



DOCUMENTARY LINGUISTICS I

prof. Nicole Nau, UAM winter 2018/2019

Fourteenth lecture
15 January 2019

TOPICS OF THE DAY: DOCUMENTING LEXICAL KNOWLEDGE AND MAKING DICTIONARIES

- ❖ Wordlist in language documentation (revision and problems)
- ❖ Lexical knowledge and knowledge related to words: how to document?
- ❖ Dictionaries as results («products») of language documentation

HOW TO COLLECT WORDS? WHAT IS / MAY BE PROBLEMATIC? (REVISION)

1. by word lists, dictionaries and translation (examples of wordlists are given in Chelliah & Reuse 2011, 232)
2. generating a list of possible wordforms
3. from texts, corpora (tool: FLEx Fieldwork Language Explorer)
4. collecting by concepts, domains (tool: FLEx Fieldwork Language Explorer; see also <http://rapidwords.net/>)



Sometimes wordlists are all that is left of a language...

Next slide: 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>

A (fictitious?) dictionary of another extinct Baltic language, **Yatvingian**, is here:

<http://www.suduva.com/virdainas/>

Source: Wikipedia

G ot	Deywis	Somobent	Sabattu
Engel	Rapa	Atte	Same
Hemel	Dangus	Leym	Lavdis
Gestirne	Lauxnos	Sant	Sindo
Wagenfou	Deymarus	let	Lapp
Schengestirne	panoran	Berg	Grabis
Sunne	Saule	Brunt	Dambo
Wonde	memg	Tal	pandanbis
Wullen	truppam	Grabe	Kuops
Helle	porails	Steyn	Stabis
Teufel	Eaur	Uier	Panno
L oz	wettan	Kole	anglis
Somer	Dagis	Functe	Samti
Herby	Eppam	Brunt	Enastis
Witter	Semo	Ayffe	pelame
woche	Samaye	Stoppayffe	pliepmis
Sontag	nadele	Fouch	Dunnis
Montag	ponadele	Brodm	Por
Dmstag	nyssafordis	Gigge	Roene
Mittewoche	passpauate	Sucstant	Boro
Dinnstag	betwre	Blut	Sari
ffirtag	pentum	flamme	Lopis
		L offe	nms

The Elbing Glossary

Contains 802 words
in German and Old
Prussian

First entries:

Got - Deywis

Engel - Rapa


Hemel - Dangus

Gestirne - Lauxnos


[...],

Sunne - Saule

thesaurus

(θɪsɔːrəs )

Collins COBUILD

Word forms: plural thesauruses 

countable noun

A **thesaurus** is a reference book in which words with similar meanings are grouped together.

Synonyms: wordbook, wordfinder, treasury, storehouse

[More Synonyms of thesaurus](#)



COBUILD Advanced English Dictionary. Copyright © HarperCollins Publishers

WHY MAKE A DICTIONARY OR A WORD LIST?

John Haviland (2006):

- ❖ «**Colonial** tradition» and «missionary tradition»
- ❖ «**ethnographic** tradition»,
- ❖ «**pedagogical** tradition»

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

? Can the three traditions be clearly distinguished (when looking at a dictionary)? Why not?

«**COLONIAL 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.»

(Haviland 2006)

«Somewhat less innocent were the wordlists created explicitly to *aid* in conversion, conquest, and control.

The friars' dictionaries of Indian languages in the New World, or vernacular vocabularies destined for colonial bureaucrats in Africa and India, represented unabashedly instrumental “documentation,” often of languages whose eventual endangerment was a byproduct of colonial expansion in the first place. Such wordlists were plainly not made “for” the speakers of the languages so documented.»

(Haviland 2006)

Modern lexicography in the «missionary tradition»: SIL

EXAMPLE FOR AN ENTRY IN A «MISSIONARY» DICTIONARY (HAVILAND 2006)

Here, for example, is a translation of the entry in Aulie and Aulie (1978) for the Chol word *ajaw*, reflex of a root which means “lord, master, God” in other Mayan languages. According to the Aulies, the Chol word means “*espíritu malo de la tierra*,” and they go on to comment:

They call it *lak tat* ‘our father.’ It is believed that a person can make a pact with it. Such a person can make requests of the spirit for or against another. The person who establishes such relations with the *ajaw* is called a “sacristán.” If a man or woman offends the sacristán, the latter appeals to the spirit to curse the other, and in a short time the other person will die.

WORD LIST – VOCABULARY - LEXICON

“The vocabulary of a language is an inventory of the items a culture talks about and has categorized in order to make sense of the world and to survive in a local ecosystem.”

(Nettle, Daniel & Suzanne Romaine. 2000. *Vanishing voices. The extinction of the world's languages*. Oxford: Oxford University Press.)

