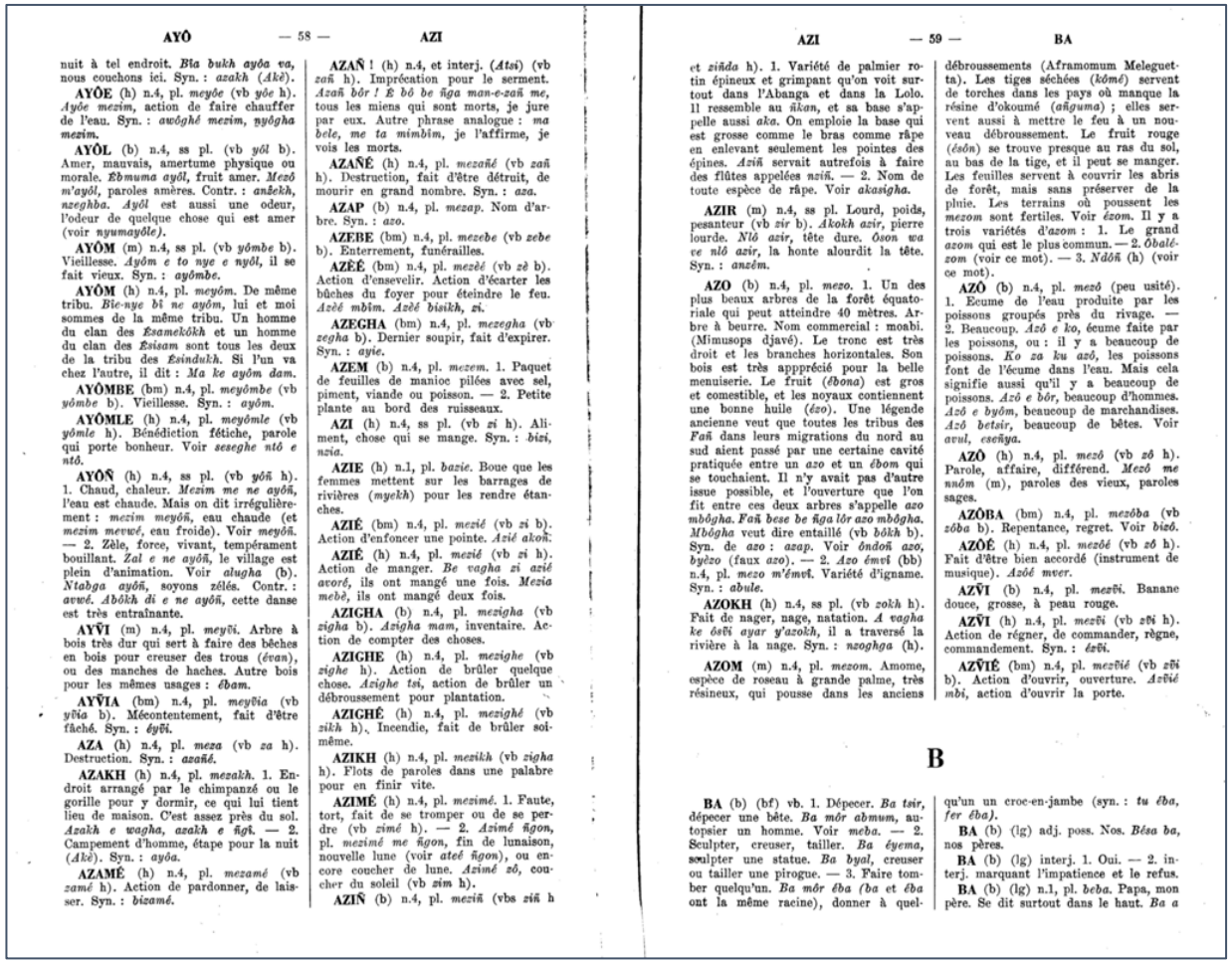
ball /bɔːl/ *noun* **1.** a round object used in playing games, for throwing, kicking or hitting ○ *They played in the garden with an old tennis ball.* ○ *He kicked the ball into the goal.* **2.** any round object ○ *a ball of wool* ○ *He crumpled the paper up into a ball.* **3.** a formal dance ○ *We've got tickets for the summer ball.* ◇ **to start the ball rolling** to start something happening ○ *I'll start the ball rolling by introducing the visitors, then you can introduce yourselves.* ◇ **to play ball** to work well with someone to achieve something ○ *I asked them for a little more time but they won't play ball.* ◇ **to have a ball** to enjoy yourself a lot ○ *You can see from the photos we were having a ball.*

ballet /'bæleɪ/ *noun* **1.** a type of dance, given as a public entertainment, where dancers perform a story to music **2.** a performance of this type of dance ○ *We went to the ballet last night.*

balloon /bə'luːn/ *noun* **1.** a large ball which is blown up with air or gas **2.** a very large balloon which rises as the air inside it is heated, sometimes with a container attached for people to travel in ■ *verb* to increase quickly in size or amount

ban /bæn/ *noun* an official statement which says that people must not do

Easier English Basic Dictionary (EEBD)



nuit à tel endroit. *Bla bukh ayóa va,* nous couchons ici. Syn. : *azakh* (*Akè*).

AYÔE (h) n.4, pl. *meyóe* (vb *yóe* h). *Ayóe mezim,* action de faire chauffer de l'eau. Syn. : *awóghé mezim, nyógha mezim.*

AYÔL (b) n.4, ss pl. (vb *yól* b). Amer, mauvais, amertume physique ou morale. *Ébmuma ayól,* fruit amer. *Mezó m'ayól,* paroles amères. Contr. : *anzekh, nzeghba. Ayól* est aussi une odeur, l'odeur de quelque chose qui est amer (voir *nyumayóle).*

AYÒM (m) n.4, ss pl. (vb *yómbe* b). Vieillesse. *Ayóm e to nye e nyól,* il se fait vieux. Syn. : *ayómbe.*

AYÒM (h) n.4, pl. *meyóm.* De même tribu. *Bie-nye bî ne ayóm,* lui et moi sommes de la même tribu. Un homme du clan des *Esamekókh* et un homme du clan des *Esisam* sont tous les deux de la tribu des *Esindukh.* Si l'un va chez l'autre, il dit : *Ma ke ayóm dam.*

AYÒMBE (bm) n.4, pl. *meyómbe* (vb *yómbe* b). Vieillesse. Syn. : *ayóm.*

AYÒMLE (h) n.4, pl. *meyómle* (vb *yómle* h). Bénédiction fétiche, parole qui porte bonheur. Voir *seseghe ntó e ntó.*

AYÓÑ (h) n.4, ss pl. (vb *yóñ* h). 1. Chaud, chaleur. *Mezim me ne ayóñ,* l'eau est chaude. Mais on dit irrégulièrement : *mezim meyóñ,* eau chaude (et *mezim mewé,* eau froide). Voir *meyóñ.* — 2. Zèle, force, vivant, tempérament bouillant. *Zal e ne ayóñ,* le village est plein d'animation. Voir *alugha* (b). *Ntabga ayóñ,* soyons zélés. Contr. : *avwé. Abókh dí e ne ayóñ,* cette danse est très entraînante.

