DOCUMENTARY LINGUISTICS I
prof. Nicole Nau, UAM winter 2016/2017

DOCUMENTING «THE LEXICON» OF A LANGUAGE

❖ making an inventory of words

❖ describing the meaning of words

❖ describing the use of words

❖ describing relations between words

Main questions: **what**? how to?

First questions: **why**? who?

WHAT IS «THE LEXICON»?

"The lexicon, in modern linguistics, has come to mean a repository of otherwise anarchic facts, an inventory of arbitrary pairings of pronunciations with bundles of features. It is where language stores its idiosyncrasies and irregularities." (Haviland 2006)

"The **lexicon** is the component of a linguistic system which can be regarded as a list or network of words or lexical entries (also lexical items, lexemes). It contains information about (a) the pronunciation, (b) the meaning, (c) morphological properties, and (d) syntactic properties of its entries. Furthermore, the lexicon must contain at least the idiosyncratic information about its entries." (Glottopedia.org)

"A lexicon is the knowledge that a native speaker has about a language. This includes information about

- the form and meanings of words and phrases

- lexical categorization

- the appropriate usage of words and phrases

- relationships between words and phrases, and

- categories of words and phrases.

Phonological and grammatical rules are not considered part of the lexicon."
(SIL linguistic glossary, http://www-01.sil.org/linguistics/glossaryoflinguisticterms/whatisalexicon.htm)

WHY MAKE A DICTIONARY OR A WORD LIST?

John Haviland (2006): «**Colonial** tradition», «**ethnographic** tradition», «**pedagogical** tradition»

"In what I call the Colonial tradition, collecting vocabularies was always a vocation of imperialists, often an accidental byproduct of exploration and conquest. Explorers collected flora and fauna, and often they also collected words.

Somewhat less innocent were the wordlists created explicitly to *aid* in conversion, conquest, and control." (Haviland 2006)

SOME QUESTIONS AND EXAMPLES

❖ What do the colonial and the ethnographic tradition have in common?

❖ Can the three traditions be clearly distinguished in dictionaries (as endproducts)? Why not?

❖ Example: Oldest dictionaries of Latvian (1638)

❖ Example of a historical word list, one of the few sources of Old Prussian: The Elbing Glossary (Elbinger Vokabular), 14th century
http://titus.fkidg1.uni-frankfurt.de/didact/idg/balt/altprbs.htm

❖ Example: a fictious dictionary of a dead language
http://www.suduva.com/virdainas/

The Elbing Glossary

Contains 802 words in German and Old Prussian

First entries:

Got – Deywis

Engel – Rapa

Hemel – Dangus

Gestirne – Lauxnos

[...],

Sunne – Saule

Source: Wikipedia

HOW TO COLLECT WORDS?

1. by word lists, dictionaries and translation
2. from texts, corpora
3. generating a list of possible wordforms
4. collecting by concepts, domains

Ad 1) For examples of wordlists, see table in Chelliah & Reuse (2011: 232)

STEDT Questionnaires http://stedt.berkeley.edu/questionnaires

SIL Comparative African Word List http://www-01.sil.org/silewp/2006/silewp2006-005.pdf

AD 4) COLLECTING BY DOMAINS

The classic thesaurus – cf. historic glossaries etc.

Modern application (by SIL): FLEx;

http://rapidwords.net/ see description next slide

Questionnaire: over 500 pages!!!

«Briefly the method involves gathering a group of native speakers of a language in a workshop. The speakers are trained in the method and then work approximately ten days to collect as many words as they can. They are provided with a copy of the list of domains. Each domain includes a few elicitation questions and sample words that help them understand the domain and the type of words that might be included in it. They then use their mental network to think of words in their language that belong to the domain. They write the words down on paper and a typist types them into the program. Approximately 200 language development projects have used the method to collect thousands of words. Most collect well over 10,000 words and one project collected over 20,000. By collecting words in semantic domains, the words are automatically classified. So the result is a classified word list (or what is more popularly called a thesaurus).»

DIGRESSION

➢ Write down 10 fish names in your native language

➢ Write down 10 tree names in your native language

Compare with other students. Discuss if it was easy/difficult and how you did it.

FISHNAMES AND PRESERVATION OF BIODIVERSITY

Pawley (2011): languages of the Pacific islands may have hundreds of fish names

https://www.youtube.com/watch?v=AVStbwJoYss

REFERENCES

- Chelliah, Shobhana L. & Reuse, Willem J. 2011. *Handbook of descriptive linguistic fieldwork.* Chapter 9: Lexicography in Fieldwork.

- Haviland, John. 2006. Documenting lexical knowledge. In: *Essentials of language documentation*.

- Pawley, Andrew. 2011. What does it take to make an ethnographic dictionary? On the treatment of fish and tree names in dictionaries of Oceanic languages. In: *Documenting endangered languages*, eds. G. Haig et al.