How to plan a dictionary

Ruth Vatvedt fjeld
University of Oslo
What is lexicography?

Lexicography is an activity which consists in observing, collecting, selecting, analyzing and describing, in a dictionary, a number of lexical items (words, word elements and word combinations) belonging to one or more languages.

Bo Svensén: A Handbook of Lexicography
Cambridge 2009
Dictionary typology

• What is the genuine purpose of the dictionary?
  – Communication-oriented
  – Knowledge-oriented
Dictionary types

• Semasiological/onomasiological
  (From word to concept/from concept to word)
• Synchronic/diachronic
• Descriptive/normative
• General/specialized
• Monolingual/bilingual/plurilingual
• Non-illustrated/illustrated/pictorial
Size

• Learner’s dictionary
• School dictionary
• Comprehensive dictionary
Function and organisation

- Reception/Production
- Monoscopal/Biscopal
- Intended users: native-speaker dictionary, foreign learner’s dictionary
Principles for data collection

- Contemporary or arcaic language
- General or technical language
- Written or spoken language
- Standard or regional language
Planning

• Project plan
  – Identifying the demand
  – The time and funding needed
• Dictionary plan
  • Which lemmas are to be included
  • Which information types shall be present
  – Preliminary plan
  – Pilot study
  – Revised plan
  – Final plan
Project plan: The money factor

- Cost of data collection
- Manuscript production
- Manuscript revision
- Production costs
Dictionary plan: an example

- Contemporary language
- General purpose
- Corpus based (100 mill. words)
- Written language
- Descriptive
- 150 000 lemmas

= comprehensive monolingual dictionary
Data collection and data selection

• Principles for data collection
  – Contemporary/arcaic; general/technical; written/spoken; standard/regional
• Calculating the size of each initial letter of the alphabet
• Which information types shall be present in the dictionary?
Format of presentation

• Establish the principles for the structure:
  – **Megastructure**: relations and order of the front and back matter and the lemma sign list
  – **Macrostructure**: the order of and relation between the lemma signs in the lemma sign list
  – **Microstructure**: the order and relation between the information types given for each lemma
Style manual

• Standards for spelling, pronunciation, inflection etc.
• Rules for sense division, collocations, idioms etc.
• How to present different information types
  – Definitions, etymology, frequence, usability, domain etc.
• List of abbreviations
• Instructions for use of dictionary editing system
• Use of corpus-analysis tools
Implementation

- Data collection
- Data selection
- Production of draft manuscript
- Revising the manuscript
- (Typesetting)
- Proofreading and correction x 2
- Printing and binding/release in an electronic format
Follow up

• Maintainance:
  – Correct errors discovered
  – Filling lemma lakunas