AYÎ (m) n.4, pl. *meyî.* Arbre à bois très dur qui sert à faire des bêches en bois pour creuser des trous (*évan*), ou des manches de haches. Autre bois pour les mêmes usages : *ébam.*

AYÎA (bm) n.4, pl. *meyîa* (vb *yîa* b). Mécontentement, fait d'être fâché. Syn. : *éyîi.*

AZA (h) n.4, pl. *meza* (vb *za* h). Destruction. Syn. : *azañé.*

AZAKH (h) n.4, pl. *mezakh.* 1. Endroit arrangé par le chimpanzé ou le gorille pour y dormir, ce qui lui tient lieu de maison. C'est assez près du sol. *Azakh e wagha, azakh e ñgó.* — 2. Campement d'homme, étape pour la nuit (*Akè*). Syn. : *ayóa.*

AZAMÉ (h) n.4, pl. *mezamé* (vb *zamé* h). Action de pardonner, de laisser. Syn. : *bizamé.*

AZAÑ ! (h) n.4, et interj. (*Atsî*) (vb *zañ* h). Imprécation pour le serment. *Azañ bór ! É bô be ñga man-e-zañ me,* tous les miens qui sont morts, je jure par eux. Autre phrase analogue : *ma bele, me ta mimbîm,* je l'affirme, je vois les morts.

AZAÑÉ (h) n.4, pl. *mezañé* (vb *zañ* h). Destruction, fait d'être détruit, de mourir en grand nombre. Syn. : *aza.*

AZAP (b) n.4, pl. *mezap.* Nom d'arbre. Syn. : *azo.*

AZEBE (bm) n.4, pl. *mezebe* (vb *zebe* b). Enterrement, funérailles.

AZÈÈ (bm) n.4, pl. *mezèè* (vb *zè* b). Action d'ensevelir. Action d'écarter les bûches du foyer pour éteindre le feu. *Azèè mbîm. Azèè bisikh, zi.*

AZEGHA (bm) n.4, pl. *mezegha* (vb *zegha* b). Dernier soupir, fait d'expirer. Syn. : *ayie.*

AZEM (b) n.4, pl. *mezem.* 1. Paquet de feuilles de manioc pilées avec sel, piment, viande ou poisson. — 2. Petite plante au bord des ruisseaux.

AZI (h) n.4, ss pl. (vb *zi* h). Aliment, chose qui se mange. Syn. : *bizi, nzia.*

AZIE (h) n.1, pl. *bazie.* Boue que les femmes mettent sur les barrages de rivières (*myekh*) pour les rendre étanches.

AZIÉ (bm) n.4, pl. *mezié* (vb *zi* b). Action d'enfoncer une pointe. *Azié akoñ.*

AZIÉ (h) n.4, pl. *mezié* (vb *zi* h). Action de manger. *Be vagha zi azié avoré,* ils ont mangé une fois. *Mezia mebé,* ils ont mangé deux fois.

AZIGHA (b) n.4, pl. *mezigha* (vb *zigha* b). *Azigha mam,* inventaire. Action de compter des choses.

AZIGHE (h) n.4, pl. *mezighe* (vb *zighe* h). Action de brûler quelque chose. *Azighe tsi,* action de brûler un débroussement pour plantation.

AZIGHÉ (h) n.4, pl. *mezighé* (vb *zikh* h). Incendie, fait de brûler soi-même.

AZIKH (h) n.4, pl. *mezikh* (vb *zigha* h). Flots de paroles dans une palabre pour en finir vite.

AZIMÉ (h) n.4, pl. *mezimé.* 1. Faute, tort, fait de se tromper ou de se perdre (vb *zimé* h). — 2. *Azimé ñgon,* pl. *mezimé me ñgon,* fin de lunaison, nouvelle lune (voir *ateé ñgon*), ou encore coucher de lune. *Azimé zó,* coucher du soleil (vb *zim* h).

AZIÑ (b) n.4, pl. *meziñ* (vbs *ziñ* h

et *ziñda* h). 1. Variété de palmier rotin épineux et grimpant qu'on voit surtout dans l'Abanga et dans la Lolo. Il ressemble au *ákan,* et sa base s'appelle aussi *aka.* On emploie la base qui est grosse comme le bras comme râpe en enlevant seulement les pointes des épines. *Aziñ* servait autrefois à faire des flûtes appelées *nziñ.* — 2. Nom de toute espèce de râpe. Voir *akasigha.*

AZIR (m) n.4, ss pl. Lourd, poids, pesanteur (vb *zir* b). *Akokh azir,* pierre lourde. *Nlô azir,* tête dure. *Ósan wa ve nlô azir,* la honte alourdit la tête. Syn. : *anzém.*

AZO (b) n.4, pl. *mezo.* 1. Un des plus beaux arbres de la forêt équatoriale qui peut atteindre 40 mètres. Arbre à beurre. Nom commercial : moabi. (Mimusops djavé). Le tronc est très droit et les branches horizontales. Son bois est très apprécié pour la belle menuiserie. Le fruit (*ébona*) est gros et comestible, et les noyaux contiennent une bonne huile (*ézo*). Une légende ancienne veut que toutes les tribus des *Fañ* dans leurs migrations du nord au sud aient passé par une certaine cavité pratiquée entre un *azo* et un *ébom* qui se touchaient. Il n'y avait pas d'autre issue possible, et l'ouverture que l'on fit entre ces deux arbres s'appelle *azo mbógha. Fañ bese be ñga lór azo mbógha. Mbógha* veut dire entaillé (vb *bókh* b). Syn. de *azo* : *asap.* Voir *óndoñ azo, byézo* (faux *azo*). — 2. *Azo émvî* (bb) n.4, pl. *mezo m'émvî.* Variété d'igname. Syn. : *abule.*

AZOKH (h) n.4, ss pl. (vb *zokh* h). Fait de nager, nage, natation. *A vagha ke ósîi ayar y'azokh,* il a traversé la rivière à la nage. Syn. : *nzoghga* (h).

AZOM (m) n.4, pl. *mezom.* Amome, espèce de roseau à grande palme, très résineux, qui pousse dans les anciens débroussements (Aframomum Meleguetta). Les tiges séchées (*kómé*) servent de torches dans les pays où manque la résine d'okoumé (*añguma*) ; elles servent aussi à mettre le feu à un nouveau débroussement. Le fruit rouge (*ésón*) se trouve presque au ras du sol, au bas de la tige, et il peut se manger. Les feuilles servent à couvrir les abris de forêt, mais sans préserver de la pluie. Les terrains où poussent les *mezom* sont fertiles. Voir *ézom.* Il y a trois variétés d'*azom* : 1. Le grand *azom* qui est le plus commun. — 2. *Óbalézom* (voir ce mot). — 3. *Ndôñ* (h) (voir ce mot).

AZÓ (b) n.4, pl. *mezó* (peu usité). 1. Écume de l'eau produite par les poissons groupés près du rivage. — 2. Beaucoup. *Azó e ko,* écume faite par les poissons, ou : il y a beaucoup de poissons. *Ko za ku azó,* les poissons font de l'écume dans l'eau. Mais cela signifie aussi qu'il y a beaucoup de poissons. *Azó e bór,* beaucoup d'hommes. *Azó e byóm,* beaucoup de marchandises. *Azó betsir,* beaucoup de bêtes. Voir *avul, cseñya.*

AZÒ (h) n.4, pl. *mezó* (vb *zó* h). Parole, affaire, différend. *Mezó me nnóm* (m), paroles des vieux, paroles sages.

AZÓBA (bm) n.4, pl. *mezóba* (vb *zóba* b). Repentance, regret. Voir *bizó.*

AZÓÉ (h) n.4, pl. *mezóé* (vb *zó* h). Fait d'être bien accordé (instrument de musique). *Azóé mver.*

AZÝI (b) n.4, pl. *mezîi.* Banane douce, grosse, à peau rouge.

AZÝI (h) n.4, pl. *mezîi* (vb *zîi* h). Action de régner, de commander, règne, commandement. Syn. : *ézîi.*

AZÝIÉ (bm) n.4, pl. *mezîié* (vb *zîi* b). Action d'ouvrir, ouverture. *Azîié mbi,* action d'ouvrir la porte.

