

Ideal Speakers and Other Speakers

The case of dative and some other cases

Höskuldur Thráinsson
University of Iceland

1. Introduction

In this paper I would like to discuss the nature of linguistic variation. The specific questions I will address are the following:

- (1) a. What is the nature of linguistic variation in general and intra-speaker variation in particular?
- b. What can we learn about internal grammars (the nature of I-language) by studying intra-speaker variation?
- c. What can we learn about the nature of linguistic change by studying intra-speaker variation?

The data come from three major variation studies:¹

- (2) a. Syntactic variation in Icelandic (IceDiaSyn) 2005–
- b. Syntactic variation in Faroese (FarDiaSyn) 2008–
- c. Phonological variation in Icelandic (RÍN) 1980s

The data from the first two surveys will center around variation case marking in general and dative variation in particular. A major point will be the pervasiveness of intra-speaker variation in case marking and data from the phonological variation study will then be presented to show that the extensive intra-speaker variation observed in the syntactic overview projects is not an artifact of the research methods used and the phenomenon is more common than we are normally willing to admit. I will then argue that we need to take intra-speaker variation more seriously than we commonly do when we develop models of speakers' internal grammars. Intra-speaker variation is not just some sort of noise in the data that should be filtered out. It is real and reflects an important aspect of the grammars of normal speakers.

2. Different approaches to intra-speaker variation

As is well known, linguists of different theoretical persuasions tend to have widely diverging ideas about the nature or even existence of intra-speaker variation. Thus many sociolinguists tend to believe that intra-speaker variation is a normal state of affairs. They claim that speakers often alternate between forms that have the same meaning or function and their

¹ The generous support of The Icelandic Research Fund to all these projects is gratefully acknowledged. IceDiaSyn and FarDiaSyn (principal investigator Höskuldur Thráinsson) were connected to the Scandinavian research networks Scandinavian Dialect Syntax (ScanDiaSyn) and Nordic Center of Excellence in Microcomparative Syntax (NORMS). Höskuldur Thráinsson and Kristján Árnason were principal investigators of RÍN. I am grateful to all my collaborators in these projects, both Icelandic, Faroese and Mainland Scandinavian ones, too numerous to mention here.

choice is then typically governed by various factors depending on the linguistic context or social situation. This concept goes back to the work of Labov (see e.g. Labov 1972) and also to the influential paper by Cedergren and D. Sankoff where it is maintained (1974:333) that optional rules can be "assigned application probabilities as functions of the structure of the input strings, possibly depending on the extralinguistic environment".

An alternative approach, although similar in spirit, was later proposed by Kroch and his associates (see e.g. Kroch 1989, 2001 and references cited there). One of Kroch's main points is that the linguistic competence of speakers is often best characterized by assuming that they have acquired two grammars (i.e., as a case of syntactic diglossia) and that these grammars are in competition, both in the linguistic community in general (where one grammar might represent a more conservative and the other a more innovative variant) and in the language of individual speakers. Thus the speakers' output may at times be more consistent with one grammar than the other and "competing forms may differ in social register" (Kroch 2001:702). In such cases we might then expect to find evidence for dialect accommodation in the sense of Trudgill (1986) and others.

A radically different tack is taken by many generative linguists working in the spirit of the principles-and-parameters framework broadly conceived. If one equates intra-speaker variation with some sort of optionality and believes in the principle of economy of derivations, which is e.g. basic to Chomsky's 1995 Minimalist Program, then true intra-speaker variation should not really exist and apparent instances of it need to be explained away. One way of doing so is to argue that certain instances of variability should be possible under economy because the variants could be equally economical and hence "the grammar doesn't mind" (cf. Biberauer and Richards 2006). Another approach maintains that intra-speaker variation may result from underspecified functional categories (cf. Adger and Smith 2010). Common to these approaches is the belief that intra-speaker variation is something special and exceptional. That is actually what one would expect under the standard approach to parametric differences, where it is assumed that parameter settings are binary (plus/minus, one/zero). But maybe that is the wrong way of looking at it.

In recent work, Yang (e.g. 2004, 2010) has argued that parameter setting in language acquisition is not "triggered" by some crucial evidence but rather proceeds in a probabilistic fashion. He has summarized the main points as follows (2004:455):

the learning model extends to a model of language change (Yang 2000), which agrees well with the findings in historical linguistics (Kroch 2001) that language change is generally (i) gradual, and (ii) exhibits a mixture of different grammatical options. But these are possible only if one adopts an SL [statistical learning] model where parameter setting is probabilistic.

Yang's work can thus be seen as a further development of the idea of "competing grammars" usually attributed to Kroch (1989), as described above. Yang puts this as follows (2000:248):

The model formalizes historical linguists' intuition of grammar competition and directly relates the statistical properties of historical texts (hence, acquisition evidence) to the direction of language change. It is important to recognize that, while sociological and other external forces clearly affect the composition of linguistic evidence, grammar competition as language acquisition (the locus of language change) is internal to the individual learner's mind/brain.

I shall argue in some detail below that data from the variation surveys mentioned above provide support for this way of looking at variation while at the same time shedding a new light on the way in which parametric settings can vary between speakers under the same conditions.

3. The methodology of the syntactic surveys

3.1 The main method and some precautionary measures

In the syntactic surveys (IceDiaSyn and FarDiaSyn) we used written questionnaires and then interviewed a (rather small) subset of the participants in the surveys. In a separate pilot study we experimented with a few different methods of questioning the subjects and then ended up with questionnaires with three separate tasks. The main task involved evaluation of sentences where the subjects were given three choices as shown in *Table 1*:²

Settu X í viðeigandi dálk:

- Já** = **Eðlileg** setning. Svona get ég vel sagt.
? = **Vafasöm** setning. Ég myndi varla segja svona.
Nei = **Ótæk** setning. Svona get ég ekki sagt.

| | | já | ? | nei | Athugasemdir |
|-------|---|----|---|-----|--------------|
| T2100 | <i>Þingmaðurinn heimsótti kjósendur.</i> Hann spurði hvort að þeir alltaf hefðu búið í kjördæminu. | | | | |

Table 1: An example from an IceDiaSyn questionnaire.