FISHNAMES AND PRESERVATION OF BIODIVERSITY



DIYADIYAYANA	Striated surgeonfish	<i>Acanthurus lineatus</i>
WULIOALAOALAU	Orangespine unicornfish	<i>Ctenochaetus striatus</i>
OSAOSA	Bullethead parrotfish	<i>Naso vlamingii</i>
OSAALAALAWA	Yellowbarred parrotfish	<i>Chlorurus variegatus</i>
HINEGAYUYU	Barred rabbitfish	<i>Scarus dimidiatus</i>
DEBI	Silver spinefoot	<i>Siganus delaisi</i>
MAMAI	Humphead Maori Wrasse	<i>Siganus argenteus</i>
LAUHAPELO	Coral Trout	<i>Chelinus undulatus</i>
SABUGABUBU	Blackspot snapper	<i>Plectropomus leaptus</i>
GILITA'ATA'AI	Blackspotted grouper	<i>Lutjanus fulviflammus</i>
AUAULI	Blacktipped grouper	<i>Cephalopholis cyanus</i>
UUTAPOTAPOI	Bigeye Bream	<i>Epinephelus spilotos</i>
SALAWA	Sabre Squirrelfish	<i>Monotaxis grandoculis</i>
LEWATAI	Moray eel	<i>Sargocentron</i>

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

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

DIGRESSION

- Write down 10 fish names in your native language
- Write down 10 tree names in your native language

WHAT IS «THE LEXICON»? HOW CAN IT BE DOCUMENTED?

«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>)

Bridging «linguistic behavior» and «metalinguistic knowledge»

«[...] for example, if you are making a thesaurus, you don't want to just find the names of different grasses in your corpus, you also want the resultant list to be discussed and gone over by speakers who are authorities on grasses to make sure you have the field properly covered and to generate good definitions. Rather, what a documentation-oriented view says is that **the discussions of grass names** should themselves be videotaped or tape recorded and **should themselves become a part of the whole corpus**; as should any and all grammatical elicitation of the traditional kind. Moreover, years from now, it will be the grass name attestations and grasses discussion tapes, and not the dictionary, that you will consider as the final document on grass names.»

(Woodbury 2003)

EXAMPLE FOR DIFFERENT CATEGORIZATIONS: «FISH» (PAWLEY 2011)

Table 5. Definitions of the generic for fish and fish-like animals in some Oceanic dictionaries

Group A. Definitions that give 'fish, or 'fish, sea creature' without further definition

Arosi:	i'a , a fish.
Cheke Holo:	sasa , fish (generic).
Marshallese:	ek , fish.
Mota:	iga , a fish.
Owa:	aiga , fish, sea creature.
Rarotongan:	ika , fish.
Roviana:	igana , the generic name for fish.

Group C. Definitions that try to be comprehensive

- Gela: **iga**, a creature of the sea, fish, mollusc, crayfish, whale, squid, sea anemone, etc.
- Paamese: **mesau**, 1. fish. 2 any sea dweller (including also turtles, dolphins, shellfish, etc.).
- Toqabaqita: **iqā**, 1. fish (generic term). 2. Also denotes a superordinate category that includes fish, whales, dolphins, turtles, dugongs.
- Wayan: **ika**, 1. Typical fish, true fish, syn. **ika dū**. This category includes all gill-breathing fish with fins, including sharks, rays and eels. 2. Fish and certain fish-like creatures. A generic which includes all true fish (see sense 1) and dolphins. Most informants also regard turtles (**ikabula**) as **ika**. Some also include octopus (**sulua**) and squid (**suluanū**). Universally excluded are crustaceans (crabs, lobsters, etc.), molluscs with shells, sea cucumbers, sea urchins and jellyfish. (There follows a full list of names of **ika**.)

RECALL: PRINCIPLES OF LANGUAGE DOCUMENTATION (BEGINNING OF 21ST CENTURY) – SLIDE FROM 2ND LECTURE


- ❖ done in **interdisciplinary teams**: linguists + other specialists,
- ❖ outsiders and insiders: speech community is involved not only as „informants”
- ❖ done for **various purposes**
- ❖ published electronically (and on paper)
- ❖ stored in digital archives
- ❖ contains written, audio and video material
- ❖ documents the use of the language in various situations

COLLECTING WORDS IN INTER-DISCIPLINARY FIELDWORK

- ❖ ethnobotany; ethnozoology;
«environmental anthropology», «conservation biology»
- ❖ ethnogeography
- ❖ ethnoastronomy
- ❖ ethnography
- ❖ ...

What is documented? Let's find examples!

What is the motivation and what are possible benefits?



