1 Typology of ordinary dictionaries

Monolingual dictionaries explain the meaning of words.

Collins COBUILD EED

Webster

Webster's Comprehensive Dictionary

A large solid hoofed quadruped / *equus caballus/*
with coarse mane and tail,
in the domestic state,
employed as a *beast of draught* and burden
and especially for riding upon.

Collins COBUILD Essential English Dictionary

A *horse* is a large animal which you can ride.
Some horses are used for pulling ploughs and carts.

for native speakers, not for language learners
for language learners, not for native speakers
The Children's Illustrated Encyclopedia

Encyclopedias:
provide real-world knowledge

Dictionaries explain meanings of words

Typology

about language

monolingual

bilingual

encyclopedias

Native

Speaker

Learner

Active

Passive

Webster Collins COBUILD EED

My personal lexicographic experience

Samoan

Teop

Samoan: monolingual school dictionary

Teop: Lexical Database & thematic dictionaries
2 Why dictionaries for endangered languages are special dictionaries

*agevana asiata* (pink siris, albizza species) spangled emperor

dictionary = wordlist + translation equivalent?

Teop

---

### Why dictionaries for endangered languages are special dictionaries

<table>
<thead>
<tr>
<th></th>
<th>ordinary dict.</th>
<th>community dict.</th>
</tr>
</thead>
<tbody>
<tr>
<td>economic basis</td>
<td>commercial</td>
<td>funding agencies</td>
</tr>
<tr>
<td>time frame</td>
<td>decades</td>
<td>3-10 years, part time</td>
</tr>
<tr>
<td>purpose</td>
<td>translating, L2 learning</td>
<td>language maintenance</td>
</tr>
<tr>
<td>users</td>
<td>general public</td>
<td>small community, academics</td>
</tr>
<tr>
<td>lexicographers</td>
<td>professionals</td>
<td>linguists, community members</td>
</tr>
<tr>
<td>linguistic resources</td>
<td>huge corpora, old dictionaries</td>
<td>language documentation</td>
</tr>
</tbody>
</table>

---

### Dictionaries for endangered speech communities: users & purpose

<table>
<thead>
<tr>
<th>Linguists</th>
<th>Community members</th>
</tr>
</thead>
<tbody>
<tr>
<td>understand and analyse texts</td>
<td>preservation of cultural memory</td>
</tr>
<tr>
<td>passive</td>
<td>language maintenance, passive &amp; active</td>
</tr>
<tr>
<td>dictionary for translation</td>
<td>learner dictionary</td>
</tr>
<tr>
<td>linguistic information</td>
<td>linguistic &amp; encyclopedic information</td>
</tr>
</tbody>
</table>

---

### 3 Content and structure of the lexical database

![SIL Toolbox Downloads](image)
3 Content and structure of the a lexical database
Not an end product, but a **dynamic tool**, containing information on

- grammatical features of lexical units  
  (morphology, word class)  
- grammatical relations between words (derivation)

- semantic features
- semantic relations to other lexical units

- translation equivalents
  - description of the meaning in Teop with an English translation

---

<table>
<thead>
<tr>
<th>Structure of the lexical entry of the TLD 2015:</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>lx</strong> head word in Teop</td>
</tr>
<tr>
<td><strong>mn</strong> crossreference to main entry</td>
</tr>
<tr>
<td><strong>mr</strong> morphological structure</td>
</tr>
<tr>
<td><strong>ge</strong> English gloss</td>
</tr>
<tr>
<td><strong>ps</strong> part of speech</td>
</tr>
<tr>
<td><strong>re</strong> reversal (to create an English wordlist)</td>
</tr>
<tr>
<td><strong>sn</strong> sense number</td>
</tr>
<tr>
<td><strong>de</strong> English translation equivalent of <strong>lx</strong>, if possible</td>
</tr>
<tr>
<td><strong>sc</strong> scientific name (for plants and animals)</td>
</tr>
<tr>
<td><strong>pc</strong> picture (link to the Teop Picture Dictionary)</td>
</tr>
<tr>
<td><strong>sd</strong> semantic domain</td>
</tr>
<tr>
<td><strong>rf</strong> reference of Teop descriptions &amp; examples</td>
</tr>
<tr>
<td><strong>xe</strong> English translation of <strong>xce</strong></td>
</tr>
<tr>
<td><strong>ve</strong> Teop description/example</td>
</tr>
<tr>
<td><strong>xe</strong> English translation of <strong>xv</strong></td>
</tr>
</tbody>
</table>