There were typically over 100 examples of this kind in each questionnaire, broken up by different tasks as will be illustrated below. The grading was explained and illustrated at the beginning of the session and the basic instructions were then repeated at the top of each page as shown here. These instructions read as follows:

- (3) Yes = A **natural** sentence. I could easily say this.
 ? = A **questionable** sentence. I would hardly say this.
 No = An **unacceptable** sentence. I could not say this.

As explained and illustrated at the beginning of each session, the purpose of the italicized sentence was just to give an appropriate context and the subjects were asked to evaluate the second (the non-italicized) sentence only. At the end of each line some space for optional remarks (*athugasemdir*) was provided.

In order to maximize the likelihood of getting reliable answers we took a number of precautions, partly following methodological suggestions found in the literature (cf. e.g. Schütze 1996 (ch. 5), Cornips and Poletto 2005). Thus we would e.g.

- (4) a. **explain the grading scale** by giving illustrative examples
 b. **vary the order** of the test sentences (e.g., reverse for half of the subjects)
 c. **test different constructions** in each overview and **include fillers**
 d. **vary the tasks** (absolute judgments, relative judgments, fill-ins ..., cf. below)

² Although the glosses of the sentences in *Table 1* are irrelevant for our present purposes, the curious linguist might want to know what they are anyway:

Þingmaðurinn *heimsótti* *kjósendur.*
 the parliamentarian visited voters
 Hann spurði hvort að þeir alltaf hefðu búið í kjördæminu.
 he asked whether that they always had lived in the district

- e. **include a break** in long sessions to prevent excessive fatigue and boredom
- f. **include context sentences** to get all subjects thinking of similar contexts
- g. **try to use natural sounding examples** (short, plausible, lexically neutral ...)
- h. test **multiple examples** of each construction to minimize unwanted effects
- i. try to make the contrasting variants maximally close to **minimal pairs**
- j. test **different types of speakers** (age groups, locations ...)
- k. throw out data from **“unreliable speakers”** (e.g. “language specialists”)
- l. get speakers to **report on their own intuition** (cf. Henry 1995, 2005a,b)

The reason for most of these should be obvious but the last one may warrant an explanation. It is sometimes maintained that investigators should ask their subjects **indirect** rather than **direct questions**, e.g., “What do people around here say?” or “What is most common in your (local) dialect?” rather than “Could you say this yourself?” The idea behind this is the belief that speakers might not want to admit that they use a particular variant themselves. In our kind of study it would have made any sense to use the indirect method. The main reason is that we wanted to compare groups of speakers so we had to try to make sure that the speakers were in fact reporting on their own intuition. We were not interested in their beliefs about the language of others. Besides, asking about such beliefs is actually a question of a very different nature as it is actually not a question about linguistic intuition at all. As will be argued in the following subsection, there is every reason to believe that most speakers answered honestly to questions about their intuitions and were in general not influenced by prescriptive ideas.

3.2 *Indicators of reliability*

Before presenting relevant results from our syntactic surveys, it is useful to give some thought to the question how one can tell whether the results are reliable. Some of the indicators of reliability are listed in (5):

- (5) a. The observed **systematic** (e.g. differences between age groups and (in a few cases) regions, etc.) and not random.
- b. Answers from **all generations seem reliable**, e.g. it is not the case that the youngest generation “accepts everything”.
- c. The **subjects answer honestly** in general and don’t seem worried by any kind of prescriptivism or the like (cf. below).
- d. Comparison of **different tasks** confirms reliability of judgments.
- e. Comparison with **corpora** confirms reliability of judgments.
- f. Comparison with **interviews** confirms reliability of judgments.
- g. Comparison of results from IceDiaSyn and FarDiaSyn with results of the **phonological interviews** in RÍN shows interesting parallels strengthening the conclusion.

By the first point we basically mean that *the proof is in the pudding*: If there had been something seriously wrong with the methodology we would not have expected the results to be systematic. Instead, we should have observed cases of random variation, which we did not find.

The second point is also worth emphasizing. We were among other things interested in detecting ongoing changes and also variants that are on their way out. Hence there was a rather large spread in the age of our subjects, the youngest group typically being around 15

years of age and the oldest one 65–70. Now it had been suggested to us that it might be difficult to get reliable judgments from the youngest group, e.g. because the youngest speakers might accept everything. As illustrated by *Figs. 1–2*, however, some variants were accepted most readily by the youngest generation whereas others were favored by the oldest one, which is obviously what we had hoped for (these are figures from IceDiaSyn):³

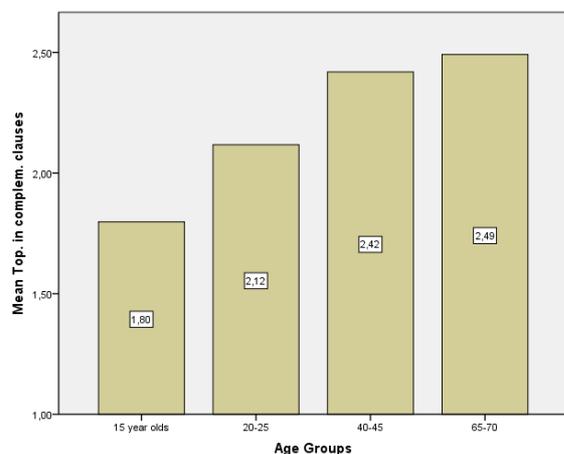


Figure 1: Mean evaluation of Topicalization in complement clauses (N > 700).
Correlation with age: $r = .466$.
Statistical significance: $p < .001$.

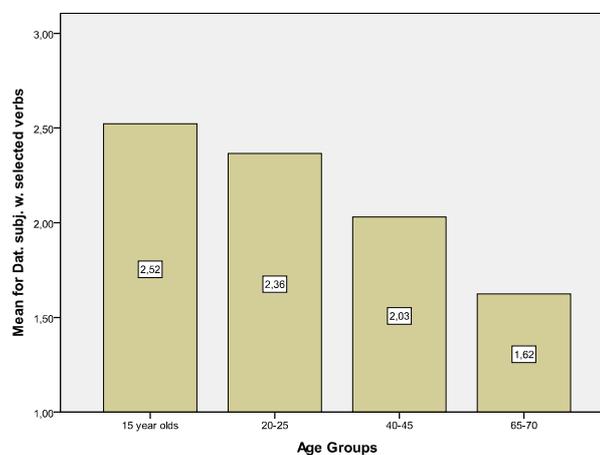


Figure 2: Mean evaluation of Dat. subjects with typical Dative Sickness verbs (N > 740).
Correlation with age: $r = -.511$.
Statistical significance: $p < .001$.

