

Icelandic Quirks: Testing Linguistic Theories and Language Technology Thórhallur Eythórsson, University of Iceland (tolli@hi.is)

Linguists working on Icelandic have brought to the fore a number of important empirical facts that at the time of their initial discussion in the theoretical literature were believed to be crosslinguistically very rare, even unattested. Among such “quirks” are the following syntactic phenomena:

- Oblique (“quirky”) subjects (Andrews 1976, Thráinsson 1979)
- Stylistic Fronting (Maling 1980)
- Long Distance Reflexivization (Thráinsson 1979)
- Object Shift of full NPs (Holmberg 1986)
- The Transitive Expletive Construction (Ottósson 1989, Jonas & Bobaljik 1993)
- The New Passive (New Impersonal) (Maling & Sigurjónsdóttir 2001, Eythórsson 2008)

These phenomena provided a testing ground for various theoretical models because they contradicted conventional views on the nature of grammatical categories and syntactic structure; some even went as far as claiming that Icelandic is “not a natural language”. This pessimistic view was authoritatively examined and dismissed by Thráinsson (1996).

The present paper takes the issue one step further, by showing how the discovery of various linguistic structures of Icelandic has led to the recognition of similar facts in other (Germanic, Indo-European and even unrelated) languages, where they had previously gone unnoticed, or had at least not been problematized in terms of linguistic theory. For example, the insight that syntactic subjects can have a morphological case other than nominative was not generally acknowledged until after the oblique subject hypothesis had been proposed for Icelandic. As a consequence, earlier theories on the relation between case and grammatical function had to be revised. Thus, numerous descriptive facts from Icelandic have advanced theoretical linguistics, in that any model of natural language must take them into account.

In addition to their synchronic status, the syntactic phenomena listed above raise questions about the historical development of such “quirks”. On the one hand, Icelandic is known to be a “conservative” language that has preserved many archaic features; on the other hand, despite its relative stability, numerous innovations are known to have taken place in Icelandic, including a number of syntactic changes. Fortunately, we are now in a position to be able to map, at least to a certain degree, the diachrony of Icelandic syntax from the earliest attested documents in the 12th century AD until the present day. This is in particular due to the existence of the Icelandic Parsed Historical Corpus (IcePaHC; Wallenberg et al. 2011), which is currently being put to use in work on Icelandic diachronic syntax. Among other things, this research tool is invaluable in distinguishing between archaisms and innovations in Icelandic syntax. A further corpus, *Greinir skáldskapar* (“Analyzer of Poetry”) (Karlsson et al. 2012), is particularly useful for the analysis of the syntax of the earliest poetic texts of Icelandic.

In conclusion, the above “quirks” present a challenge both to Linguistic Theory and Language Technology. This paper illustrates, by means of selected examples, how this challenge has been successfully met and how advances in linguistic research proceed in a constant interplay between description and theorizing.