

## Basic search tips

1. The data in the lexicon can be searched by selecting the tool **Headword Search**.
2. The tool **Headword Search** allows you to search for a particular word or by part of a word. The lexicon provides information on individual words (e.g. *sveikata* ('health') and *nesveikas* ('unhealthy')) and multi-word units (e.g. *valgyk j sveikatą* ('eat to your health')).
3. Some words may be followed by numbers (e.g. SVEIKAS\_1 and sveikas\_2). The numbers are used to indicate homonyms, which are words that have the same form but are different in meaning. Homonyms are presented in separate entries.

The part of speech is indicated next to each word, and its frequency is provided in brackets, indicating how many times this word is used in the Pedagogic Corpus. The frequency of multi-word units is not provided (see, for example, *sveikas kaip ridikas* ('as fit as a fiddle'), *sveikas sugrįžęs* ('welcome back')).

4. Words which are highly frequent in the Pedagogic Corpus are capitalised, for example, SVEIKATA, SVEIKAS\_1 ('health, healthy'). Such words appear in the Pedagogic Corpus with the minimum frequency of 100 times. For such highly frequent words, the **full-record entry** is provided which includes **usage patterns**. Usage patterns represent different word meanings. By clicking on SVEIKAS\_1 and selecting the tab **Meanings**, you can see that this adjective has one meaning, which is realized with three usage patterns (pattern 1.1, pattern 1.2, and pattern 1.3). If the pattern is not expanded, only one example is displayed next to it. To expand the pattern and to view more examples, click on the plus sign.
5. After expanding a specific pattern, a three-colour table with three columns is displayed. Each column represents **three different levels of the pattern: grammatical, semantic, and lexical**. The pattern is read from top to bottom: the first column represents the grammatical level (marked in blue), the second column represents the semantic level (marked in pink), and the third column represents the lexical level (marked in purple). The word described in the pattern is capitalized.

For example, in the case of SVEIKAS\_1 usage pattern 1.1. indicates that at the grammatical level *sveikas* ('healthy') is used as an agreeing attribute with a modified item (Mod), which typically refers to an abstract entity or an individual (at the semantic level). The most common words in the category of abstract entities are *sveika+gyvensena* ('healthy+lifestyle'), *sveikas+protas* ('healthy+mind'), and *sveikas+gyvenimo būdas* ('healthy+lifestyle'); these collocates are indicated at the lexical level of the model. All such usage features recorded in the patterns are represented in examples, so the number of examples depends on the structure and frequency of the pattern.

6. The lowercase letters indicate words and phrases that are used less than 100 times in the Pedagogic Corpus, such as *sveikti* ('to get better') and *sveikas sugrįžęs* ('welcome back'). For such lexical items, the **short-record entries** are provided, giving only examples associated with specific meanings.
7. The patterns (with examples) in the **full-record entries** and the examples in the **short-record entries** can be downloaded as a Word document.
8. In the tool **Headword Search**, the tab **Word forms** provides the forms in which the headword is used in the Pedagogic Corpus; these forms are accentuated, and the frequency of each form and their morphological features are indicated in parentheses (for a full list of conventions used for the mark-up of morphological categories, see the **List of morphological specifications**).
9. By clicking on the tab **Phonetic Information**, you can access the accented form of the headword or a multi-word unit, followed by a phonetic transcription in square brackets and an audio recording (click on the triangle icon).