

LEXUS Lexicon Tool - an introduction -

Marc Kemps-Snijders

Peter Wittenburg

Max-Planck-Institute for Psycholinguistics

peter.wittenburg@mpi.nl

www.mpi.nl

www.mpi.nl/DOBES

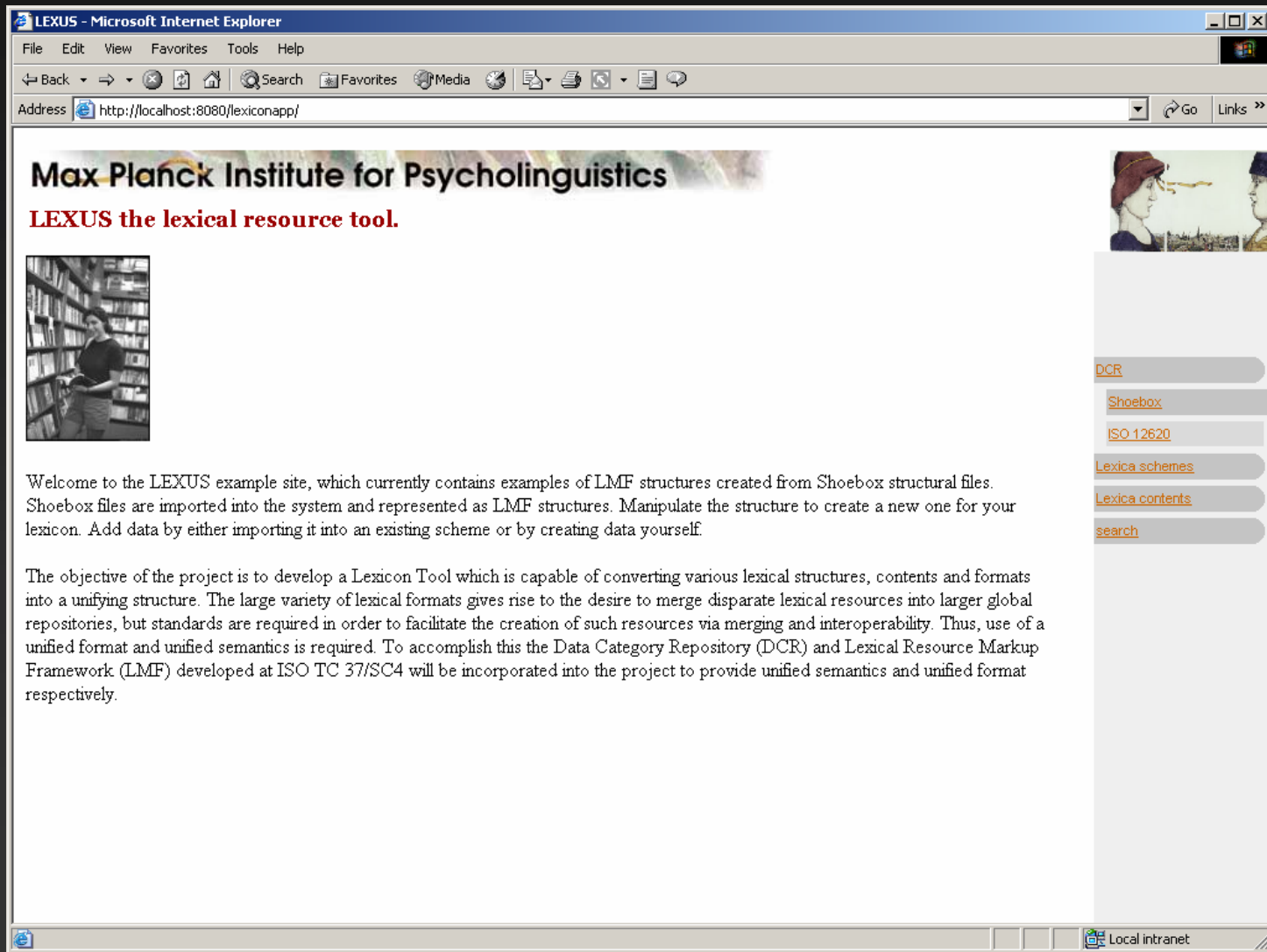
Lexical Heterogeneity

- analysis showed that almost every field researcher creates a suitable lexicon dependent on the language in focus and dependent on linguistic theory and purpose
- in DOBES the 7 pilot teams had in total 10 different structures
- an analysis by ISO for the language engineering area came to the same conclusions
- it was concluded that for both areas, FL and LE, a lexicon framework is needed that is flexible enough to handle all possible lexicon structures
- at an ISO preparation workshop the association with LEGO bricks was worked out
- a flexible framework would give the chance to execute searches across several lexica and to easily compare lexica on computers
- this was the reason to develop the Lexicon Markup Framework within ISO TC37/SC4 that defines a blueprint for such a flexible framework
- a proposal of the LMF standard is being worked out (MPI specialists are actively participating)