B

BA (b) (bf) vb. 1. Dépecer. *Ba tsir,* dépecer une bête. *Ba mór abmum,* autopsier un homme. Voir *meba.* — 2. Sculpter, creuser, tailler. *Ba éyema,* sculpter une statue. *Ba byal,* creuser ou tailler une pirogue. — 3. Faire tomber quelqu'un. *Ba mór éba (ba* et *éba* ont la même racine), donner à quelqu'un un croc-en-jambe (syn. : *tu éba, fer éba).*

BA (b) (lg) adj. poss. Nos. *Bésa ba,* nos pères.

BA (b) (lg) interj. 1. Oui. — 2. interj. marquant l'impatience et le refus.

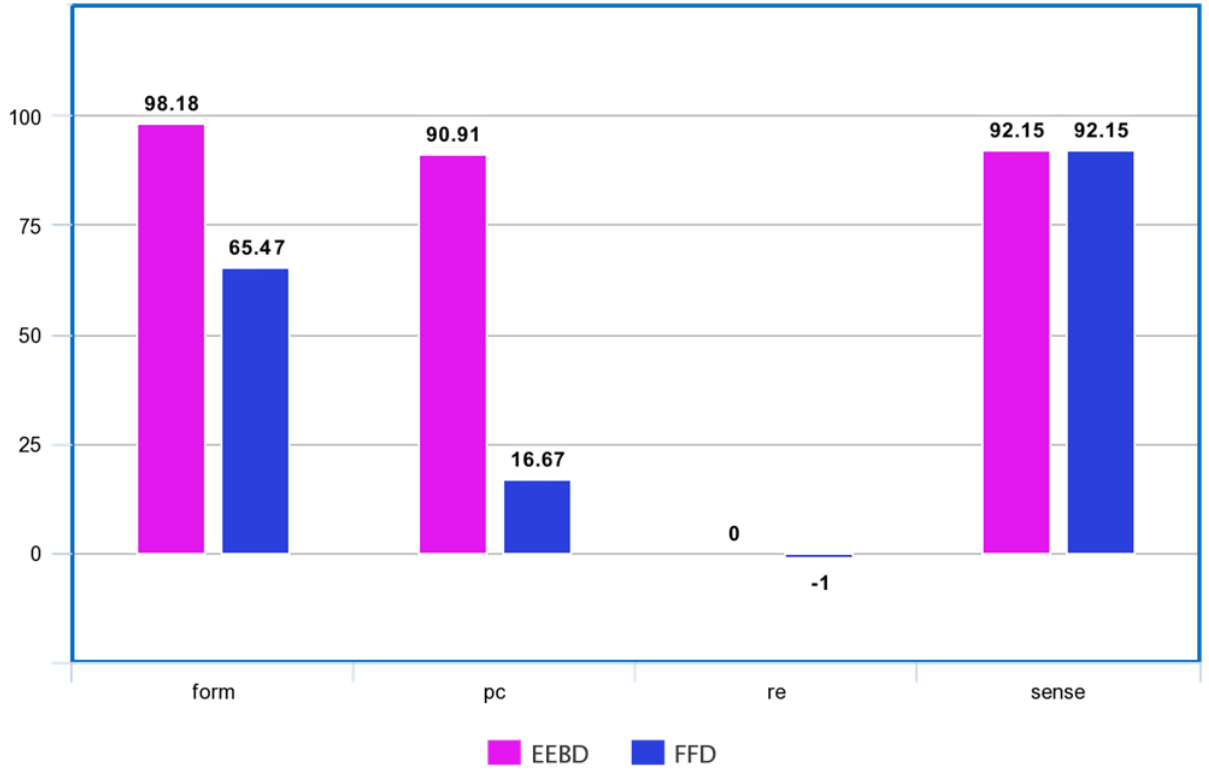
BA (b) (lg) n.1, pl. *beba.* Papa, mon père. Se dit surtout dans le haut. *Ba a*

Fang-French Dictionary (FFD)

Evaluation: Token Level - F1 Score

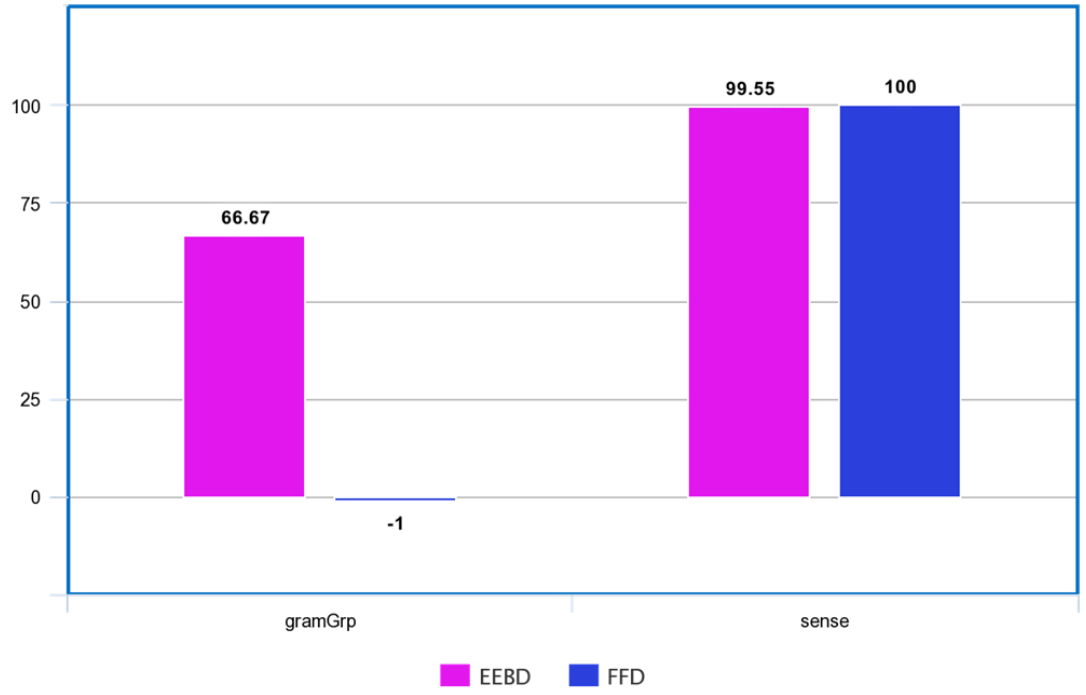
Lexical Entry

- EEBD: 100 LE (8 pages): 76 training, 24 evaluation
- FFD: 71 LE (3 pages): 47 training, 24 evaluation