The third point has to do with possible reluctance of the subjects to admit that they find stigmatized variants acceptable. Fortunately, very few of the variants that we were interested in had figured at all in the prescriptive discussion. One notable exception, however, was the so-called "Dative Sickness", namely the tendency to replace accusative subjects with dative subjects for a particular class of ("impersonal") verbs (cf. *Fig. 2*). This particular variant has been frowned upon in schools for a few decades at least, although there is very little evidence that this prescriptivism has slowed down this particular change, as *Fig. 2* indicates. As mentioned above, the questionnaires in IceDiaSyn typically involved different tasks. One of the tasks was the absolute evaluation of sentences explained in the discussion around *Table 1* above. Another was a fill-in task where the subjects were asked to fill in blanks in a short passage by using the pronouns 'I, he, she' as appropriate for the context. Some of the blanks were for subjects of typical Dative Sickness (DS) verbs. This method for investigating subject case marking with this particular class of verbs was originally developed by Ásta Svavarsdóttir (1982). It was found to work well and has been used several times since then (see e.g. Jónsson and Eythórsson 2003, 2005). These studies have suggested that the speakers were generally unaware of the fact that they were providing information about case marking preferences by filling the blanks with pronominal forms. Given this, one might have expected that the acceptance rate of sentences with dative subjects of the relevant verbs would have been lower than the relative frequency of dative subjects selected for these verbs in the (disguised) fill-ins and in spontaneous speech. But this is not what we found. The acceptance rate of sentences with Dative subjects for typical DS verbs was consistently higher in the

³ We made three different surveys using written questionnaires in IceDiaSyn with the number of participants in each ranging from 714 to 772. The means in the graphs are "mean grades" from the evaluation (cf. *Table 1*), where 3 = 'all examples found to be natural' and 1 = 'no examples found to be natural'.

judgment tasks than the selection rate of Dat. subject pronominal forms in the fill-in tasks. This is illustrated in the first row of *Table 2* with numbers from IceDiaSyn for the verb *langa* 'want', which is one of the most common DS verbs. The other rows show figures from a couple of other studies, as will be explained below:

| Study | All age groups | | | Youngest age group | |
|---------------------------|----------------|------------|------------|--------------------|------------|
| | Judgments | Fill-ins | Corpora | Judgments | Fill-ins |
| IceDiaSyn | 68% accept | 19% select | 7%/25% us. | 77% accept | 35% select |
| Jónsson & Eythórsson 2003 | | | | | 40% select |
| Svavarsdóttir 1982 | | | | | 32% select |
| Friðriksson 2008 | | | 15% use | | |

Table 2: Acceptance, selection and usage of Dat. subjects with *langa* 'want' in various studies.

The figure 68% in the first row shows the acceptance rate by all age groups combined in IceDiaSyn for sentences with a Dat. subject with *langa* 'want' and the figure 19% indicates how many of these subjects actually selected a Dat. pronominal subject in a fill-in task in the same survey. Corresponding figures for the youngest age group separately are 77% and 35%, respectively. In a spoken language corpus considered for comparative purposes the corresponding figure is 25%, as shown in the middle column, but in a corpus which is mostly based on written (published) material the figure is only 7%, as Dat. subjects of *langa* are typically weeded out by proofreaders.

The remaining numbers in *Table 2* are also of some comparative interest. First, note that some 40% of the youngsters tested on a fill-in task by Jónsson and Eythórsson (2003) selected Dat. form of the subject. A corresponding figure from Svavarsdóttir's original study (1982) was 32%. Since the tasks and the age groups were comparable (12 year olds mostly, i.e. somewhat younger than the youngest group in IceDiaSyn), this indicates a slow but steady increase in Dative Sickness among Icelandic teenagers during the 20 years that had passed between the studies. Finally, spontaneous speech recordings made by Friðriksson (2008) showed only 15% usage of Dat. (as opposed to Acc.) subjects. This is lower than the 25% usage rate reported for a spontaneous speech corpus by Svavarsdóttir (2006, 2010) in connection with IceDiaSyn and closer to the 19% selection rate for the fill-ins in IceDiaSyn. The observed differences might be due to a number of reasons, one being difference in the age of the speakers involved in the different studies, another the topic of the conversation, as it has often been observed that speakers are more likely to use 3rd person Dat. subjects than 1st person (see e.g. Svavarsdóttir 1982).

The most important aspect of these results is the fact that they show very clearly that the speakers have no problem with accepting Dat. subjects of a DS verb like *langa* 'want' in a judgment task (the acceptance rate ranges from 68% for the group as a whole to 77% for the youngest age group). But this does not necessarily mean that all the speakers who accept the Dat. variant will actually select a Dat. subject form in a fill-in task. As it turns out, many of the speakers actually prefer an Acc. subject with *langa* even if they find the Dat. subject natural. This is what the selection figures show: 19% of the whole group select the Dat. subject in the fill-in task and 35% of the youngest group. This is consistent with the finding that the Acc. subject variant with *langa* actually gets an even higher acceptance rate than the Dat. subject variant in the judgment task (88% of the whole group accepted the Acc. variant and 68% the Dat. one).

These results suggest, then, that the subjects in IceDiaSyn were largely unaffected by prescriptivism since Dative Sickness was actually our main concern from that point of view and here the subjects had no problem admitting that they found the (stigmatized?) Dat.

subjects with the DS verbs natural.⁴ In addition, although the acceptance rate was considerably higher than the selection rate, there was a **strong and highly significant correlation** between the judgments and the selection: $r = .570$, $p < .001$. This is obviously encouraging and lends further support to our belief that the judgments were reliable.