---

Lexical Database - Toolbox
link to picture
description
example
description
4 Lexical database and printed dictionary

<table>
<thead>
<tr>
<th>Lexical database</th>
<th>Printed dictionary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space</td>
<td>unlimited</td>
</tr>
<tr>
<td></td>
<td>restricted</td>
</tr>
<tr>
<td>Macro-structure</td>
<td>multi-dimensional</td>
</tr>
<tr>
<td></td>
<td>linear</td>
</tr>
<tr>
<td>Purpose</td>
<td>dynamic unbiased</td>
</tr>
<tr>
<td></td>
<td>resource and tool</td>
</tr>
<tr>
<td></td>
<td>for researchers</td>
</tr>
<tr>
<td>Content</td>
<td>moderately</td>
</tr>
<tr>
<td></td>
<td>selective</td>
</tr>
<tr>
<td>Meanings</td>
<td>text meanings</td>
</tr>
<tr>
<td></td>
<td>meaning potentials</td>
</tr>
<tr>
<td>examples</td>
<td>citations</td>
</tr>
<tr>
<td></td>
<td>illustrative examples</td>
</tr>
</tbody>
</table>

4 Lexical database, printed and e-dictionary

(e-dictionary: green colour)

<table>
<thead>
<tr>
<th>Lexical database</th>
<th>Printed dictionary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space</td>
<td>unlimited</td>
</tr>
<tr>
<td></td>
<td>restricted</td>
</tr>
<tr>
<td>Macro-structure</td>
<td>multi-dimensional</td>
</tr>
<tr>
<td></td>
<td>linear</td>
</tr>
<tr>
<td>Purpose</td>
<td>dynamic unbiased</td>
</tr>
<tr>
<td></td>
<td>resource and tool</td>
</tr>
<tr>
<td></td>
<td>for researchers</td>
</tr>
<tr>
<td>Content</td>
<td>moderately</td>
</tr>
<tr>
<td></td>
<td>selective</td>
</tr>
<tr>
<td>Meanings</td>
<td>text meanings</td>
</tr>
<tr>
<td></td>
<td>meaning potentials</td>
</tr>
<tr>
<td>examples</td>
<td>citations</td>
</tr>
<tr>
<td></td>
<td>illustrative examples</td>
</tr>
</tbody>
</table>
Lexical database | Printed dictionary
---|---
orthography | can easily deal with variation
standardization preferable
grammar | citations contain speech errors, interferences from dominant languages
standard forms preferable

requires editing with native speakers

**Problem 1:**
**Text meaning vs. meaning potential**

Not the meaning of the lexemes is captured, but only the various senses the word has in the particular contexts of the corpus.

And what we consider as distinct senses is influenced by the translation equivalents.

Example: **naovana**

‘and Gaivaa became a **naovana** that we call seagull’

‘then you get two cockatoo feathers, this is a white **naovana**, ...

**naovana** ‘bird’ but other **naovana**: flying foxes, insects

**Lexical database**
captures the different aspects of word meaning that are activated in its particular context.

**Dictionary**
“contains lists of meaning potentials”

(Hanks 2000/2008)

**How can we capture the relevant “meaning potentials” and present them in a sensible way in the dictionary?**
1 meaning, but 2 (or more) translation equivalents

naovana  animal that flies or moves on the ground

translation equivalent:
the word a native speaker would use in the particular context in question.

What kind of ______ is this?

naovana n.a. animal (that lives on the land); bird; insect; ....

Problem 2:
not all citations are suitable as example sentences.