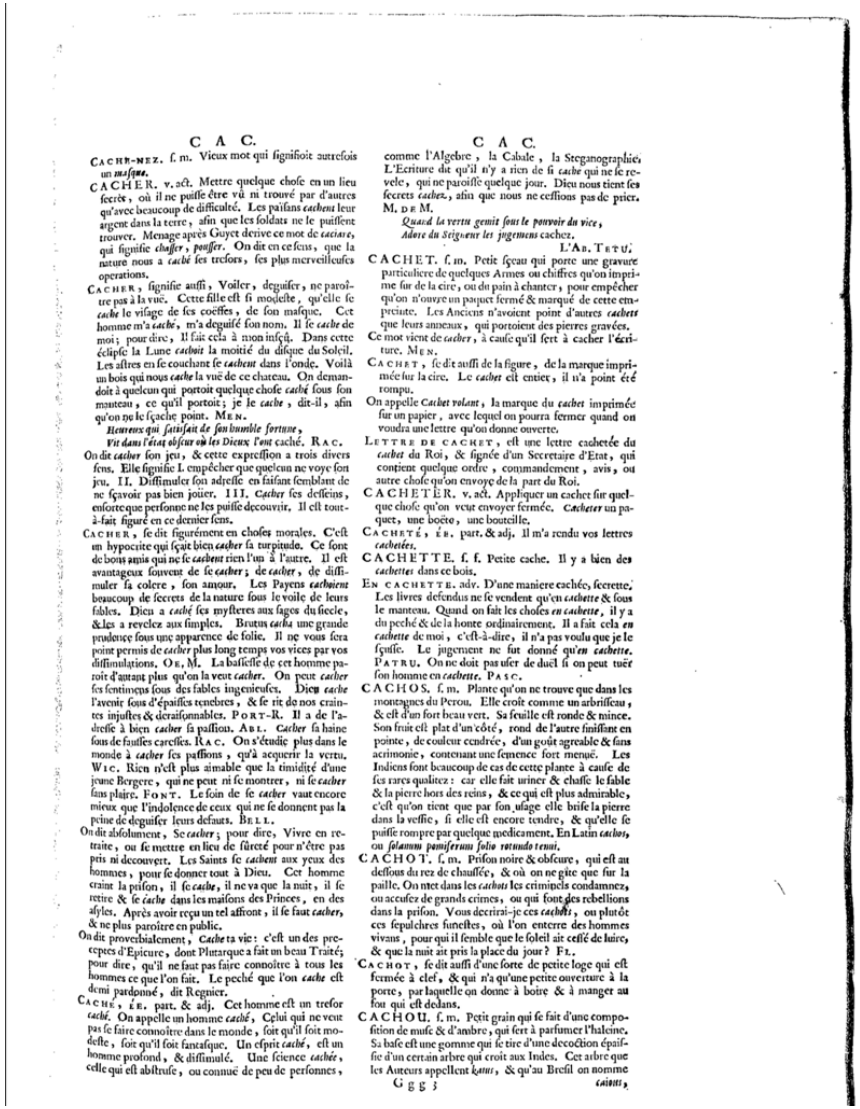


Sense

- EEBD: 30 blocks (6 pages): 15 training, 15 evaluation
- FFD: 90 LE (4 pages): 71 training, 19 evaluation



Further complexities



C A C.

CACHE-NEZ. ſ. m. Vieux mot qui ſignifioit autrefois un *maſque*.

CACHER. v. act. Mettre quelque choſe en un lieu ſecrète, où il ne puiſſe être vû ni trouvé par d'autres qu'avec beaucoup de difficulté. Les paiſans *cachent* leur argent dans la terre, afin que les ſoldats ne le puiſſent trouver. Menage après Guyet derive ce mot de *cacisare*, qui ſignifie *chaſſer, pouſſer*. On dit en ce ſens, que la nature nous a *caché* les treſors, ſes plus merveilleuſes operations.

CACHER, ſignifie auſſi, Voiler, deguiſer, ne paroître pas à la vuë. Cette fille eſt ſi modeſte, qu'elle ſe *cache* le viſage de ſes coëffes, de ſon maſque. Cet homme m'a *caché*, m'a deguiſé ſon nom. Il ſe *cache* de moi; pour dire, Il fait cela à mon inſçû. Dans cette éclipſe la Lune *cachoit* la moitié du diſque du Soleil. Les aſtres en ſe couchant ſe *cachent* dans l'onde. Voilà un bois qui nous *cache* la vuë de ce chateau. On demandoit à quelcun qui portoit quelque choſe *caché* ſous ſon manteau, ce qu'il portoit; je le *cache*, dit-il, afin qu'on ne le ſçache point. Men.

Heureux qui ſatisfait de ſon humble fortune,
Vit dans l'état obſcur où les Dieux l'ont caché. Rac.

On dit *cacher* ſon jeu, & cette expreſſion a trois divers ſens. Elle ſignifie I. empêcher que quelcun ne voye ſon jeu. II. Diſſimuler ſon adreſſe en faiſant ſemblant de ne ſçavoir pas bien jouer. III. *Cacher* ſes deſſeins, enſorte que perſonne ne les puiſſe découvrir. Il eſt tout-à-fait figuré en ce dernier ſens.

CACHER, ſe dit figurément en choſes morales. C'eſt un hypocrite qui ſçait bien *cacher* ſa turpitude. Ce ſont de bons amis qui ne ſe *cachent* rien l'un à l'autre. Il eſt avantageux ſouvent de ſe *cacher*; de *cacher*, de diſſimuler ſa colere, ſon amour. Les Payens *cachoient* beaucoup de ſecrets de la nature ſous le voile de leurs fables. Dieu a *caché* ſes myſteres aux ſages du ſiecle, & les a revelez aux ſimples. Brutus *cacha* une grande prudence ſous une apparence de folie. Il ne vous ſera point permis de *cacher* plus long temps vos vices par vos diſſimulations. Oe, M. La baſſeſſe de cet homme paroit d'autant plus qu'on la veut *cacher*. On peut *cacher* ſes ſentimens ſous des fables ingenieuſes. Dieu *cache* l'avenir ſous d'épaiſſes tenebres, & ſe rit de nos craintes injuſtes & deraiſonnables. Port-R. Il a de l'adreſſe à bien *cacher* ſa paſſion. Abl. *Cacher* ſa haine ſous de fauſſes careſſes. Rac. On s'étudie plus dans le monde à *cacher* ſes paſſions, qu'à acquerir la vertu. Wic. Rien n'eſt plus aimable que la timidité d'une jeune Bergere, qui ne peut ni ſe montrer, ni ſe *cacher* ſans plaire. Font. Le ſoin de ſe *cacher* vaut encore mieux que l'indolence de ceux qui ne ſe donnent pas la peine de deguiſer leurs defauts. Bell.

On dit abſolument, Se *cacher*; pour dire, Vivre en retraite, ou ſe mettre en lieu de ſûreté pour n'être pas pris ni découvert. Les Saints ſe *cachent* aux yeux des hommes, pour ſe donner tout à Dieu. Cet homme craint la priſon, il ſe *cache*, il ne va que la nuit, il ſe retire & ſe *cache* dans les maiſons des Princes, en des aſyles. Après avoir reçu un tel affront, il ſe faut *cacher*, & ne plus paroître en public.

On dit proverbialement, *Cache ta vie*: c'eſt un des preceptes d'Epicure, dont Plutarque a fait un beau Traité; pour dire, qu'il ne faut pas faire connoître à tous les hommes ce que l'on fait. Le peché que l'on *cache* eſt demi pardonné, dit Regnier.

CACHÉ, ÉE. part. & adj. Cet homme eſt un treſor *caché*. On appelle un homme *caché*, Celui qui ne veut pas ſe faire connoître dans le monde, ſoit qu'il ſoit modeſte, ſoit qu'il ſoit fantaſque. Un eſprit *caché*, eſt un homme profond, & diſſimulé. Une ſcience *cachée*, celle qui eſt abſtruſe, ou connuë de peu de perſonnes,

C A C.

comme l'Algebre, la Cabale, la Steganographie, L'Ecriture dit qu'il n'y a rien de ſi *caché* qui ne ſe revele, qui ne paroiſſe quelque jour. Dieu nous tient ſes ſecrets *cachez*, afin que nous ne ceſſions pas de prier. M. de M.

Quand la vertu gemit ſous le pouvoir du vice,
Adore du Seigneur les jugemens cachez.
L'Ab. Tetu.

CACHET. ſ. m. Petit ſçeau qui porte une gravure particuliere de quelques Armes ou chiffres qu'on imprime ſur de la cire, ou du pain à chanter, pour empêcher qu'on n'ouvre un paquet fermé & marqué de cette empreinte. Les Anciens n'avoient point d'autres *cachets* que leurs anneaux, qui portoient des pierres gravées. Ce mot vient de *cacher*, à cauſe qu'il ſert à cacher l'écriture. Men.

CACHET, ſe dit auſſi de la figure, de la marque imprimée ſur la cire. Le *cachet* eſt entier, il n'a point été rompu.

On appelle *Cachet volant*, la marque du *cachet* imprimée ſur un papier, avec lequel on pourra fermer quand on voudra une lettre qu'on donne ouverte.