- another problem for interoperability is the large variation in attribute and value naming
 - the concept “noun” can be labeled as “n”, “no”, “noun”
- to increase the interoperability a lexicon framework should use lexicon attributes from well-accepted Data Category Registries such as being worked out by ISO, GOLD and as offered by Shoebox
- for existing lexica this semantic unification can be achieved by creating a “language profile” which includes mappings to DCRs
 - this method is applied in the FIELD tool from E-MELD and in the OntoELAN tool created by WayneState University
- for new lexica this could also be achieved by referring directly to the DCRs
 - this method is applied by Shoebox and in the first version of LEXUS
- in addition ontologies can be used that relate various attributes such that search engines can be used to exploit the relations

- Lexus is an implementation of LMF as being worked out in ISO
- since Shoebox is widely used in FL LEXUS imports Shoebox structure definitions and contents
- an export to Shoebox will be supported – but creates some problems due to Shoebox limitations
- LEXUS is a tool working locally on PC/MAC, but also remotely via the web
- it is programmed based on Java and uses only Open Source components
- LEXUS will have the following major functionality
 - flexible structure definition, structure import and dynamic modification
 - linking to data categories from registries such as from ISO, Shoebox, etc
 - lexicon content import from structured lexica and content editing
 - XML-output according to LMF specifications
 - support for Shoebox files
 - cross-lexical operations such as searches
 - synchronization features for lexical creation by teams
 - input methods for various languages such as in ELAN
 - be-directional interaction with the ELAN annotation tool
 - integration of LEXUS into the Flexible Archive Exploitation Framework that is currently being developed at the MPIs Nijmegen and Leipzig
 - integration of multimedia contents

(the features marked in yellow are already implemented)



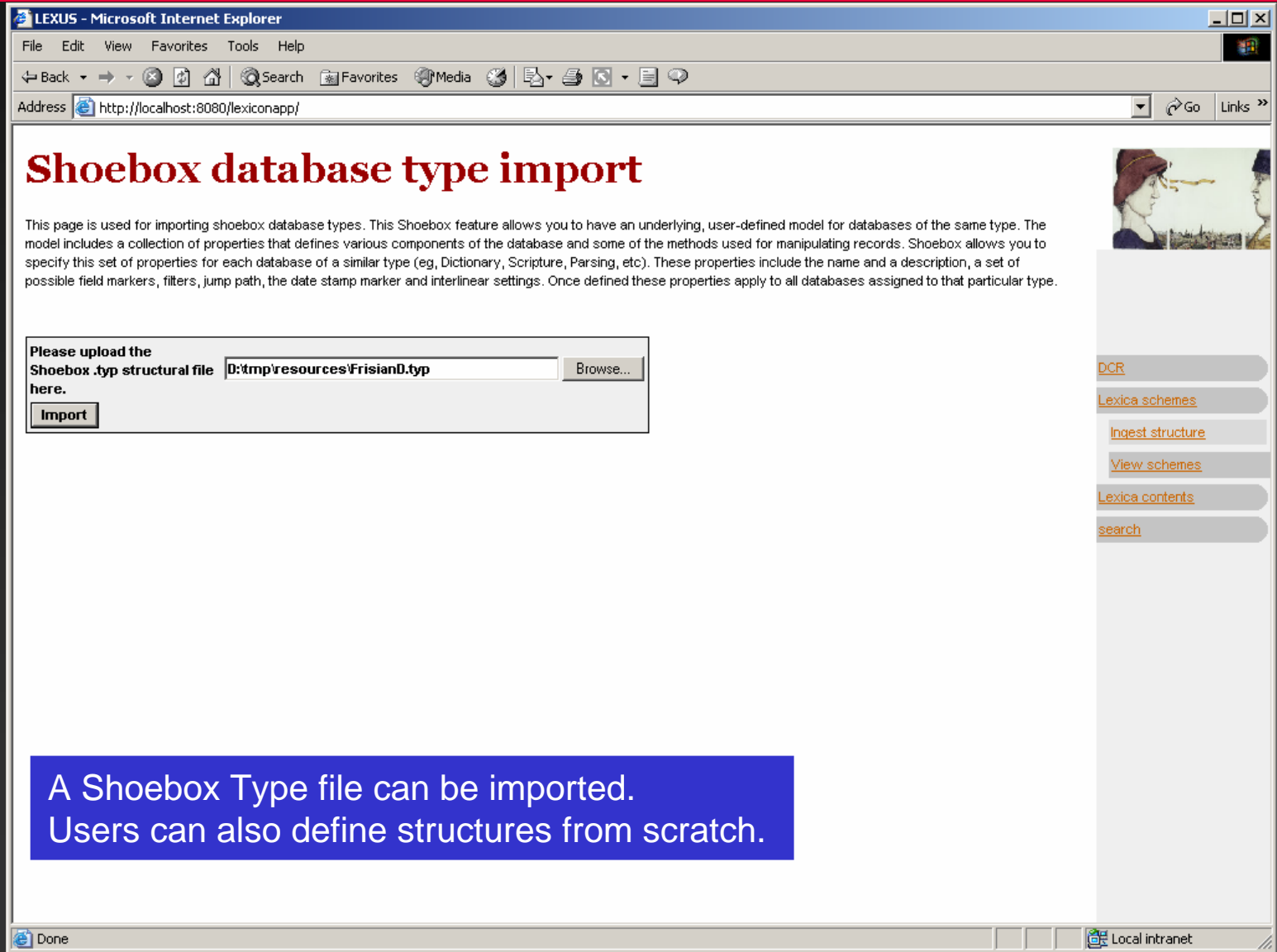
Max Planck Institute for Psycholinguistics
LEXUS the lexical resource tool.