Problem 3:
English translation equivalent in unknown

2008

Applications
The eova is a black bird. It does not fly high. It only stays down near the ground here in the bush. It lays its eggs in big tree stumps which it digs out. Its eggs are bigger than chicken eggs. Bean tara

2013

Applications
Melanesian scrubfowl, Melanesian megapode, a big, black bird that lays its eggs in the ground
the creation of a dictionary is a different job
- revision of sense discrimination
- systematic ordering of senses
- revision of examples
-moderate standardization of orthography
 and grammar

The same applies to electronic dictionaries !
You need two versions:
 a) database as a tool for researchers and lexicographers
 b) version for the general public

Problem 4: time management
no comprehensive dictionary in 5 years
never promise such a dictionary

5. Transforming a lexical data base into small thematic dictionaries

Criteria for selecting themes:
• which themes do the elders and teachers consider as important?
• which are most endangered?
• which are most suitable for indigenous lexicographers?
• which are easy to understand for the linguist?

Problem 5: specialised vocabulary is under-represented in the corpus

Supplementary recordings in Teop for the House-Dictionary:
How to build a men’s house
How to make bamboo walls
How to make the floor from the wood of the kabuu palm
How to make the thatch

Teop language documentation project

Themes of Teop mini-dictionaries:
completed (2003-2015):
house building, fishing, fishes, shellfish, trees and other plants,
texts typed and proof read:
land animals, food, body and health,

Further advantages of mini-dictionaries:
specialised vocabulary is less frequently used
➢ more endangered
➢ less polysemous

endangered species & word

Supplementary recordings
How to make bamboo walls

How to make the thatch from sago palm leaves

vavaagoroho
‘sleeping pattern’

vavaassun
‘standing pattern’

How to make the floor from kobuu palm wood
Problem 6: specialised vocabulary

- it is difficult to translate
- linguists have no expertise in ethno-sciences
- indigenous experts lack the necessary proficiency
  in the target language (English/French/Spanish etc.)

  vernacular encyclopedic descriptions
  with translations

Problem 7: descriptions are not an indigenous genre

but show the expressive potential of the language
reflect the native speaker's metalinguistic knowledge
can show the native speaker's conceptualisation of extra-
linguistic phenomena, e.g. taxonomies
are useful to understand the meaning of words

  training:  “Explain what X is to a child.”

Conventional and idiosyncratic language use
in indigenous lexicography

Let different people work on items
that are presumably described in a similar way; e.g.

  house & canoe  or  fishes & birds

similarities of house and canoe descriptions:
  topic = things
  predicate = put s.th. somewhere

similarities of fishes and birds descriptions:
  topic = animal
  predicate = habits

= properties like size, colour, shape
6 What linguists can discover in thematic-dictionaries

- register analysis:
  - conventional constructions for descriptions of things, properties and events
- systematic patterns of polysemy, homonymy and word formation

Example for systematic conversion/homonymy

<table>
<thead>
<tr>
<th>Noun</th>
<th>Verb</th>
</tr>
</thead>
<tbody>
<tr>
<td>'part of the house'</td>
<td>'add this part (e.g. wall) onto the house'</td>
</tr>
<tr>
<td>'rafter'</td>
<td>'put up the rafters'</td>
</tr>
<tr>
<td>'bamboo wall'</td>
<td>'put up the bamboo wall'</td>
</tr>
<tr>
<td>'floor'</td>
<td>'make the floor of s.th. (a house) from s.th. (wood; planks)'</td>
</tr>
</tbody>
</table>

7 Conclusions

A lexical database and a dictionary are two very different things
- a dictionary is not a by-product
- a dictionary requires hard work and active involvement of the speech community

Start with a mini-dictionary.
Use vernacular encyclopedic descriptions.
Be rewarded by a culturally and linguistically interesting, and completed little dictionary!
Make sure that the format of electronic dictionaries is regularly updated

References

*The Oxford Guide to Practical Lexicography.* OUP.


*Practical Lexicography. A Reader.* OUP, pp. 125-134

In Sylviana Granger & Magali Paquot. *Electronic Lexicography,* Oxford:OUP
References (cont.)