Lettre de cachet, eſt une lettre cachetée du *cachet* du Roi, & ſignée d'un Secretaire d'Etat, qui contient quelque ordre, commandement, avis, ou autre choſe qu'on envoye de la part du Roi.

CACHETER. v. act. Appliquer un cachet ſur quelque choſe qu'on veut envoyer fermée. *Cacheter* un paquet, une boëte, une bouteille.

CACHETÉ, ÉE. part. & adj. Il m'a rendu vos lettres *cachetées*.

CACHETTE. ſ. f. Petite cache. Il y a bien des *cachettes* dans ce bois.

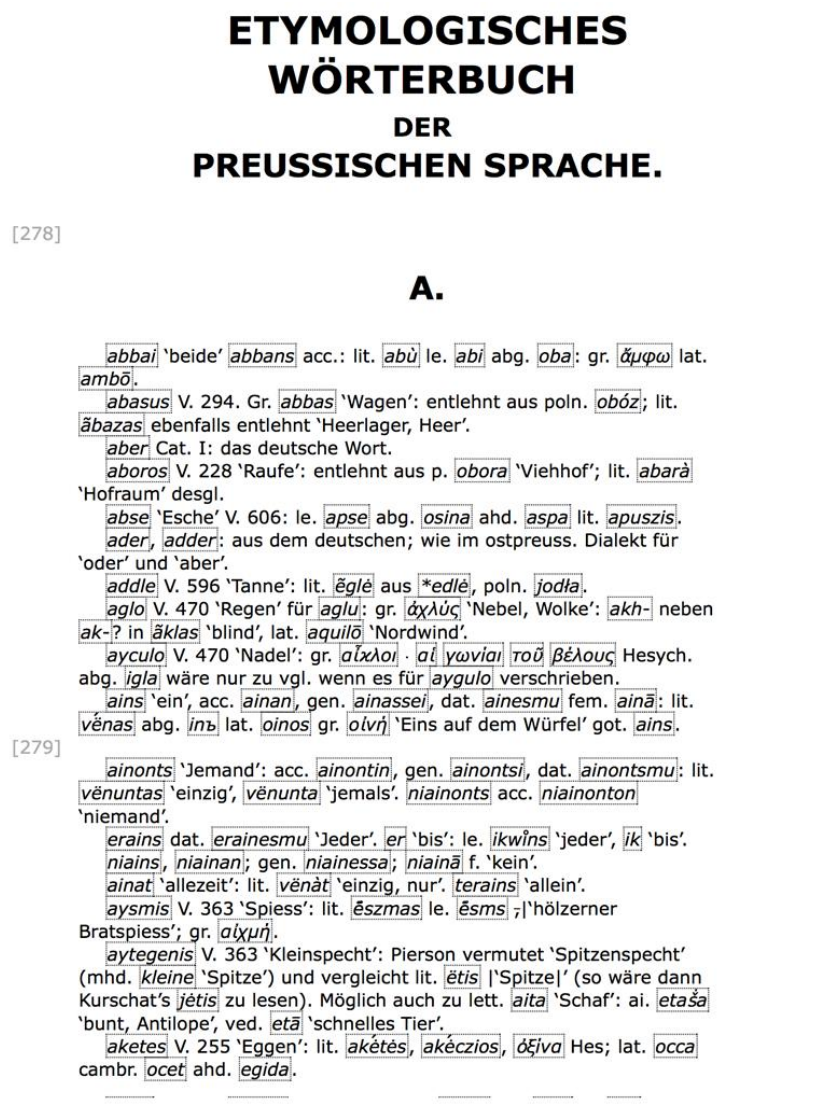
En cachette. adv. D'une maniere cachée, ſecrette. Les livres defendus ne ſe vendent qu'en *cachette* & ſous le manteau. Quand on fait les choſes *en cachette*, il y a du peché & de la honte ordinairement. Il a fait cela *en cachette* de moi, c'eſt-à-dire, il n'a pas voulu que je le ſçuſſe. Le jugement ne fut donné qu'*en cachette*. Patru. On ne doit pas uſer de duel ſi on peut tuër ſon homme en *cachette*. Pasc.

CACHOS. ſ. m. Plante qu'on ne trouve que dans les montagnes du Perou. Elle croit comme un arbriſſeau, & eſt d'un fort beau vert. Sa feuille eſt ronde & mince. Son fruit eſt plat d'un côté, rond de l'autre finiſſant en pointe, de couleur cendrée, d'un goût agreable & ſans acrimonie, contenant une ſemence fort menuë. Les Indiens font beaucoup de cas de cette plante à cauſe de ſes rares qualitez: car elle fait uriner & chaſſe le ſable & la pierre hors des reins, & ce qui eſt plus admirable, c'eſt qu'on tient que par ſon uſage elle briſe la pierre dans la veſſie, ſi elle eſt encore tendre, & qu'elle ſe puiſſe rompre par quelque medicament. En Latin *cachos*, ou *ſolanum pomiferum folio rotundo tenui*.

CACHOT. ſ. m. Priſon noire & obſcure, qui eſt au deſſous du rez de chauſſée, & où on ne gite que ſur la paille. On met dans les *cachots* les criminels condamnez, ou accuſez de grands crimes, ou qui font des rebellions dans la priſon. Vous decrirai-je ces *cachots*, ou plutôt ces ſepulchres funeſtes, où l'on enterre des hommes vivans, pour qui il ſemble que le ſoleil ait ceſſé de luire, & que la nuit ait pris la place du jour? Fl.

Cachot, ſe dit auſſi d'une ſorte de petite loge qui eſt fermée à clef, & qui n'a qu'une petite ouverture à la porte, par laquelle on donne à boire & à manger au fou qui eſt dedans.

CACHOU. ſ. m. Petit grain qui ſe fait d'une compoſition de muſc & d'ambre, qui ſert à parfumer l'haleine. Sa baſe eſt une gomme qui ſe tire d'une decoction épaiſſie d'un certain arbre qui croit aux Indes. Cet arbre que les Auteurs appellent *katus*, & qu'au Breſil on nomme *cainetti*,



ETYMOLOGISCHES WÖRTERBUCH DER PREUSSISCHEN SPRACHE.

[278]

A.

abbai 'beide' *abbans* acc.: lit. *abù* le. *abi* abg. *oba*: gr. *ἄμφω* lat. *ambō*.

abasus V. 294. Gr. *abbas* 'Wagen': entlehnt aus poln. *obóz*; lit. *ābazas* ebenfalls entlehnt 'Heerlager, Heer'.

aber Cat. I: das deutsche Wort.

aboros V. 228 'Raufe': entlehnt aus p. *obora* 'Viehhof'; lit. *abarà* 'Hofraum' desgl.

abse 'Esche' V. 606: le. *apse* abg. *osina* ahd. *aspa* lit. *apuszis*.

ader, *adder*: aus dem deutschen; wie im ostpreuss. Dialekt für 'oder' und 'aber'.

addle V. 596 'Tanne': lit. *ẽglė* aus **edlė*, poln. *jodła*.

aglo V. 470 'Regen' für *aglu*: gr. *ἀχλύς* 'Nebel, Wolke': *akh-* neben *ak-*? in *ãklas* 'blind', lat. *aquilō* 'Nordwind'.

ayculo V. 470 'Nadel': gr. *αἶχλοι · αἱ γωνίαι τοῦ βέλους* Hesych. abg. *igla* wäre nur zu vgl. wenn es für *aygulo* verschrieben.

ains 'ein', acc. *ainan*, gen. *ainassei*, dat. *ainesmu* fem. *ainā*: lit. *vẽnas* abg. *inъ* lat. *oinos* gr. *οἰνή* 'Eins auf dem Würfel' got. *ains*.