Welcome to the LEXUS example site, which currently contains examples of LMF structures created from Shoebox structural files. Shoebox files are imported into the system and represented as LMF structures. Manipulate the structure to create a new one for your lexicon. Add data by either importing it into an existing scheme or by creating data yourself.

The objective of the project is to develop a Lexicon Tool which is capable of converting various lexical structures, contents and formats into a unifying structure. The large variety of lexical formats gives rise to the desire to merge disparate lexical resources into larger global repositories, but standards are required in order to facilitate the creation of such resources via merging and interoperability. Thus, use of a unified format and unified semantics is required. To accomplish this the Data Category Repository (DCR) and Lexical Resource Markup Framework (LMF) developed at ISO TC 37/SC4 will be incorporated into the project to provide unified semantics and unified format respectively.

- [DCR](#)
- [Shoebox](#)
- [ISO 12620](#)
- [Lexica schemes](#)
- [Lexica contents](#)
- [search](#)



LEXUS - Microsoft Internet Explorer

File Edit View Favorites Tools Help

← Back → Stop Home Search Favorites Media

Address <http://localhost:8080/lexiconapp/> Go Links »

Shoebox database type import

This page is used for importing shoebox database types. This Shoebox feature allows you to have an underlying, user-defined model for databases of the same type. The model includes a collection of properties that defines various components of the database and some of the methods used for manipulating records. Shoebox allows you to specify this set of properties for each database of a similar type (eg, Dictionary, Scripture, Parsing, etc). These properties include the name and a description, a set of possible field markers, filters, jump path, the date stamp marker and interlinear settings. Once defined these properties apply to all databases assigned to that particular type.

Please upload the Shoebox .typ structural file

[DCR](#)

[Lexica schemes](#)

[Ingest structure](#)

[View schemes](#)

[Lexica contents](#)

[search](#)

Done Local intranet

A Shoebox Type file can be imported.
Users can also define structures from scratch.

LEXUS - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address <http://localhost:8080/lexiconapp/>

LEXUS-Lexicon Scheme Viewer

Our Lexica

- Lexicon testing
 - LexicalEntry
 - Frisian Word

Lexical Extension: Frisian Word

Model name: to be filled out


Cardinality:

Description:
test description

Admin info:
to be filled out

Content features:

name	description	category reference
Frisian Word	test description	fri
Frisian Alternate		a
Comment		com
English		eng
English Alternate		enga
Environment		env
Gloss		g
Morpheme Cooccurrence Constraint		mcc
Morpheme Property		mp
Part of Speech		ps
Underlying Form		u



[DCR](#)

[Lexica schemes](#)

[Ingest structure](#)

[View schemes](#)

[Lexica contents](#)

[search](#)

A simple structure is shown with the viewer. The LMF blueprint model for lexical extensions can be seen. Complex categories including other lexical extensions are indicated by tree icons. Complex categories only including primitive categories are indicated by sheet icons.

LEXUS - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address <http://localhost:8080/lexiconapp/>

LEXUS-Lexicon Scheme Viewer

- Our Lexica
 - Lexicon testing
 - LexicalEntry
 - Lexeme
 - Subentry
 - Etymology (proto form)
 - Part of speech
 - Paradigm
 - Paradigm label
 - Sense number
 - Cross-reference
 - Lexical function
 - Lexical function lex
 - Reference
 - Example (v)
 - Variant form(s)

Model name: to be filled out

Cardinality:


Description:
Used for the vernacular lexeme in a lexical function network. The 'lv' field "points to" the vernacular lexeme (a morpheme, word, or phrase) that is semantically related to the current headword as mapped or cataloged by the label in the 'Vf' field.

Admin info:
to be filled out

Content features:

| name | description | category reference |
|----------------------------|--|--------------------|
| Lexical function lexeme | Used for the vernacular lexeme in a lexical function network. The 'lv' field "points to" the vernacular lexeme (a morpheme, word, or phrase) that is semantically related to the current headword as mapped or cataloged by the label in the 'Vf' field. | lv |
| Lexical function gloss (E) | This is for giving the English gloss of the vernacular lexeme referenced by the lexical function. | le |
| Lexical function gloss (n) | This is for giving the national language gloss of the vernacular lexeme referenced by the lexical function. | ln |
| Lexical function gloss (r) | This is for giving the regional language gloss of the vernacular lexeme referenced by the lexical function. | lr |

Save



[DCR](#)

[Lexica schemes](#)

[Ingest structure](#)

[View schemes](#)

[Lexica contents](#)

[search](#)

LEXUS - Microsoft Internet Explorer

Address: http://localhost:8080/lexiconapp/

LEXUS-Lexicon Scheme Viewer

- Our Lexica
 - Lexicon testing
 - LexicalEntry
 - Lexeme
 - Subentry
 - Etymology (proto form)
 - Part of speech
 - Paradigm
 - Paradigm label
 - Sense number
 - Cross-reference
 - Lexical function
 - Lexical function lex
 - Reference
 - Example (v)
 - Variant form(s)

Model name: to be filled out

Cardinality:

Description:
This is a generic reference marker used to link together any two related entries in the lexicon. The content is a vernacular lexeme. If the relationship is known, the lexical function 'lf' field is a better way to cross-reference two lexemes.

