

TAGGER

A tagger is a program for tagging texts, which means that it analyzes each word in the text for its word class, together with various other properties of the word, such as the inflectional form. The result is called the tag of the word or the tag of the text. In Tungutorg the tagging of the Icelandic source text is a common first step for the translation from Icelandic to Danish and the translation from Icelandic to English. The tag of an Icelandic text, therefore, is an extra product of these translations, and the users of Tungutorg are offered the possibility to call the tag separately without an accompanying translation.

The tagging primarily follows the method applied in *Íslenskri orðtíðnibók* (Frequency Dictionary of Icelandic) from 1991, such that e.g. ordinal numerals are classified with adjectives, prepositions are classified as adverbs that govern case, and exclamations are registered as a subclass of adverbs. An important difference is that at Tungutorg the middle form of a verb is treated as a separate verb, such that their base form, the infinitive, ends in *st*. And furthermore, numerals are classified in several subclasses. In Tungutorg tagging of Icelandic text is performed in this way in order that it serves the machine translation in the best possible way.

An overview of the Icelandic tagset follows.

Icelandic tagset

| | |
|----------|--------------------------|
| n | NOUN |
| k | masculine |
| v | feminine |
| h | neuter |
| e | singular |
| f | plural |
| n | nominative |
| o | accusative |
| þ | dative |
| e | genitive |
| g | with suffixed article |
| m | person name |
| ö | place name |
| s | other proper noun |
| l | ADJECTIVE |
| k | masculine |
| v | feminine |
| h | neuter |
| e | singular |
| f | plural |
| n | nominative |
| o | accusative |
| þ | dative |
| e | genitive |
| s | strong declension |
| v | weak declension |
| o | indeclineable |
| f | positive |
| m | comparative |
| e | superlative |
| f | PRONOUN |
| a | demonstrative |
| b | indefinite demonstrative |
| e | possessive |
| o | indefinite |
| p | personal |
| s | interrogative |
| t | relative |
| k | masculine |
| v | feminine |
| h | neuter |
| 1 | 1st person |
| 2 | 2nd person |
| e | singular |
| f | plural |
| n | nominative |
| o | accusative |
| þ | dative |
| e | genitive |
| g | ARTICLE |
| k | masculine |
| v | feminine |
| h | neuter |
| e | singular |
| f | plural |
| n | nominative |
| o | accusative |
| þ | dative |
| e | genitive |

| | |
|----------|-----------------------------------|
| t | NUMERAL |
| f | cardinal |
| a | year |
| p | percentage |
| o | other numeral |
| k | masculine |
| v | feminine |
| h | neuter |
| e | singular |
| f | plural |
| n | nominative |
| o | accusative |
| þ | dative |
| e | genitive |
| s | VERB (not past participle) |
| n | infinitive |
| b | imperative |
| f | indicative |
| v | subjunctive |
| s | supine |
| l | present participle |
| g | active voice |
| m | middle voice |
| 1 | 1st person |
| 2 | 2nd person |
| 3 | 3rd person |
| e | singular |
| f | plural |
| n | present |
| þ | past |
| s | VERB (past participle) |
| þ | past participle |
| g | active voice |
| m | middle voice |
| k | masculine |
| v | feminine |
| h | neuter |
| e | singular |
| f | plural |
| n | nominative |
| o | accusative |
| þ | dative |
| e | genitive |
| a | ADVERB |
| a | does not govern case |
| o | governs accusative |
| þ | governs dative |
| e | governs genitive |
| u | exclamation |
| m | comparative |
| e | superlative |
| c | CONJUNCTION |
| n | sign of infinitive |
| t | relative conjunction |

unknown word