Ethnobiology is the scientific study of dynamic relationships among peoples, biota, and environments (Salick et al. 2003). This discipline has departed from being descriptive and now attempts to use the full spectrum of scientific methodologies and tools **to understand and explain cultural differences and similarities in the knowledge and use of biota and environments** (Bale'e 1994).

(Macclatchey in Thieberger, ed. 2011)

ETHNOBOTANY: FOLK TAXONOMIES (BERLIN 1992)

unique beginner (*plant*)

life form (*tree*)

intermediate

generic (*oak*)

specific

varietal

ETHNOGRAPHER'S INTERESTS (FRANCHETTO 2006)

Semantic domains (p. 187-188):

- Kinship (example p. 209-210)
- Body parts
- Material culture, artefacts
- Terms relating to subsistence activities
- Shamanism
- ...

Topics in texts interesting for ethnographers: p. 189-190

LEXICAL VS. ENCYCLOPEDIAIC KNOWLEDGE, OR: HOW TO DESCRIBE THE MEANING OF *SHEEP*?

“In explicating the meaning of a lexical unit, are there principled grounds for drawing a line between linguistic and encyclopaedic knowledge, or between definitions and cultural information? If so, what are the distinguishing criteria? If not, what practical criteria should be invoked? **These questions haunt every writer of descriptive or ethnographic dictionaries**”
(Pawley 2011)

WHAT IS A CARP?

«A carp is a kind of fish that lives in lakes and rivers.»

COBUILD Advanced English Dictionary



By Dezidor - Own work, CC BY-SA 3.0,
<https://commons.wikimedia.org/w/index.php?curid=2387371>

WHAT IS A CARP? DICTIONARY VS. ENCYCLOPEDIA

«a freshwater teleost food fish, *Cyprinus carpio*, having a body covered with cycloid scales, a naked head, one long dorsal fin, and two barbels on each side of the mouth: family Cyprinidae»
(*Colins English Dictionary*)

«a large variable Asian soft-finned freshwater cyprinid fish (*Cyprinus carpio*) of sluggish waters that is often raised for food and has been widely introduced into U.S. waters» (*Merriam-Webster*)

«**Carp** are various species of oily^[1] freshwater fish from the family Cyprinidae, a very large group of fish native to Europe and Asia.»
(*Wikipedia*)

«OAK» IN DICTIONARIES

«An oak or an oak tree is a large tree that often grows in woods and forests and has strong, hard wood.» (COBUILD)

«any deciduous or evergreen tree or shrub of the fagaceous genus *Quercus*, having acorns as fruits and lobed leaves» (Colins)

«any of a genus (*Quercus*) of trees or shrubs of the beech family that produce acorns» (Merriam-Webster)

«OAK» IN WIKIPEDIA

[HTTPS://EN.WIKIPEDIA.ORG/WIKI/OAK](https://en.wikipedia.org/wiki/Oak)

An **oak** is a [tree](#) or [shrub](#) in the [genus](#) ***Quercus*** (/ˈkwɜːrkəs/^[1] [Latin](#) "oak tree") of the [beech](#) family, [Fagaceae](#). There are approximately [600 extant species of oaks](#). The common name "oak" also appears in the names of species in related genera, notably [Lithocarpus](#) (stone oaks), as well as in those of unrelated species such as [Grevillea robusta](#) (silky oaks) and the [Casuarinaceae](#) (she-oaks). The genus *Quercus* is native to the Northern Hemisphere, and includes [deciduous](#) and [evergreen](#) species extending from [cool temperate](#) to [tropical](#) latitudes in the Americas, Asia, Europe, and North Africa. North America contains the largest number of oak species, with approximately 90 occurring in the United States, while Mexico has 160 species of which 109 are endemic. The second greatest center of oak diversity is China, which contains approximately 100 species.

“The question: ‘What is the meaning of *sheep*?’ is probably wrong-headed. When formulating a dictionary explication of sheep, it makes more sense to ask ‘Of the many characteristics of sheep known to English speakers, which are the most salient?’ and ‘For the various users of the dictionary, what is likely to be the most useful information to include?’”
(Pawley 2011)

Define 1-2 of the fish or tree names you gathered!

FOR DISCUSSION

What words may be more difficult to describe («define») than others?

What other information about words (related to their meaning) may be useful in a dictionary?

What information do you find useful in dictionaries of a language you are learning?

What information do you find essential for a documentation of a lesser used language?

EXAMPLES OF DICTIONARIES OF ENDANGERED LANGUAGES

<https://dictionaria.cld.org/contributions>

<https://www.nationalgeographic.org/archive/projects/enduring-voices/talking-dictionaries/>

cf. Harrison in <https://www.youtube.com/watch?v=GxZAbA3NuEM>

<http://www.webonary.org/> (SIL)

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.