[279]

ainonts 'Jemand': acc. *ainontin*, gen. *ainontsi*, dat. *ainontsmu*: lit. *vënuntas* 'einzig', *vënunta* 'jemals'. *niainonts* acc. *niainonton* 'niemand'.

erains dat. *erainesmu* 'Jeder'. *er* 'bis': le. *ikwĩns* 'jeder', *ik* 'bis'.

niains, *niainan*; gen. *niainessa*; *niainā* f. 'kein'.

ainat 'allezeit': lit. *vënàt* 'einzig, nur'. *terains* 'allein'.

aysmis V. 363 'Spiess': lit. *ẽszmas* le. *ẽsms* ,|'hölzerner Bratspiess'; gr. *αἰχμή*.

aytegenis V. 363 'Kleinspecht': Pierson vermutet 'Spitzenspecht' (mhd. *kleine* 'Spitze') und vergleicht lit. *ëtis* |'Spitze|' (so wäre dann Kurschat's *jėtis* zu lesen). Möglich auch zu lett. *aita* 'Schaf': ai. *etaša* 'bunt, Antilope', ved. *etā* 'schnelles Tier'.

aketes V. 255 'Eggen': lit. *akėtės*, *akėczios*, *ὀξίνα* Hes; lat. *occa* cambr. *ocet* ahd. *egida*.

parceque ſa chair eſt plus ferme que celle des autres pêches.

DURANDAL. ſ. m. eſt le nom de l'épée de Roland Chevalier Heros de l'Arioſte. On s'en ſert en cette phraſe proverbiale: pour expliquer qu'une viande eſt fort dure, on dit que c'eſt *durandal*, l'épée de Roland.

DURANT. Prepoſition. Pendant, tandis qu'une choſe ſubſiſtera. *Durant* qu'on eſt dans l'emploi il faut faire ſa fortune. Il faut faire ſes proviſions *durant* l'été. N'ai-je pas ſçu que *durant* vôtre voyage, vous avez été de la plus belle humeur du monde? LET. PORTUG. Si jamais la voye du Chretien eſt étroite, c'eſt durant les perſecutions. FL. Ce mot ſe met quelquefois après le nom qu'il regit. J'ay été malade ſix ans *durant*. On lui a aſſigné une penſion ſa vie *durant*.

DURCIR. v. act. & n. Rendre dur. On *durcit* le fer à force de le battre. Le ſoleil *durcit* l'ambre, *durcit* les perles. L'air *durcit* le corail. Un œuf trop cuit ſe *durcit*. La viande *durcit* pendant la gelée.

DURCIR, ſe dit auſſi figurément de l'eſprit, & ſignifie, Rendre ferme. Cela *durcit* l'eſprit.

DURCI, IE. part.

DURE. ſ. f. On ne le dit qu'en cette phraſe: Coucher ſur la *dure*; c'eſt-à-dire, ſur la terre, ou ſans matelas.

DURE'E. ſ. f. Perſeverance des choſes dans leur être; temps meſuré par la ſubſiſtance de quelque choſe. Le temps eſt defini par les Philoſophes, La *durée* d'un mouvement. Dieu a promis à ſes éleus une gloire d'éternelle *durée*. Cette fougue eſt trop violente, elle ne ſera pas de *durée*. Nous ne jouïſſons de la vie qu'à meſure que nous la perdons: chaque moment en abrege la *durée*. MORALE DE P. On juge de la *durée* du temps ſelon la diſpoſition où l'on ſe trouve: celui qui eſt accablé de triſteſſe s'ennuye de la *durée* du temps, parcequ'elle lui eſt penible, & qu'il y fait plus d'attention. MALEB. La *durée* des heures, au regard de l'ennui, & du chagrin, ſe fait plus ſentir que celle des années. BOU. Les Dieux ne ſont immortels que par la *durée* de leurs plaiſirs. DAC. Les paſſions veulent être conduites avec art pour en étendre la *durée*, afin qu'elles ne s'épuiſent pas trop tôt. LE CH. D'H. La *durée* de nos paſſions ne depend pas plus de nous, que la *durée* de nôtre vie. LA ROCH. Je ne meſure pas ma vie par la *durée* du temps; mais par la *durée* de la gloire. BOU. Les Dames pour l'ordinaire trouvent leurs maris de longue *durée*. LE CH. DE M. Cette femme s'eſt mis dans l'eſprit d'égaler la *durée* de ſon deuil à celle de ſa vie, & a choiſi cette triſte, & fatigante voye pour acquerir de la reputation. M. ESP.

Il n'eſt rien ici bas d'éternelle durée. MALH.
Cette tendre amitié par tant de fois jurée,
Qui devoit ſurpaſſer les ſiecles en durée,
A la fin s'eſt éteinte. VOI.

DUREMENT, adv. D'une maniere dure. Il a été traité *durement* par ſon Maître, &c. Ces Religieux ſont couchez bien *durement*. Luther s'eſt exprimé *durement* en parlant de la predeſtination. CL. Il ne faut pas dire *durement* les choſes dures. NIC.

DURE-MERE. ſ. f. Terme d'Anatomie. C'eſt la membrane qui enveloppe le cerveau. Membrane du cerveau groſſe & dure, qui eſt attachée à l'os du crâne.

DURER. v. n. Subſiſter pendant quelque eſpace de temps. Une femme ſe defait de ſon galant quand elle veut; mais il faut qu'elle garde ſon mari tant qu'il *dure*. LE CH. DE M. L'abſence, pour peu qu'elle *dure*, nuit à l'amitié auſſi bien qu'à l'amour. ID. Rien n'approche de l'ennui que donne une paſſion qui *dure* trop.

ST. EV. Un engagement qui doit *durer* juſqu'à la mort, ne ſe doit jamais faire qu'avec de grandes precautions. MOL. L'amour *duroit* un monde au bon vieux temps. MAR. Le monde a déjà *duré* cinq à ſix mille ans. Ce que Malherbe écrit *dure* éternellement. MALH.

DURER, ſe dit auſſi de ce qui eſt ſolide; qui ſubſiſte long temps; qui eſt fort; qui s'uſe difficilement. Le drap d'Eſpagne eſt d'un bon uſer, il *dure* long temps. Ce meuble *durera* un ſiecle, cela *durera* juſqu'au bout.

DURER, avec la negative, ſignifie, Reſiſter, ſouffrir quelque mal, quelque peine, quelque incommodité. On ne ſçauroit *durer* avec cette femme-là, tant elle eſt criarde. Je ne puis plus *durer* avec cette colique. On ne ſçauroit *durer* à la maiſon par ce beau temps-là. On ne ſçauroit *durer* en ce poſte, il eſt trop expoſé à l'artillerie. On n'y *dure* point, on n'y peut tenir. MOL. Penſez-vous que je puiſſe *durer* avec toutes ſes turlupinades? ID. On dit auſſi ne pouvoir *durer* de chaud & de froid &c. pour dire, être extrêmement incommodé du chaud, du frold &c.

On dit proverbialement, Il faut faire vie qui *dure*, lorſqu'on parle de menage, & qu'on veut empêcher la diſſipation. On dit d'un niais qui n'a point vû le monde, qu'il eſt bien neuf, qu'il *durera* long temps. On dit que le temps *dure* à quelcun; pour dire, qu'il lui ennuye, qu'il attend quelque choſe avec grande impatience. On dit auſſi, qu'un homme ne ſçauroit *durer* en ſa peau, qu'il ne peut *durer* en place; pour dire, qu'il eſt inquiet & inconſtant.

DURET, ETTE. adj. diminutif de *dur*. L'oiſeau étoit bon, mais il étoit un peu *duret*. Il eſt du ſtile bas.

DURETE'. ſ. f. Solidité, qualité de ce qui eſt dur. C'eſt la reſiſtance que font les corps à la diviſion, & à la ſeparation des parties dont ils ſont compoſez. Le repos, la liaiſon, & la contiguité des parties qui ſe touchent immediatement ſans ſe mouvoir, fait la *dureté* des corps. ROH. On a trouvé l'invention de donner au plâtre la *dureté* du marbre. La *dureté* des diamans fait la meilleure partie de leur valeur. Les viandes gelées ont de la *dureté*.