As a final indicator that the subjects' judgments are generally reliable consider the following: As mentioned above, a subset of those who had filled in the written questionnaires were interviewed later. In these interviews we wanted among other things to try to elicit production data to compare to the judgment data that we had obtained. The so-called New Passive Construction (or the New Impersonal, cf. Maling and Sigurjónsdóttir 2002, Thráinsson 2007:273ff.) seemed to offer an opportunity to do this since here there was a great difference between the age groups: The youngest speakers would typically accept the NewPass examples whereas the older ones would virtually all reject them. A typical NewPass example is given in (6c):

- (6) a. *Krakkarnir hrintu mér í frímínútunum.* (active)
 the kids pushed me(D) in break
 'The kids pushed me during the break.'
- b. *Mér var hrint í frímínútunum.* (regular passive)
 me(D) was pushed in break
 'I was pushed during the break.'
- c. *Það var hrint mér í frímínútunum.* (new passive)
 there was pushed me(D) in break
 'I was pushed during the break.'

The regular passive in (6b) is characterized by the auxiliary 'be' and the past participle of the main verb 'push' and the theme (or patient) of the predicate occurs in subject position.⁵ In the NewPass we also get the auxiliary 'be' and the past participle of the main verb, but the theme (patient) stays in situ (in apparent object position) and the subject position is typically filled by an expletive *það* 'there' (or some other element occurs in initial position).

In the interview the subjects were first presented with a model pair of sentences like (7a, b) and it was pointed out to them that a sentence like (7a) could be paraphrased by starting with (the expletive) *það* 'there' as in (7b) (a natural expletive passive for all speakers):

- (7) a. *Einhverjir köstuðu tómötum í söngvarann.*
 some people threw tomatoes in the singer
 'Some people threw tomatoes at the singer.'
- b. *Það var kastað tómötum í söngvarann.*
 there was thrown tomatoes in the singer
 'Tomatoes were thrown at the singer.'

⁴ Although prescriptivists often talk about Dative Sickness and school teachers have tried to eradicate it for decades, it seems that there is very little awareness of the phenomenon among the general public. In interviews conducted in connection with IceDiaSyn we sometimes asked the subjects if they had heard about Dative Sickness. While many of them said that they had, and knew that it had something to do with the use of Dative, they were usually unable to give relevant examples.

⁵ Since Dat. objects have a lexically assigned case, they keep their case in the passive construction. Acc. objects do not in the regular passive, as is well known, but in the NewPass the theme/patient argument would retain an Acc. case (cf. e.g. Maling and Sigurjónsdóttir 2002, Thráinsson 2007:273ff.).

The subjects were then presented with an example like (8a) and asked to paraphrase it by a sentence beginning with *það* 'there' and they typically had no problem coming up with a paraphrase like (8b), which is a natural expletive sentence for all speakers of Icelandic:

- (8) a. *Einhverjir fóru að syngja í rúttunni.*
 some people began to sing in the bus
 b. *Það fóru einhverjir að syngja í rúttunni.*
 there began some people to sing in the bus

But when the subjects were presented with (9a), only the youngest speakers could paraphrase it as (9b), which is a NewPass example as explained above, whereas the older speakers were at a loss:

- (9) a. *Krakkarnir hrintu mér í frímínúttunum.*
 the kids pushed me in break
 b. *Það var hrint mér í frímínúttunum.*
 there was pushed me in break

Here there was almost a perfect correlation between the judgments of the speakers and their production: $r = .989$, $p < .001$. This means that virtually all the speakers who had accepted NewPass sentences in the questionnaires could produce such sentences under these circumstances whereas the ones who had rejected NewPass examples could not produce them either. One could not really ask for a stronger confirmation of reliability of the judgments.

4. Observed variation in dative case marking — and elsewhere

In this section I will first present evidence for inter-speaker variation in datives and then turn to the extensive intra-speaker variation. At the end of the section I will then present some data from phonology to show that the observed intra-speaker variation in case marking is by no means an isolated phenomenon.

Before we turn to the examples it should be pointed out that variation in case marking is arguably more “pure” from a syntactic point of view than many other instances of variation as it typically seems devoid of any semantic or pragmatic nuances (the same is probably true of agreement but not necessarily of word order variation, use of reflexives/non-reflexives, etc.). Hence case alternations are useful for determining the nature of variation. But before we continuing our discussion of variation in case marking, it should be emphasized that for most verbs in Icelandic there is no variation in subject nor object case marking. The same is true of Faroese. So variation in case marking is the exception and not the rule. Nevertheless, these exceptions are important in the present context.

4.1 *Inter-speaker variation in case marking*

Dative case variation in Icelandic is mainly of two kinds, i.e. variation in subject case and object case. Some speakers accept (and use) Dat. subjects with experiencer-type verbs that used to take Acc. or even Nom. subjects. The acceptance rate for selected verbs of this kind in IceDiaSyn is shown in *Table 3*:

| Verb | Nom. subject | Acc. subject | Dat. subject | (N+)A+D |
|-------------------------------------|--------------|--------------|--------------|---------|
| <i>hlakka til</i> 'look forward to' | 48.6% | 59.7% | 44.2% | 152.5 |
| <i>langa</i> 'want, long for' | | 88.3% | 68.2% | 148.7 |
| <i>vanta</i> 'need, lack' | | 92.1% | 56.6% | 156.5 |

Table 3: Acceptance rate of different cases in subject position for selected verbs (N > 740).

As the reader will note, the figures in Table 3 do not only indicate inter-speaker variation but also intra-speaker variation. If there was no intra-speaker variation, the figures in each row (acceptance of Nom. + Acc. + Dat. subjects) would add up to 100% but they add up to around 150 in each case. We will return to this issue in section 4.2 below.

Originally, the verb *hlakka til* 'look forward to' took a Nom. subject whereas *vanta* 'need, lack' and *langa* 'want, long for' took an Acc. subject. As Table 3 shows, Acc. is (still) the most widely accepted subject case for all of these verbs although it represents an innovation for *hlakka til* and has been corrected in schools. Two thirds of the oldest generation (66.0%) accepted (the original) Nom. subject with *hlakka til* but only one third of the youngest generation did (32.5%). But if we compare the acceptance by different age groups of Acc. subjects and Dat. subjects for the typical DS verbs *vanta* and *langa*, we see that the Dat. subject is most widely accepted by the youngest generation but there is very little difference between age groups in the acceptance of Acc. subjects with these verbs. This is illustrated in Table 4:

| Verb | 15 year olds | 20–25 | 40–45 | 65–70 |
|-------------------|--------------|-------|-------|-------|
| <i>langa</i> Acc. | 85.2% | 88.2% | 90.8% | 89.3% |
| <i>langa</i> Dat. | 79.6% | 78.0% | 66.7% | 43.8% |
| <i>vanta</i> Acc. | 86.7% | 96.1% | 93.8% | 91.7% |
| <i>vanta</i> Dat. | 46.9% | 35.1% | 15.4% | 7.5% |

Table 4: Acceptance rate of Acc. and Dat. subjects with selected DS verbs (N > 740).

