

# Syntactic government patterns in the Sketch Engine and in syntagmatic dictionaries for Estonian

Jelena Kallas

Institute of the Estonian Language, Tallinn University  
jelena.kallas@eki.ee

## Sketch Engine (<http://www.sketchengine.co.uk/>) and the extraction of government patterns

The Sketch Engine is a corpus query system which allows the user to view word sketches, thesaurally similar words and “sketch differences”.

**Word sketches** are one-page automatic, corpus-based summaries of a word’s grammatical and collocational behaviour (Kilgarriff, Adam; Rychly, Pavel; Smrž, Pavel; Tugwell, David (2004). The Sketch Engine. In: *Proceedings Euralex*. Lorient, France, July: 105-116).

Estonian word sketches take as input ca 250mln tagged Estonian Reference Corpus.

The system shows explicitly:

1. grammatical relations which correspond to POS-tag and morphological inflections (subject, object, oblique objects, adverbials, modifiers);
2. case, adposition and infinitive government patterns for substantives and adjectives;
3. object, case, adposition and infinitive government for verbs;
4. case government for particle verbs;
5. adposition government for prepositional verbs and case government for adverbs;
6. constructions with conjunctions *ja/või* ‘and/or’, *kui/nagu* ‘as’;
7. predicatives (complements of the copula-like verb *olema* ‘to be’).

## “Basic Estonian Dictionary (2010-2013)”

Basic Estonian Dictionary contains 4,000 most common Estonian words (core vocabulary).

The dictionary is being compiled for Estonian language learners at the beginner (A1, A2) and lower-intermediate (B1) levels.

Dictionary focuses on productive tasks and one of the aims of the dictionary is to give a comprehensive description of syntactic government patterns of Estonian substantives, adjectives, verbs, adverbs and adpositions. Patterns included in the dictionary are provided by word sketches and are chosen on the basis of the criteria of frequency and complexity of their structure. Verb sketches are used also in detecting the sentence patterns of Estonian. Each government pattern is illustrated by example sentence(s).

The dictionary is compiled using the EELex workbench **EELex**, developed at the Institute of the Estonian Language.

## Syntactic and collocational properties of Estonian nouns, adjectives, adverbs and verbs

for **noun** → case, adposition, infinitive government and noun+noun, adjective+noun, noun (subject)+verb, noun (object)+verb collocations;

for **adjective** → case, adposition, infinitive government and adverb+adjective collocations;

for **verb** → object, case, adposition, infinitive government and verb+adverb collocations;

for **adverb** → case government and adverb+adverb collocations.

## Syntactic government patterns of nouns and verbs in the Sketch Engine and in dictionary entries

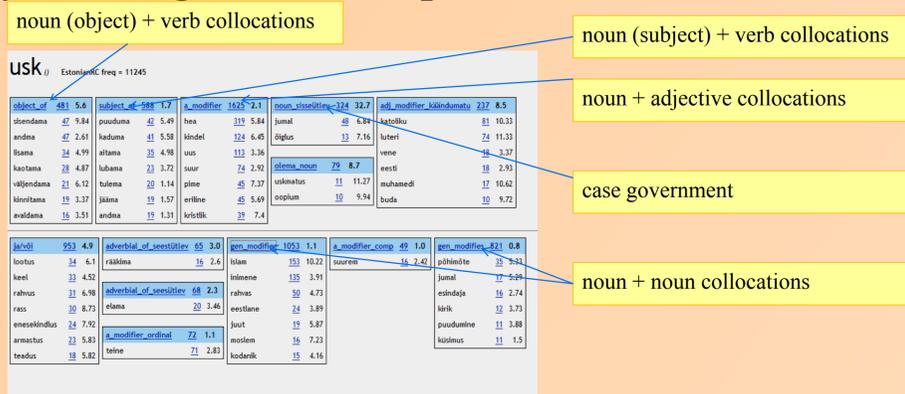


Figure 1. Word sketch for noun “faith”

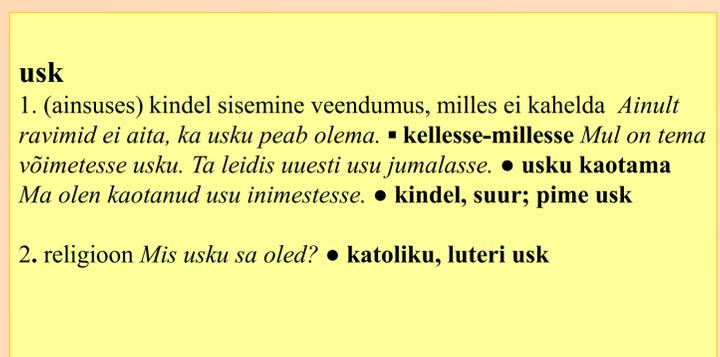


Figure 2. Dictionary entry for noun “faith”

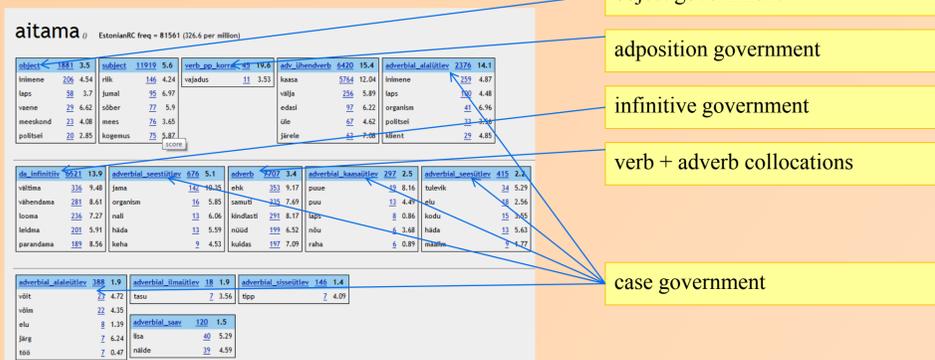


Figure 3. Word sketch for verb “to help”

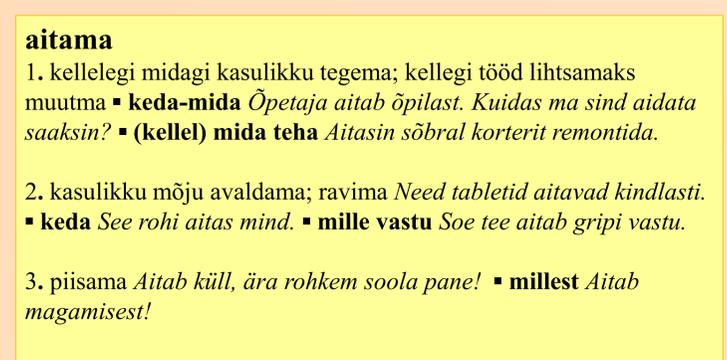


Figure 4. Dictionary entry for verb “to help”

## Acknowledgments

This research was supported by European Social Fund's Doctoral Studies and Internationalisation Programme DoRa