Admin info:
to be filled out

Content features:

| name | description | category reference | |
|----------------------|--|--------------------|---|
| Cross-reference | This is a generic reference marker used to link together any two related entries in the lexicon. The content is a vernacular lexeme. If the relationship is known, the lexical function 'lf' field is a better way to cross-reference two lexemes. | cf | |
| Cross-ref. gloss (E) | Gives the English gloss(es) for the vernacular lexeme referenced by the preceding 'cf' field. | ce | |
| Cross-ref. gloss (n) | Gives the national language gloss (es) for the vernacular lexeme referenced by the preceding 'cf' field. | cn | |
| Cross-ref. gloss (r) | Gives the regional language gloss (es) for the vernacular lexeme referenced by the preceding 'cf' field. | cr | <input type="checkbox"/> text
<input checked="" type="checkbox"/> Shbx
<input type="checkbox"/> ISO |

DCR

Lexica schemes

Inqest structure

View schemes

Lexica contents

search

The modification of an existing structure or the creation of a new structure is shown. By selecting "add" an empty attribute is added. Users can select whether it is a new category or whether it should be taken from a DCR.

LEXUS - Microsoft Internet Explorer

File Edit View Favorites Tools Help

← Back → Search Favorites Media

Address <http://localhost:8080/lexiconapp/> Go Links

Shoebox Data Category Repository

Shoebox data categories are listed below.

| Data categorie | Description | Reference |
|-------------------------------|--|-----------|
| First dual | Used to give the vernacular for this particular paradigm form. (It is better to use the new 'pdl' field set for this.) | 1d |
| First plural exclusive | Used to give the vernacular for this particular paradigm form. (It is better to use the new 'pdl' field set for this.) | 1e |
| First plural inclusive | Used to give the vernacular for this particular paradigm form. (It is better to use the new 'pdl' field set for this.) | 1i |
| First plural | Used to give the vernacular for this particular paradigm form. (It is better to use the new 'pdl' field set for this.) | 1p |
| First singular | Used to give the vernacular for this particular paradigm form. (It is better to use the new 'pdl' field set for this.) | 1s |
| Second dual | Used to give the vernacular for this particular paradigm form. (It is better to use the new 'pdl' field set for this.) | 2d |
| Second plural | Used to give the vernacular for this particular paradigm form. (It is better to use the new 'pdl' field set for this.) | 2p |
| Second singular | Used to give the vernacular for this particular paradigm form. (It is better to use the new 'pdl' field set for this.) | 2s |
| Third dual | Used to give the vernacular for this particular paradigm form. (It is better to use the new 'pdl' field set for this.) | 3d |
| Third plural | Used to give the vernacular for this particular paradigm form. (It is better to use the new 'pdl' field set for this.) | 3p |
| Third singular | Used to give the vernacular for this particular paradigm form. (It is better to use the new 'pdl' field set for this.) | 3s |
| Non-animate dual | Used to give the vernacular for this particular paradigm form. (It is better to use the new 'pdl' field set for this.) | 4d |
| Non-animate plural | Used to give the vernacular for this particular paradigm form. (It is better to use the new 'pdl' field set for this.) | 4p |
| Non-animate singular | Used to give the vernacular for this particular paradigm form. (It is better to use the new 'pdl' field set for this.) | 4s |
| Antonym | Used to reference an antonym of the lexeme, but using the 'lf' (lexical function) field for this is better practice. | an |
| Bibliography | Used to record any bibliographic information pertinent to the lexeme. MDF adds the label 'Read:' to this field. | bb |
| Borrowed word (loan) | Used for denoting the source language of a borrowed word. | bw |
| Cross-ref. gloss (E) | Gives the English gloss(es) for the vernacular lexeme referenced by the preceding 'vcf' field. | ce |
| Cross-reference | This is a generic reference marker used to link together any two related entries in the lexicon. The content is a vernacular lexeme. If the relationship is known, the lexical function 'lf' field is a better way to cross-reference two lexemes. | cf |
| Cross-ref. gloss (n) | Gives the national language gloss(es) for the vernacular lexeme referenced by the preceding 'vcf' field. | cn |
| Cross-ref. gloss (r) | Gives the regional language gloss(es) for the vernacular lexeme referenced by the preceding 'vcf' field. | cr |
| Definition (E) | Used to fully express the semantic domains of each sense of a lexeme in English. May be verbose. Other fields ('vee', 'uee', and 'loe') provide for expanded information. | de |

[DCR](#)

[Shoebox](#)

[ISO 12620](#)

[Lexica schemes](#)

[Lexica contents](#)

[search](#)

<http://localhost:8080/lexiconapp/initShbxDCR.htm> Local intranet

LEXUS - Microsoft Internet Explorer

File Edit View Favorites Tools Help

← Back → Search Favorites Media

Address <http://localhost:8080/lexiconapp/> Go Links

ISO 12620 Data Category Repository

ISO 12620 data categories are listed below.