This becomes even clearer on the bar chart in Figure 3, where the mean evaluation of Acc. subjects for examples with the verbs *langa* and *vanta* is broken up according to age groups. Figure 4 shows comparable results for Dat. subjects with the same verbs:

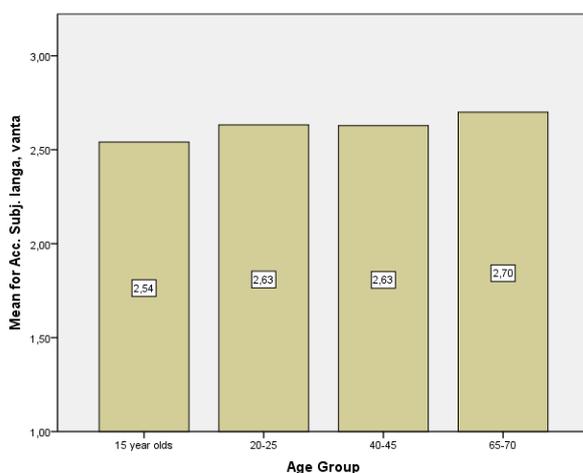


Figure 3: Mean evaluation of Acc. subjects with *langa* 'want' and *vanta* 'need'. Correlation w. age: $r = .133$.

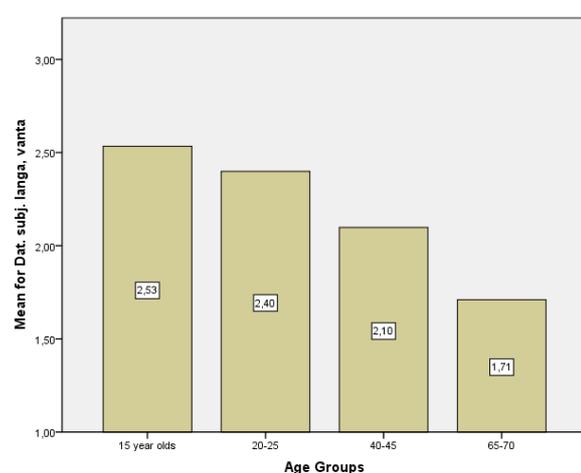


Figure 4: Mean evaluation of Dat. subjects with *langa* 'want' and *vanta* 'need'. Correlation with age: $r = -.459$.

These figures show that the younger age groups are more "bidialectal" than the older ones in the sense that they typically accept both the innovative case marking (Dat.) and the traditional one (Acc.) whereas the older generations (especially the oldest one) only accept the traditional value. This is an interesting result from a theoretical point of view since it gives an idea of how linguistic change may spread and suggests that intra-speaker variation may be a more common state of affairs under such circumstances than often assumed. But before we look more closely at intra-speaker variation, it is useful to look at an instance of variation in object case for comparison.

While it is standardly assumed that Acc. is the default (or structural) object case in Icelandic and Dat. is a lexically assigned case, it is also well known that Dat. object case shows some regularities (see e.g. Barðdal 2001, Maling 2002, Svenonius 2002, Woolford 2006, Thráinsson 2007:156ff.). What has been less well documented in the literature is the fact that for some verbs there is a variation between Dat. and Acc. on objects. This is obviously of interest for theories that attempt to relate Dat. object case marking to semantic or thematic notions, since there is no reason to assume that this variation correlates with differences of interpretation by the speakers involved. First, consider the figures in *Table 5*:

| Verb | Acc. object | Dat. object | Total Acc.+Dat. |
|-----------------------------------|-------------|-------------|-----------------|
| faxa 'faxa' | 91.3% | 23.8% | 115.1 |
| framlengja 'extend' | 82.7% | 61.0% | 143.7 |
| negla 'nail (a ball into a goal)' | 66.5% | 72.6% | 139.1 |
| rústa 'demolish' (lit. and fig.) | 22.3% | 82.4% | 104.1 |

Table 5: Acceptance of Acc. and Dat. objects with selected verbs.

As pointed out in connection with *Table 3*, the combined acceptance of the different cases indicates intra-speaker variation, namely that many speakers accept both Acc. and Dat. objects with these verbs. It is interesting to note, however, that there is a considerable difference between the verbs involved. Thus there is hardly any intra-speaker variation in object case selection of *rústa* 'demolish' but large variation in the case of *framlengja* 'extend', for instance.

It is also important to note in this connection that we were able to discover another important difference in case variation by combining the methods of elicitation we used. Recall that we collected information about case marking preferences in two different ways: The speakers were both asked to evaluate sentences and to fill in blanks in short narratives. As discussed in connection with *Table 2* above, this combination of methods showed in some instances that speakers who accepted a Dat. subject with a given verb might nevertheless select an Acc. subject when filling in the blanks. *Table 6* shows an interesting difference between the subject case of *langa* 'want' and object case of *rústa* 'demolish' in this respect:

| Verb | Dat. selected | Dat. accepted |
|--|---------------|---------------|
| Object case w. <i>rústa</i> 'demolish' | 88.1% | 83.6% |
| Subject case w. <i>langa</i> 'want' | 19.2% | 68.2% |

As *Table 6* shows, there is hardly any difference between the selection rate and acceptance rate of Dat. objects with *rústa* 'demolish' whereas there is considerable difference between corresponding figures for the subject case of *langa* 'want'. This means, then, that if a speaker accepts Dat. object case with *rústa* (s)he will also use it but a speaker who accepts Dat. subject case with *langa* may not use it. The other side of the coin is that a speaker who does

not select Dat. subject case with *langa* may nevertheless accept it. The reason for this difference is that speakers are much less likely to be bidialectal with respect to object case of *rústa* (cf. also the figures in *Table 5*) but they may very well be w.r.t. subject case of *langa*, as already discussed. Assuming that the selection task under discussion mirrors production (you select the form that you would use), this clearly shows that production data do not necessarily tell us the whole story about the internalized grammar of speakers. Facts of this sort should thus help dispel the myth that it would be best if we could rely on “natural data” in syntactic studies (“real examples” found in spontaneous speech).

