

Language contact without contact: A nationwide study of digital minoritization

Sigríður Sigurjónsdóttir (University of Iceland), Iris E. Nowenstein (University of Iceland),
Sigríður M. Björnsdóttir (University of Tromsø), Ásgrimur Angantýsson (University of Iceland),
Anton Karl Ingason (University of Iceland) & Eiríkur Rögnvaldsson (University of Iceland)

We present a project which aims to investigate and model the linguistic consequences of digital language contact, using the rise of English in the Icelandic language community as a test case. In this paper we focus on the effect of increased amounts of English input due to digital media on the acquisition of various syntactic and morphosyntactic variables. First, we investigate variables that have been undergoing internal change in the language, to determine what effect increased English input has on existing internal variation and change. Some examples are the New Impersonal Construction (e.g. Maling & Sigurjónsdóttir 2002) and Dative Substitution (e.g. Svavarsdóttir 1982; Nowenstein 2015), which is seen in (1a). Second, we investigate whether increased English input leads to variation in younger speakers, which represents loss of native syntactic and morphosyntactic elements, such as Verb Second, see (2), and lexical Case, see (1b). Third, we investigate the transfer of English structural elements, such as pseudo passives and tough-movement. When choosing the variables, we considered previous knowledge on incomplete acquisition and language attrition of heritage Icelandic in North America, identifying variables which seem to be contact sensitive (e.g. oblique subject case) and variables which seem to be stable (e.g. canonical word order), see Björnsdóttir (2014) and Thráinsson (in press.).

- (1) a. **Mig** → **Mér** vantar hníf b. **Mig** → **Ég** vantar hníf
me.ACC me.DAT needs a knife ('I need a knife') me.ACC INOM needs a knife
- (2) **Í gær** **var** kennarinn lasinn → **Í gær** kennarinn **var** lasinn
Yesterday was the.teacher sick Yesterday the.teacher was sick

The data is drawn from an online survey reaching a randomized sample of 5,000 Icelandic speakers (approximately 1.5% of the population, stratified by age, 2-65+ years old), of which 400 speakers undergo much more extensive testing. Although Icelandic is the target of this project, we hope that the methodological advances resulting from the project will provide a framework and guidelines to investigate digital minoritization of other languages which, like Icelandic, coexist with English. One of the main theoretical impact of the project is of a broader interest as well, since we aim to integrate sociological factors and bilingualism into the evolving field of models which derive the linguistic knowledge of speakers from the quantified distribution of input in acquisition and hypothesized constraints on possible languages. In particular, we will extend Yang's (2002) Variational Model of Language Acquisition as well as his Tolerance Principle (2016), a framework well equipped for capturing gradual emergence of properties throughout the learning period and across generations of learners.

Background. In recent years it has often been claimed that Icelandic is losing ground to the globally dominant English, which has penetrated Icelandic society through the rapid rise of interactive technology. Many languages are subject to the pressure of English even though they fulfill all requisites for language vitality according to criteria such as the Expanded Graded Intergenerational Disruption Scale, EGIDS, and the UNESCO Language Vitality and Endangerment scales. It has been claimed that at least 21 of Europe's official languages are in danger of 'digital extinction' (see Rehm and Uzokoreit 2012). The perceived effect of these altered conditions is much discussed in popular media and commentators have, for instance, exchanged anecdotes about Icelandic children and youths having entire conversations in English. Still, no systematic accounts of these contact effects exist so far.

Input profiles and variable testing. Based on screening input information from the larger online survey, we will select 400 participants for which we will establish so-called *input profiles*. To create the input profile for each participant, we will estimate the proportion each language occupies in her language space. This will be done by mapping the participant's daily activities to the input they receive, dividing up the hours of their day between the languages they are exposed to. This initial quantification of the input will be complemented by a qualitative assessment in which the input will be classified according to relevant criteria, for example passive or interactive, private or in groups, formal or informal register and written or spoken. The medium of the input (e.g. online chat or physical conversation) will be documented and, when relevant, the relationship in each instance of language exchange (e.g. Icelandic peer or L2 learner of Icelandic) will be assessed. Two resulting measures are computed for each participant: (1) the percentage of each language in their input based on daily activities and (2) the average level of input intensity for each language. By correlating these data with the results from the testing of grammatical variables, which we test through acceptability judgments, sentence completion, elicited examples and naturally occurring speech, we hope to reach important insights into the relation between input and linguistic knowledge.

Conclusion. The project presents a unique opportunity to collect linguistic data on a nationwide scale, investigating an understudied type of language contact and capturing aspects of syntactic and morphosyntactic changes in progress. We test the hypothesis that the status of a native language and global English in the linguistic repertoire of a speaker is primarily a mechanical reflex of the distribution of input in a given type of social environment. The hypothesis will be tested against the alternative hypothesis that changes in the status of the two languages are primarily driven by other factors, such as the attitudes of the speakers involved, taking into account potential links between different types of effects. We argue that a promising way to do this consists in rigorous analysis of input factor's relation to different linguistic variables, and predict that our input measures will be positively correlated with well-known internal variation as well as changes which can be linked to the contact with English. We expect speakers of Icelandic to pattern with speakers of North-American Icelandic, where morphosyntactic variables have the most prominent contact sensitivity.

References

- Björnsdóttir, Sigríður M. 2014. Fallmörkun og samræmi í vesturíslensku erfðarmáli [Morphosyntactic change in North American Icelandic]. MA-thesis, University of Iceland.
- Maling, Joan & Sigríður Sigurjónsdóttir. 2002. The "New Impersonal" Construction in Icelandic. *Journal of Comparative Germanic Linguistics* 5(1): 97–142.
- Nowenstein, Iris E. 2015. Acquiring intra-speaker variation. Poster presented at the 2015 GALANA conference, University of Maryland, College Park.
- Rehm, G., & Uzkoreit, H. (eds.). 2012 *META-NET White Paper Series*. Springer, Berlin.
- Svavarsdóttir, Ásta. 1982. „Þágufallssýki“ [Dative Sickness]. *Íslenskt mál* 4: 19–62.
- Yang, Charles. 2002. *Knowledge and Learning in Natural Language*. Oxford University Press, Oxford.
- Yang, Charles. 2016. *The Price of Linguistic Productivity*. MIT Press, Cambridge.
- Thráinsson, Höskuldur. In press. The North-American Icelandic language and culture. Háskólaútgáfan, Reykjavík.