| Data categorie | Description | Reference |
|-----------------------------|--|-----------|
| abbreviation | A designation formed by omitting words or letters from a longer form and designating the same concept. | |
| abkhazian | the alpha-3 (bibliographic) and alpha-2 (terminologic) code | |
| achinese | the alpha-3 (bibliographic) and alpha-2 (terminologic) code | |
| acoli | the alpha-3 (bibliographic) and alpha-2 (terminologic) code | |
| acronym | An abbreviation made up of the initial letters of the components of the full form of the designation or from syllables of the full form and pronounced syllabically. | |
| adangme | the alpha-3 (bibliographic) and alpha-2 (terminologic) code | |
| adjectiveClass | A categorization of an adjective indicating whether it pertains to a single object or to a class of objects. | |
| administrativeStatus | The status of a term with respect to its assignment to an administrative level within a certain working environment. | |
| admittedTerm | A term rated according to the scale of a term acceptability rating as a synonym for a preferred term. | |
| admittedTermAdmnSts | A term rated according to the scale of a term acceptability rating as a synonym for a preferred term. | |
| adyghe | the alpha-3 (bibliographic) and alpha-2 (terminologic) code | |
| afar | the alpha-3 (bibliographic) and alpha-2 (terminologic) code | |
| afrihili | the alpha-3 (bibliographic) and alpha-2 (terminologic) code | |
| afrikaans | the alpha-3 (bibliographic) and alpha-2 (terminologic) code | |
| afroAsiatic(other) | the alpha-3 (bibliographic) and alpha-2 (terminologic) code | |
| akan | the alpha-3 (bibliographic) and alpha-2 (terminologic) code | |
| akkadian | the alpha-3 (bibliographic) and alpha-2 (terminologic) code | |
| albanian | the alpha-3 (bibliographic) and alpha-2 (terminologic) code | |
| albanian | the alpha-3 (bibliographic) and alpha-2 (terminologic) code | |
| aleut | the alpha-3 (bibliographic) and alpha-2 (terminologic) code | |
| algonquianLanguages | the alpha-3 (bibliographic) and alpha-2 (terminologic) code | |
| altaic(other) | the alpha-3 (bibliographic) and alpha-2 (terminologic) code | |

[DCR](#)

[Shoebbox](#)

[ISO 12620](#)

[Lexica schemes](#)

[Lexica contents](#)

[search](#)

<http://localhost:8080/lexiconapp/initISODCR.htm> Local intranet

ISO DCR categories can be selected to be used.

LEXUS - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address http://localhost:8080/lexiconapp/

LEXUS-Lexicon Scheme Viewer

- Our Lexica
 - Lexicon testing
 - LexicalEntry
 - Lexeme
 - Subentry
 - Etymology (proto form)
 - Part of speech
 - Paradigm
 - Paradigm label
 - Sense number
 - Cross-reference
 - Lexical function
 - Lexical function lex
 - Reference
 - Example (v)
 - Variant form(s)

Cardinality:

Description:
This is a generic reference marker used to link together any two related entries in the lexicon. The content is a vernacular lexeme. If the relationship is known, the lexical function 'lf' field is a better way to cross-reference two lexemes.


Admin info:
to be filled out

Content features:

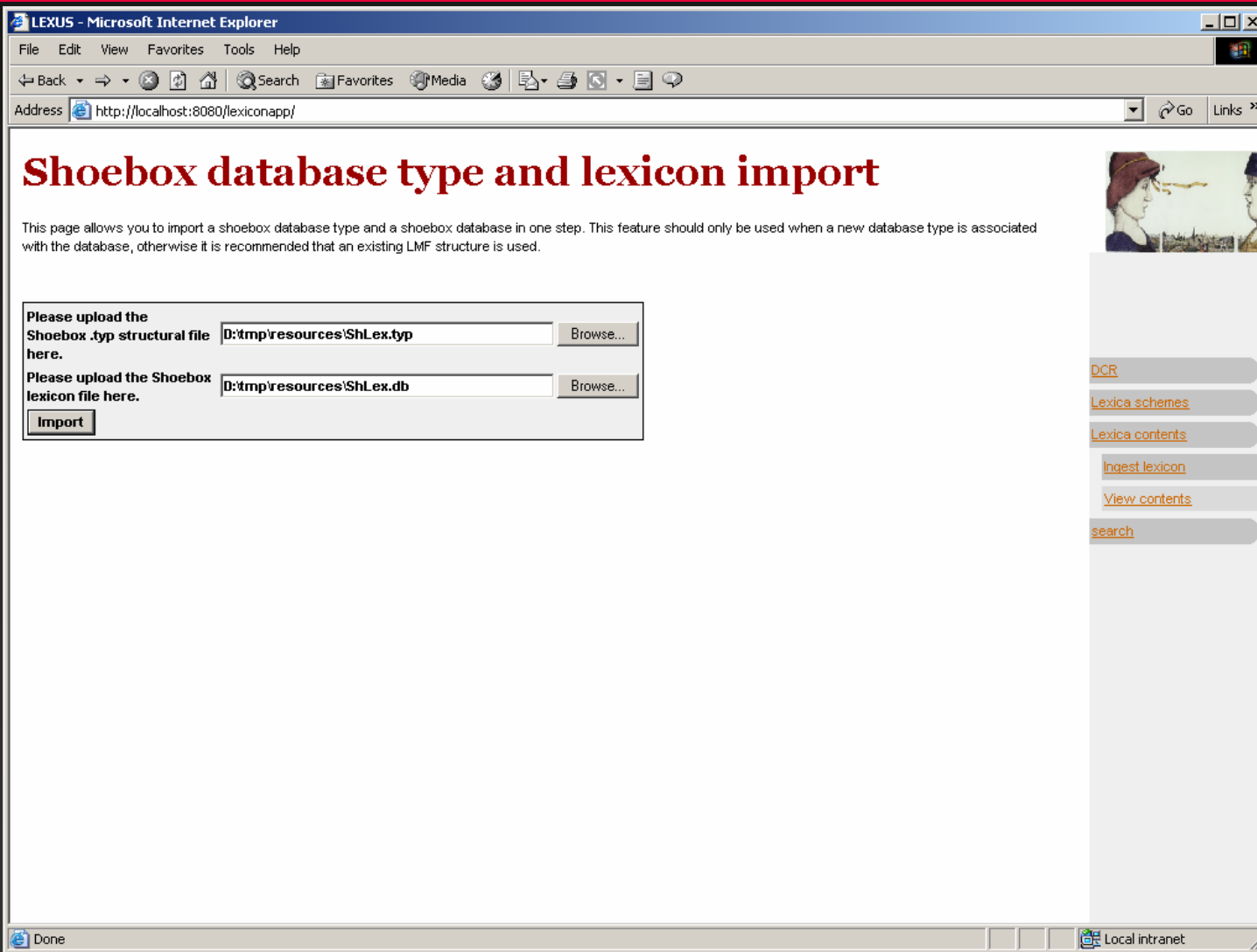
| name | description | category | reference |
|----------------------|--|----------|-----------|
| Cross-reference | This is a generic reference marker used to link together any two related entries in the lexicon. The content is a vernacular lexeme. If the relationship is known, the lexical function 'lf' field is a better way to cross-reference two lexemes. | cf | |
| Cross-ref. gloss (E) | Gives the English gloss(es) for the vernacular lexeme referenced by the preceding 'lcf' field. | ce | |
| Cross-ref. gloss (n) | Gives the national language gloss (es) for the vernacular lexeme referenced by the preceding 'lcf' field. | cn | |
| broaderSynonym | Pertaining to a term that is roughly synonymous to the subject concept, but whos intension is somewhat broader than that of the subject concept. | | |

Add Remove

Save



- [DCR](#)
- [Lexica schemes](#)
- [Inqest structure](#)
- [View schemes](#)
- [Lexica contents](#)
- [search](#)



Shoebox database type and lexicon import

This page allows you to import a shoebox database type and a shoebox database in one step. This feature should only be used when a new database type is associated with the database, otherwise it is recommended that an existing LMF structure is used.

Please upload the Shoebox .typ structural file here.

Please upload the Shoebox lexicon file here.

DCR
Lexica schemes
Lexica contents
Ingest lexicon
View contents
search

Done Local intranet

Lexical contents can be imported. At this moment only a Shoebox import is available. Other I/O modules will be added. Users can also start adding their own content when a schema was defined.

LEXUS - Microsoft Internet Explorer














File Edit View Favorites Tools Help


← Back → Search Favorites Media

Address <http://localhost:8080/lexiconapp/> Go Links

Overview of ingested lexical entries

Below the elements are listed which have been ingested into the system based on the dictionary file you have provided. By clicking on an element you will be directed to the detailed view of the entry.

| Lexical Entry | |
|--|--|
| Lexeme: a
Date (last edited): 24/Mar/1997_sh v3.0 400 SHLEX SAMPLE |  |
| Lexeme: a
Date (last edited): 24/Mar/1997 |  |
| Lexeme: -a
Date (last edited): 23/Feb/1998
Homonym number: 1 |  |
| Lexeme: -a
Date (last edited): 23/Feb/1998
Homonym number: 2 |  |
| Lexeme: -abana
Date (last edited): 22/Jan/1990 |  |
| Lexeme: abat
Date (last edited): 26/Feb/1990 |  |
| Lexeme: -abili
Date (last edited): 01/Feb/1990 |  |
| Lexeme: abu
Date (last edited): 08/Jun/1990 |  |
| Lexeme: -abuk
Date (last edited): 20/Feb/1990 |  |
| Lexeme: -abutak
Date (last edited): 20/Feb/1990 |  |
| Lexeme: -adai
Date (last edited): 30/Oct/1989 |  |
| Lexeme: adar
Date (last edited): 17/Mar/1990 |  |
| Lexeme: -aditi
Date (last edited): 14/Mar/1990 |  |



[DCR](#)

[Lexica schemes](#)

[Lexica contents](#)

[Ingest lexicon](#)

[View contents](#)

[search](#)

Lexical entries can be viewed sorted by the lexemes. More flexible sorting options will be made available so that for example any lexical attribute could be selected for this purpose and that sort orders can be specified.