With this in mind, we can now turn to a more detailed investigation of intra-speaker variation in case marking.

4.2 Intra-speaker variation in case marking

The differences in intra-speaker variation between *langa* 'want' and *rústa* 'demolish' can be visualized in histograms like the following (here we are only considering two examples with Dat. arguments for each verb):

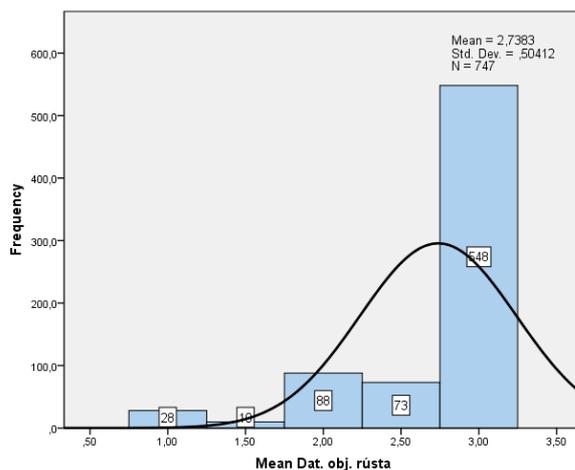


Figure 5: Distribution of means for Dat. objects of *rústa* 'demolish'.

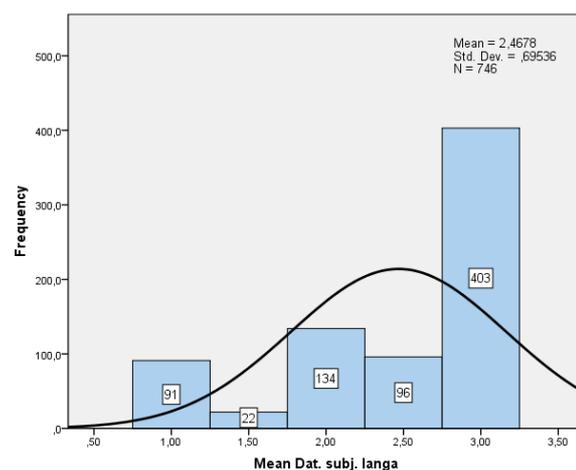


Figure 6: Distribution of means for Dat. subjects of *langa* 'want'.

If there was no intra-speaker variation, the only bars that would show up on these histograms would be the ones for 1 (= no Dat. examples found natural) and 3 (= all Dat. examples found natural). Thus the values in between show how many speakers accepted some but not all of the examples under consideration, i.e. for how many speakers some sort of intra-speaker variation was found. By adding up the numbers we can see that this holds for 171 (out of 747) speakers for Dat. objects of *rústa* and for 252 (out of 746) speakers for Dat. subjects of *langa*.

Now the intra-speaker variation observed in *Figures 5–6* may not look all that impressive. After all, the majority of the speakers found all of the Dat. examples natural, both for objects of *rústa* 'demolish' and subjects of *langa* 'want'. But if we add more examples and include the verb *vanta* 'need' in the picture, the other most typical DS verb in Icelandic, we see that the intra-speaker variation is actually more pervasive. This is shown in *Figure 7*. In *Figure 8* I have added comparable information for comparable verbs in Faroese:

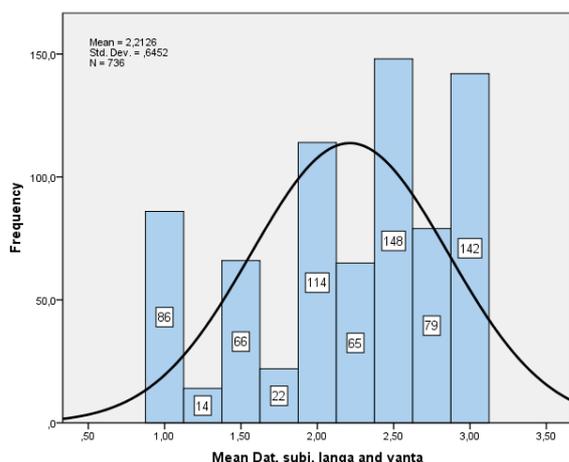


Figure 7: Distribution of means for Dat. subjects of Icel. *langa* and *vanta*.

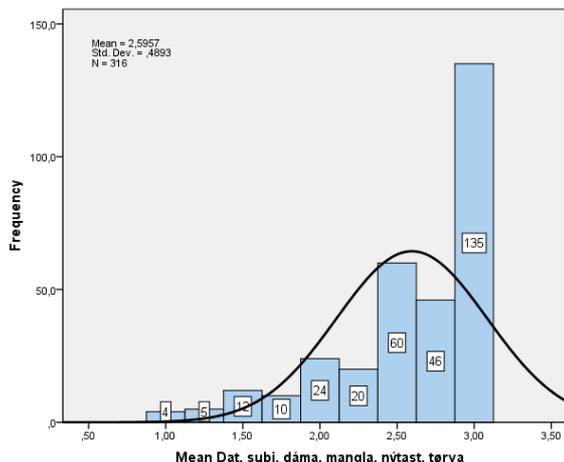


Figure 8: Distribution of means for Dat. subjects of Far. *dáma*, *mangla*, *nýttast* and *tøvva*.

What these figures reveal is that there is much more intra-speaker variation in subject case marking of *langa* 'want' and *vanta* 'need' in Icelandic than there is for the comparable verbs *dáma* 'like', *mangla* 'lack', *nýttast* 'need' and *tøvva* 'need' in Faroese. As Figure 8 shows, about 43% of the Faroese speakers tested (136 out of a total of N 316 in this instance) accepted all examples of Dat. subjects with the verbs in question whereas only about 19% of the Icelandic speakers accepted all the Dat. subject examples with *langa* and *vanta*.