En termes de Medecine on appelle *duretez*, certaines tumeurs ou calloſitez de corps & d'humeurs qui s'endurciſſent. On ſent des *duretez* dans les mains des hommes de travail. ABLAN. On dit auſſi, une *dureté* de ventre, quand on eſt conſtipé; une *dureté* d'oreille, quand on eſt preſque ſourd.

DURETÉ, ſe dit figurément en choſes ſpirituelles & morales, & ſignifie, Indocilité, inſenſibilité, cruauté. La *dureté* du cœur des Juifs obligea Moïſe à leur permettre le divorce. LE MAI. Il a une *dureté* d'eſprit qui fait qu'il ne peut rien comprendre; une *dureté* de cœur, qui fait qu'il n'aime perſonne. Nous joindrons nos forces pour attaquer la *dureté* de ſon humeur. MOL. Le cœur, & le temperament des Stoïciens ne s'accommodoient pas toûjours de la *dureté* philoſophique dont ils faiſoient profeſſion. OE. M. Les opinions de Seneque ont trop de *dureté*. ST. EV. Un peu de *dureté* ſied bien aux grandes ames. CORN. Penſez-vous que je vous pardonne toutes les *duretez* que vous m'avez dites? Vous avez eu la *dureté* de me dire que la converſation de cette Dame vous avoit plû. LET. PORTUG. La *dureté* des termes choque d'autant plus, qu'elle enferme quelque ſorte d'indifference, & de mepris. NIC.

Je renonce à la vanité
De cette dureté farouche,
Que l'on appelle fermeté. QUIN.

parceque ſa chair eſt plus ferme que celle des autres pêches.

DURANDAL. ſ. m. eſt le nom de l'épée de Roland Chevalier Heros de l'Arioſte. On s'en ſert en cette phraſe proverbiale: pour expliquer qu'une viande eſt fort dure, on dit que c'eſt *durandal*, l'épée de Roland.

DURANT. Prepoſition. Pendant, tandis qu'une choſe ſubſiſtera. *Durant* qu'on eſt dans l'emploi il faut faire ſa fortune. Il faut faire ſes proviſions *durant* l'été. N'ai-je pas ſçu que *durant* vôtre voyage, vous avez été de la plus belle humeur du monde? LET. PORTUG. Si jamais la voye du Chretien eſt étroite, c'eſt durant les perſecutions. FL. Ce mot ſe met quelquefois après le nom qu'il regit. J'ay été malade ſix ans *durant*. On lui a aſſigné une penſion ſa vie *durant*.

DURCIR. v. act. & n. Rendre dur. On *durcit* le fer à force de le battre. Le ſoleil *durcit* l'ambre, *durcit* les perles. L'air *durcit* le corail. Un œuf trop cuit ſe *durcit*. La viande *durcit* pendant la gelée.

DURCIR, ſe dit auſſi figurément de l'eſprit, & ſignifie, Rendre ferme. Cela *durcit* l'eſprit.

DURCI, IE. part.

DURE. ſ. f. On ne le dit qu'en cette phraſe: Coucher ſur la *dure*; c'eſt-à-dire, ſur la terre, ou ſans matelas.

DURE'E. ſ. f. Perſeverance des choſes dans leur être; temps meſuré par la ſubſiſtance de quelque choſe. Le temps eſt defini par les Philoſophes, La *durée* d'un mouvement. Dieu a promis à ſes éleus une gloire d'éternelle *durée*. Cette fougue eſt trop violente, elle ne ſera pas de *durée*. Nous ne jouïſſons de la vie qu'à meſure que nous la perdons: chaque moment en abrege la *durée*. MORALE DE P. On juge de la *durée* du temps ſelon la diſpoſition où l'on ſe trouve: celui qui eſt accablé de triſteſſe s'ennuye de la *durée* du temps, parcequ'elle lui eſt penible, & qu'il y fait plus d'attention. MALEB. La *durée* des heures, au regard de l'ennui, & du chagrin, ſe fait plus ſentir que celle des années. BOU. Les Dieux ne ſont immortels que par la *durée* de leurs plaiſirs. DAC. Les paſſions veulent être conduites avec art pour en étendre la *durée*, afin qu'elles ne s'épuiſent pas trop tôt. LE CH. D'H. La *durée* de nos paſſions ne depend pas plus de nous, que la *durée* de nôtre vie. LA ROCH. Je ne meſure pas ma vie par la *durée* du temps; mais par la *durée* de la gloire. BOU. Les Dames pour l'ordinaire trouvent leurs maris de longue *durée*. LE CH. DE M. Cette femme s'eſt mis dans l'eſprit d'égaler la *durée* de ſon deuil à celle de ſa vie, & a choiſi cette triſte, & fatigante voye pour acquerir de la reputation. M. ESP.

Il n'eſt rien ici bas d'éternelle durée. MALH.
Cette tendre amitié par tant de fois jurée,
Qui devoit ſurpaſſer les ſiecles en durée,
A la fin s'eſt éteinte. VOI.

DUREMENT, adv. D'une maniere dure. Il a été traité *durement* par ſon Maître, &c. Ces Religieux ſont couchez bien *durement*. Luther s'eſt exprimé *durement* en parlant de la predeſtination. CL. Il ne faut pas dire *durement* les choſes dures. NIC.

DURE-MERE. ſ. f. Terme d'Anatomie. C'eſt la membrane qui enveloppe le cerveau. Membrane du cerveau groſſe & dure, qui eſt attachée à l'os du crâne.

DURER. v. n. Subſiſter pendant quelque eſpace de temps. Une femme ſe defait de ſon galant quand elle veut; mais il faut qu'elle garde ſon mari tant qu'il *dure*. LE CH. DE M. L'abſence, pour peu qu'elle *dure*, nuit à l'amitié auſſi bien qu'à l'amour. ID. Rien n'approche de l'ennui que donne une paſſion qui *dure* trop.

ST. EV. Un engagement qui doit *durer* juſqu'à la mort, ne ſe doit jamais faire qu'avec de grandes precautions. MOL. L'amour *duroit* un monde au bon vieux temps. MAR. Le monde a déjà *duré* cinq à ſix mille ans. Ce que Malherbe écrit *dure* éternellement. MALH.

DURER, ſe dit auſſi de ce qui eſt ſolide; qui ſubſiſte long temps; qui eſt fort; qui s'uſe difficilement. Le drap d'Eſpagne eſt d'un bon uſer, il *dure* long temps. Ce meuble *durera* un ſiecle, cela *durera* juſqu'au bout.

DURER, avec la negative, ſignifie, Reſiſter, ſouffrir quelque mal, quelque peine, quelque incommodité. On ne ſçauroit *durer* avec cette femme-là, tant elle eſt criarde. Je ne puis plus *durer* avec cette colique. On ne ſçauroit *durer* à la maiſon par ce beau temps-là. On ne ſçauroit *durer* en ce poſte, il eſt trop expoſé à l'artillerie. On n'y *dure* point, on n'y peut tenir. MOL. Penſez-vous que je puiſſe *durer* avec toutes ſes turlupinades? ID. On dit auſſi ne pouvoir *durer* de chaud & de froid &c. pour dire, être extrêmement incommodé du chaud, du frold &c.

On dit proverbialement, Il faut faire vie qui *dure*, lorſqu'on parle de menage, & qu'on veut empêcher la diſſipation. On dit d'un niais qui n'a point vû le monde, qu'il eſt bien neuf, qu'il *durera* long temps. On dit que le temps *dure* à quelcun; pour dire, qu'il lui ennuye, qu'il attend quelque choſe avec grande impatience. On dit auſſi, qu'un homme ne ſçauroit *durer* en ſa peau, qu'il ne peut *durer* en place; pour dire, qu'il eſt inquiet & inconſtant.