LEXUS - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media Print Copy Paste

Address http://localhost:8080/lexiconapp/ Go Links

LEXUS-Lexical Entry Viewer


Lexical Object: Sense number

Our Lexica

- Lexicon test
 - LexicalEntry
 - Lexeme
 - Subentry
 - Etymology (proto form)
 - Part of speech
 - Paradigm
 - Paradigm label
 - Paradigm label
 - Sense number
 - Cross-reference
 - Lexical function
 - Lexical function lex
 - Reference
 - Example (v)
 - Reference
 - Example (v)
 - Variant form(s)

Linguistic information units:

| name | value | |
|-----------------------|---|--|
| Encyclopedic info (E) | This is uc: not limited to coconut groves but is used for mangoes, etc. | |
| Gloss (E) | grove | |
| Gloss (n) | dusun | |



- [DCR](#)
- [Lexica schemes](#)
- [Lexica contents](#)
- [Ingest lexicon](#)
- [View contents](#)
- [search](#)



By selecting a lexeme and a tree element the lexical content can be viewed. This is the first viewing options. Other more comfortable viewers will be added.

LEXUS - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Address <http://localhost:8080/lexiconapp/>

LEXUS-Lexical Entry Viewer


Lexical Object: Sense number

Our Lexica

- Lexicon test
 - LexicalEntry
 - Lexeme
 - Subentry
 - Etymology (proto form)
 - Part of speech
 - Paradigm
 - Paradigm label
 - Paradigm label
 - Sense number
 - Cross-reference
 - Lexical function
 - Lexical function lex
 - Reference
 - Example (v)
 - Reference
 - Example (v)
 - Variant form(s)

| Linguistic information units: | |
|-------------------------------|---|
| name | value |
| Encyclopedic info (E) | This is uc: not limited to coconut groves but is used for mangoes, etc. |
| Gloss (E) | grove |
| Gloss (n) | <input type="text" value="my own value"/> |

Save



[DCR](#)

[Lexica schemes](#)

[Lexica contents](#)

[Ingest lexicon](#)

[View contents](#)

[search](#)

The value of an element can be easily changed and new entries can be added.

LEXUS - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media

Address <http://localhost:8080/lexiconapp/> Go Links >>

LEXUS-Lexical Entry Viewer

Lexical Object: Sense number


Our Lexica

- Lexicon test
 - LexicalEntry
 - Lexeme
 - Subentry
 - Etymology (proto form)
 - Part of speech
 - Paradigm
 - Paradigm label
 - Paradigm label
 - Sense number
 - Cross-reference
 - Lexical function
 - Lexical function lex
 - Reference
 - Example (v)
 - Reference
 - Example (v)
 - Variant form(s)

| Linguistic information units: | |
|-------------------------------|---|
| name | value |
| Encyclopedic info (E) | This is uc: not limited to coconut groves but is used for mangoes, etc. |
| Gloss (E) | grove |
| Gloss (n) | my own value |

text | url

Save



[DCR](#)

[Lexica schemes](#)

[Lexica contents](#)

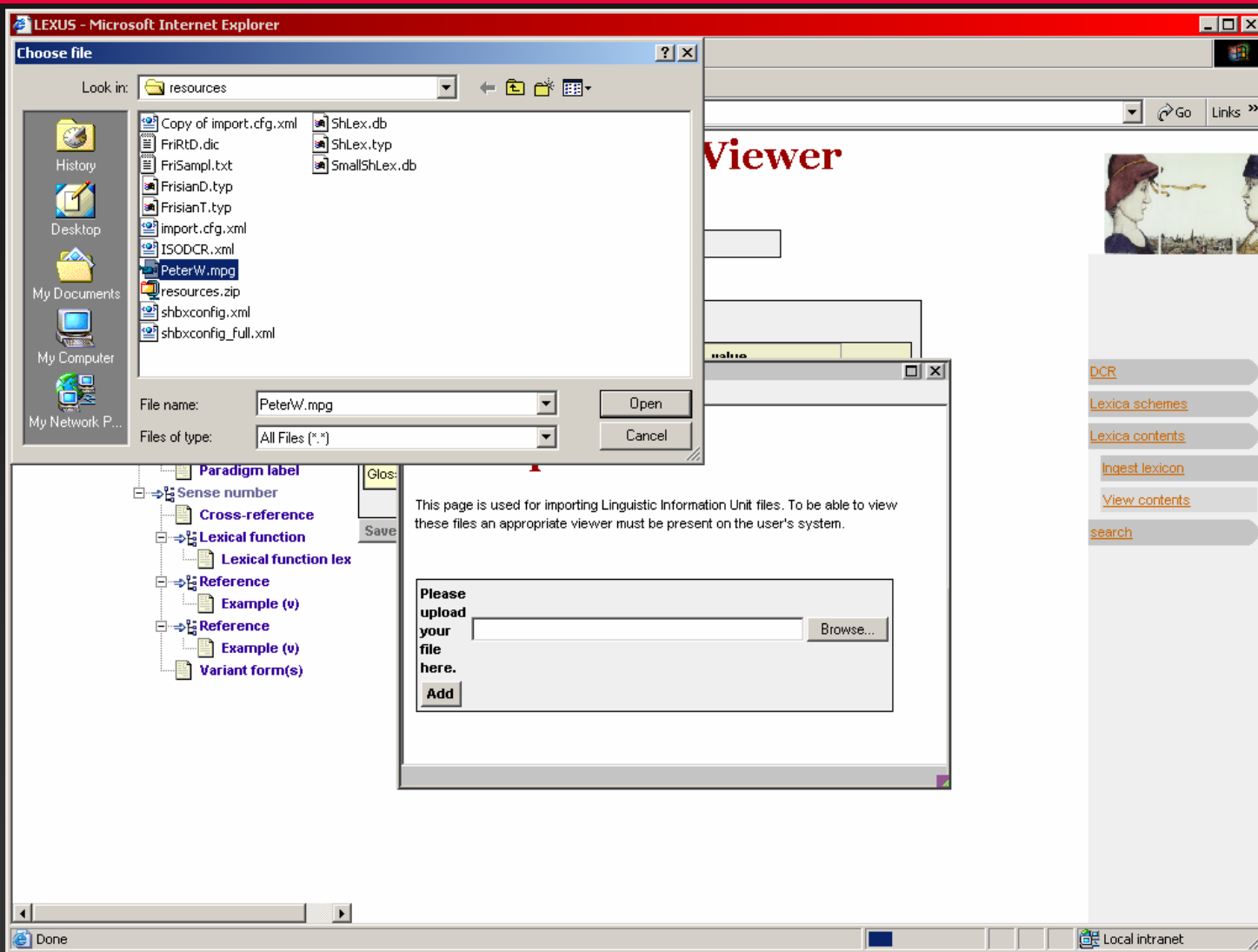
[Ingest lexicon](#)