It is important to note in this connection, however, that the historical development has been rather different in the two languages in this area. The Icelandic verbs *langa* and *vanta* originally took Acc. subjects and the Dat. subjects are an innovation still frowned upon in schools. This is the so-called Dative Sickness, a change in progress, as we have seen. In Faroese, on the other hand, the Acc. subjects have virtually disappeared (see e.g. Thráinsson et al. 2004:252ff., Jónsson and Eythórssón 2005). Here, too, many verbs that originally took Acc. subjects now take Dat. subjects (including *vanta*; the verb *langa* is not used anymore in Faroese). But a further development is now taking place in Faroese as Nom. subjects are replacing Dat. ones for many verbs. Now the Faroese verbs *dáma* 'like' and *nýttast* 'need' presumably took Dat. from early on (as their Icelandic counterparts still do). The verb *tøvva* 'need', on the other hand, corresponds to Old Norse (and Icelandic) *þurfa*, which took a Nom. subject, so here the Dat. subject is an innovation. The same is presumably true for *mangla* 'lack', which is apparently a loanword from Danish (d. *mangle*). Thus the Dat. case marking of the subject of the last two verbs indicates a strong tendency to regularize thematic case marking of a subclass of experiencer verbs in earlier stages of Faroese, although it is now beginning to give way to structural Nom. case.

As a final illustration of intra-speaker variation in case marking, consider the following: There is a strong correlation between Dative Sickness and age in Icelandic. The prediction is then that this should be reflected in different intra-speaker variation patterns for different generations. As Figures 9 and 10 show, this prediction is borne out:

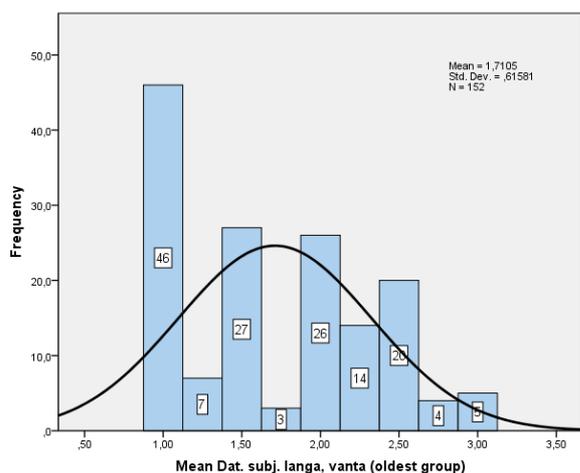


Figure 9: Distribution of means for Dat. subj. of *langa* and *vanta* for the oldest group.

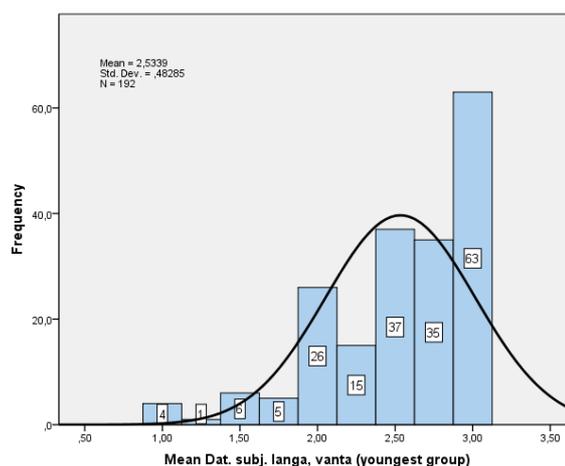


Figure 10: Distribution of means for Dat. subj. of *langa* and *vanta* for the youngest group.

As these figures show, the means cluster to the left (very little Dative Sickness, 46 subjects with none at all) for the oldest group whereas they cluster to the right (more Dative Sickness, 63 subjects who found all the examples natural) for the youngest group.

We have now seen considerable evidence for the existence of widespread intra-speaker variation in case marking. As pointed out above, some theoretical linguists would prefer that intra-speaker variation did not exist at all. Hence it is important to show that it is not at all restricted to case marking nor is the variation observed here an artifact of the research methods used. So we will turn to a totally different area in the next subsection.

4.3 Intra-speaker variation in phonology

One of the rather unexpected results of the survey of phonological variation in Icelandic mentioned above (RÍN, see e.g. Thráinsson and Árnason 1992, Árnason and Thráinsson 2003) was the extensive intra-speaker variation found for many of the phonological variables investigated. It should be emphasized here that this variation was found even when the situation and social context was held constant so it does not have anything to do with socially conditioned dialects. The extent to which it shows up depends, however, on the nature of the phonological variables in question. It would take us too far afield to go into this in any detail, but a demonstration of two cases should suffice to show the similarities to the intra-speaker variation observed in case marking.

Although it has proved to be very difficult to find any geographical variation in Icelandic syntax, such variation can still be found in Icelandic phonology. Two phonological features that basically characterize North-Eastern Icelandic are voiced sonorants (i.e. /l,m,n/) and voiced /ð/ before /p,t,k/ (the so-called voiced pronunciation) and aspirated stops after long vowels (the so-called hard pronunciation). The intra-speaker variation of these features can be seen in *Figures 11* and *12* (note that here the means range from 100 (= does not occur at all) to 200 (= occurs in every possible instance)):

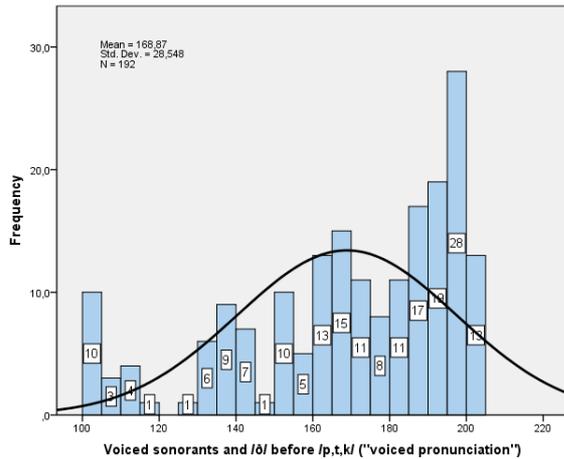


Figure 11: Distribution of means for voiced sonorants and /ð/ before /p,t,k/ in an area in NE Iceland (N = 192).