DURET, ETTE. adj. diminutif de *dur*. L'oiſeau étoit bon, mais il étoit un peu *duret*. Il eſt du ſtile bas.

DURETE'. ſ. f. Solidité, qualité de ce qui eſt dur. C'eſt la reſiſtance que font les corps à la diviſion, & à la ſeparation des parties dont ils ſont compoſez. Le repos, la liaiſon, & la contiguité des parties qui ſe touchent immediatement ſans ſe mouvoir, fait la *dureté* des corps. ROH. On a trouvé l'invention de donner au plâtre la *dureté* du marbre. La *dureté* des diamans fait la meilleure partie de leur valeur. Les viandes gelées ont de la *dureté*.

En termes de Medecine on appelle *duretez*, certaines tumeurs ou calloſitez de corps & d'humeurs qui s'endurciſſent. On ſent des *duretez* dans les mains des hommes de travail. ABLAN. On dit auſſi, une *dureté* de ventre, quand on eſt conſtipé; une *dureté* d'oreille, quand on eſt preſque ſourd.

DURETÉ, ſe dit figurément en choſes ſpirituelles & morales, & ſignifie, Indocilité, inſenſibilité, cruauté. La *dureté* du cœur des Juifs obligea Moïſe à leur permettre le divorce. LE MAI. Il a une *dureté* d'eſprit qui fait qu'il ne peut rien comprendre; une *dureté* de cœur, qui fait qu'il n'aime perſonne. Nous joindrons nos forces pour attaquer la *dureté* de ſon humeur. MOL. Le cœur, & le temperament des Stoïciens ne s'accommodoient pas toûjours de la *dureté* philoſophique dont ils faiſoient profeſſion. OE. M. Les opinions de Seneque ont trop de *dureté*. ST. EV. Un peu de *dureté* ſied bien aux grandes ames. CORN. Penſez-vous que je vous pardonne toutes les *duretez* que vous m'avez dites? Vous avez eu la *dureté* de me dire que la converſation de cette Dame vous avoit plû. LET. PORTUG. La *dureté* des termes choque d'autant plus, qu'elle enferme quelque ſorte d'indifference, & de mepris. NIC.

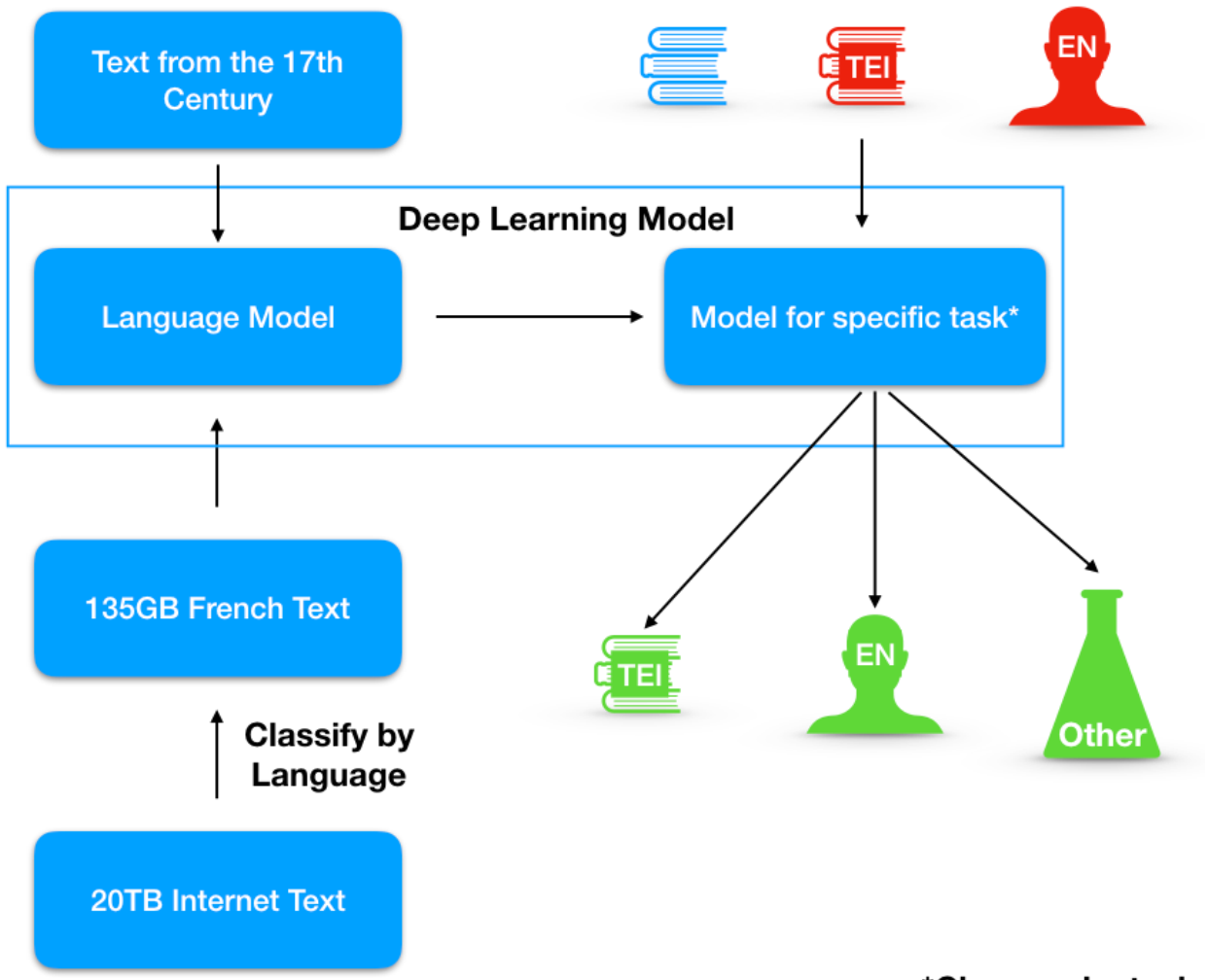
Je renonce à la vanité
De cette dureté farouche,
Que l'on appelle fermeté. QUIN.

Experiments: Lexical Entry Model

	Sample 1			*Sample 2*		
TEI element	**Precision**	**Recall**	**F1**	**Precision**	**Recall**	**F1**
<etym>	87.5	60	71.19	73.68	71.79	72.73
<form>	94.44	92.73	93.58	92.24	96.4	94.27
<pc>	90.91	69.44	78.74	88.97	80.13	84.32
<re>	33.33	9.09	14.29	55.56	22.73	32.26
<sense>	67.65	59.28	63.19	77	76.65	76.84
<xr>	100	80	88.89	100	100	100

Table 3: Field Level Evaluation of the Lexical Entry Model

GROBID: the next generation



Pedro Javier Ortiz Suárez, Laurent Romary, Benoît Sagot. Preparing the Dictionnaire Universel for Automatic Enrichment. *10th International Conference on Historical Lexicography and Lexicology (ICHLL)*, Jun 2019, Leeuwarden, Netherlands. ⟨hal-02131598⟩

Wrapping up

- Dictionaries are cool things
  - But we all share this…
- A great deal of standardisation work has already been done
  - A strong basis for improving interoperability
  - Further convergence work is needed
- Huge expectation around automatic annotation
  - Cf. Basnage: a drop in the legacy ocean (e.g. SIL)
    - But lack of generalisation across dictionaries
  - Future:
    - Families of dictionaries
    - Deep learning
- Towards a wealth of dictionary sources
  - Changing scale for more lexical knowledge
  - Towards a lexical time-space machine => bringing back knowledge to the lexicographic folk

Merci pour votre attention!