[View contents](#)

[search](#)

Done
Local intranet

This screenshot simply shows the result of a modification.



The screenshot shows a Microsoft Internet Explorer window titled 'LEXUS - Microsoft Internet Explorer'. A 'Choose file' dialog box is open, showing the 'resources' folder. The file 'PeterW.mpg' is selected. The dialog box has 'File name: PeterW.mpg' and 'Files of type: All Files (*.*)'. The background page is titled 'Viewer' and contains a navigation menu with links: [DCR](#), [Lexica schemes](#), [Lexica contents](#), [Ingest lexicon](#), [View contents](#), and [search](#). The main content area of the page displays a tree structure of linguistic information types: **Paradigm label**, **Sense number**, **Cross-reference**, **Lexical function**, **Lexical function lex**, **Reference**, **Example (v)**, **Reference**, **Example (v)**, and **Variant form(s)**. A text box on the page says: 'This page is used for importing Linguistic Information Unit files. To be able to view these files an appropriate viewer must be present on the user's system.' Below this text is a form with the text 'Please upload your file here.' and a 'Browse...' button. The status bar at the bottom shows 'Done' and 'Local intranet'.

LEXUS - Microsoft Internet Explorer

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media


Address <http://localhost:8080/lexiconapp/> Go Links

LEXUS-Lexical Entry Viewer

Lexical Object: Sense number

Our Lexica

- Lexicon test
 - LexicalEntry
 - Lexeme
 - Subentry
 - Etymology (proto form)
 - Part of speech
 - Paradigm
 - Paradigm label
 - Paradigm label

| Linguistic information units: | |
|-------------------------------|---|
| name | value |
| Encyclopedic info (E) | This is uc: not limited to coconut groves but is used for mangoes, etc. |
| Gloss (E) | grove |
| Gloss (n) |  Please click here to open file.. text url |

[DCR](#)

[Lexica schemes](#)


[Lexica contents](#)

[Ingest lexicon](#)

[View contents](#)

[search](#)

<http://localhost:8080/lexiconapp/resources/127.0.0.110...>



This fake example shows the result of the inclusion of a video clip. By clicking on the reference a player can be started for example.

- a first LEXUS version has been finished within 4 months
- the user interface will need a re-design based on user wishes
- APIs will be provided to allow an integration of LEXUS functionality
- eventual changes in the LMF design will be implemented
- a tool will be provided to map lexica in some XML format into LEXUS structures
- the current version is already very promising
- in particular the link with Data Category Repositories is assumed to improve interoperability
- in particular the user interface needs a major revision
- a close collaboration with ISO TC37/SC4 and E-Meld is intended
- the integration of Shoebox functionality is intended as well



- ISO (2004): *Lexical Markup Framework*. ISO Document in progress
- N. Ide, A. Lenci, N. Calzolari (2003): *RDF Instantiation of ISLE/Mile Lexical Entries*. LDC Workshop. Philadelphia
- P. Wittenburg, W. Peters, S. Drude (2002): *Analysis of Lexical Structures from Field Linguistics and Language Engineering*. LREC 2002 Conference. Las Palma, Mai
- P. Wittenburg (2001): *Lexical Structures*. MPI Technical Report. MPI Nijmegen
- J. Bell, S. Bird (2000): *A Preliminary Study of the Structure of Lexicon Entries*. Workshop on Web-Based Language Documentation and Description. Philadelphia.
- Ide, N., Kilgarriff, A. and Romary, L. (2000), *A Formal Model of Dictionary Structure and Content*, Euralex, Stuttgart

We would like to thank the colleagues in DOBES and ISO TC37/SC4 for the intensive and very fruitful discussions.

We would also like to thank VolkswagenFoundation for the support of this work.