

When the users jump to conclusions: Presenting prescriptive information

Kristín Ingibjörg Hlynsdóttir & Kristín Bjarnadóttir

The Árni Magnússon Institute for Icelandic Studies

The Database for Modern Icelandic Inflection (DMII) is an online reference and a resource for language technology. Originally, the DMII was mostly descriptive but the online users' demand for prescriptive information was met with notes on usage. With the addition of a new classification system (Kristín Bjarnadóttir and Kristín Ingibjörg Hlynsdóttir 2020), the analysis of variation and errors was expanded greatly. New features of the system include errors encompassing whole paradigms and errors in specific inflectional forms, with normalization, error analysis, and references to the correct forms. The database includes over 300,000 paradigms (including errors) and about 50,000 non-standard word forms (as single inflectional forms).

The online DMII is not only used as a reference on inflection; it is also used to check spelling, acceptability, etc. In these cases, the users may only go as far as the search list, using the least possible number of clicks and stopping when they think they have the answer, without seeing the usage notes. The result is a misinterpretation of the data: All items in the search list are considered to be equally good. This is emphatically not the case in the DMII where the principle is to show the language "as is", including errors. The pair *aðjunkt/aðjúnt* is a case in point. A user only going as far as the search list will assume that both are correct. In fact, the first spelling is correct, the second is not.

The presentation will contain a description of the methods used to convey the features of the classification system to the online users, with reference to a survey of online queries to be completed in January. The goal is to make access to the data as efficient as possible, even when the users are in the clutches of the laziness principle.

References

Kristín Bjarnadóttir, Kristín Ingibjörg Hlynsdóttir & Steinþór Steingrímson. 2019. DIM: The Database of Icelandic Morphology. *Proceedings of the 22nd Nordic Conference on Computational Linguistics*, pp. 146-154. Turku, Finland.

Kristín Bjarnadóttir og Kristín Ingibjörg Hlynsdóttir. 2020. Online Data on Icelandic Inflection: Descriptive to Prescriptive: "Why, for whom, by whom" and how? *Nordiska studier i lexikografi* 15. Rapport från 15 konferensen om lexikografi i Norden. Helsingfors 4-5 juni 2019, pp. 71-79.

Keywords: morphology; inflection; error analysis; normalization