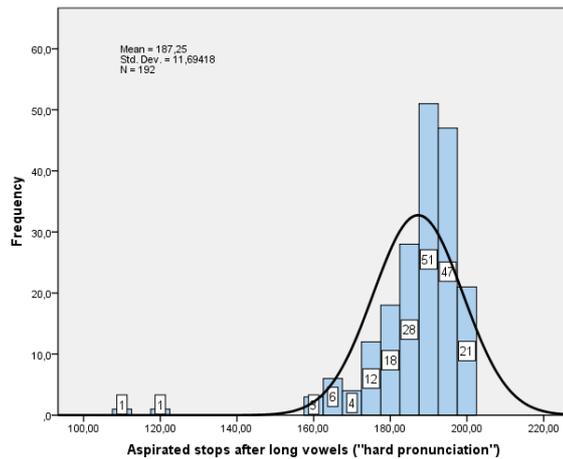


Figure 12: Distribution of means for aspirated stops after long vowels in an area in NE Iceland (N = 192).

These figures show that there is much greater intra-speaker variation in the voiced pronunciation than there is in the hard pronunciation. This reflects the fact that the former is on the way out whereas the latter is relatively stable. The pattern observed here is remarkably like the one found in the syntax although the data come from very different parts of the grammar and the method of elicitation was radically different (the results from the phonological survey are all based on production data elicited in structured interviews). Hence it is clear that the intra-speaker variation observed in the syntax is neither an artifact of the elicitation method used nor a special characteristic of case marking.

5. Conclusion

I would now like to summarize the main results and claims of this paper as follows:

- (10) a. Intra-speaker variation is common and pervasive in those aspects of grammar that are undergoing change.
- b. Intra-speaker variation can be observed in (spontaneous) speech production but sometimes even more clearly in (syntactic) judgments.
- c. Intra-speaker variation is probably by nature a transitional stage, caused by inconsistent and conflicting input, but it needs to be taken seriously in models of grammar.

By the last point I want to claim that when the input (or the primary linguistic data, PLD) is relatively consistent and uniform, as it at least occasionally is, there is no reason to expect intra-speaker variation. Hence we should find areas of grammar without any significant intra-speaker variation. But when the input is inconsistent and even conflicting, as it typically is when a change is ongoing, be it a spreading Dative Sickness or a spreading devoicing of sonorants before /p,t,k/, intra-speaker variation will be a natural outcome. But since most changes eventually reach an end point, the relevant inconsistency in the input should disappear in the end and so should the related intra-speaker variation, at least in theory.

With this in mind, we can now go back to the different ideas about variation discussed in section 2 above and try to determine which ones seem most suitable to deal with facts of the sort discussed in this paper. The concept of variable rules along the lines suggested by

Labov (1972) and his followers seems too performance-oriented on the one hand, as it can make reference to extra-linguistic contextual or situational features. On the other it does not really seem to be an adequate model of the situation found here, where there is a clear difference between speakers in the "amount" of intra-speaker variation characteristic of their grammars under the same circumstances. The same can be said about the proposals of Biberauer and Richards (2006) on the one hand ("the grammar doesn't mind") and Adger and Smith (2010) on the other ("underspecified functional categories"). Besides, it seems that the intra-speaker variation observed is more general than these linguists assume, as it is difficult to argue that it only occurs when the grammar doesn't (or shouldn't) mind or when we can convincingly postulate underspecified functional categories. It seems more promising to seek an account along the lines proposed by Yang (2000, 2004, 2010), which can partly be seen as an extension or further development of the original idea about "competing grammars" proposed by Kroch (1989). Contrary to ideas argued for by Lightfoot (1999) and Hale (2007), for instance, it does not seem that speakers settle relatively early on a grammar that is consistent and uniform in all respects. Rather, it seems that certain areas of their grammar may remain incompletely specified for a long time, even their whole lifetime. But it is not completely unspecified either. Some speakers are more likely than others to use voiced sonorants although they do not do so all the time, and some speakers are more likely than others to use Dat. subjects with experiencer-type verbs although they do not do so all the time. This is a real fact and an important one and it is not just a matter of random performance or performance influenced by some social situation or context. It is a part of these speakers' competence and it can be reflected in their (syntactic or phonological) performance and also in their evaluation or judgments of sentences. Hence it needs to be taken into account in our models of linguistic competence, even if it means that we cannot allow ourselves to deal only with ideal speakers but will have to consider other speakers as well.

Finally, there are probably several ways of modeling this aspect of our competence. One question is whether or to what extent this can be related to parameter settings (cf. the discussions in Yang's work referred to above). I find it likely that this may vary from case to case (no pun intended). One way to find out is to investigate whether there is any relationship between variation in more than one area, so to speak. If there is such a relationship or correlation, then one might want to look for a suitable parameter. At present I do not know if there is any correlation between variation in case marking of the sort reported on here and anything else in the grammars of the speakers involved. But I do know that although there is considerable inter-speaker and intra-speaker variation in verb placement in Faroese, this variation correlates to some extent with available subject position, availability of transitive expletives and probably also Stylistic Fronting. This suggests that some sort of parametric variation is involved, but it would take us too far afield to go into this here (see Thráinsson 2010b).

References

- Adger, David, and Jennifer Smith. 2010. Variation in agreement: A lexical feature-based approach. *Lingua* 120:1109–1134.
- Árnason, Kristján, and Höskuldur Thráinsson. 2003. Fonologiske dialekttræk på Island: Generationer og geografiske områder. In Gunnstein Akselberg, Anne Marit Bødal and Helge Sandøy (eds): *Nordisk dialektologi*, pp. 151–196. Novus, Oslo.

- Barðdal, Jóhanna. 2001. *Case in Icelandic — A Synchronic, Diachronic and Comparative Approach*. Doctoral dissertation, University of Lund.
- Biberauer, Theresa, and Marc Richards. 2006. True optionality: When the grammar doesn't mind. Cedric Boeckx (ed.): *Minimalist Essays*, pp. 35–66. Benjamins, Amsterdam.
- Cedergren, Henrietta, and David Sankoff. 1974. Variable rules: Performance as a statistical reflection of competence. *Language* 50:333–355.
- Chomsky, Noam. 1995. *The Minimalist Program*. MIT Press, Cambridge.
- Cornips, Leonie, and Cecilia Poletto. 2005. On standardizing syntactic elicitation techniques (part 1). *Lingua* 115:939–957.
- Friðriksson, Finnur. 2008. *Language Change vs. Stability in Conservative Language Communities: A Case Study of Icelandic*. Doctoral dissertation, University of Gothenburg, Gothenburg.
- Hale, Mark. 2007. *Historical Linguistics: Theory and Method*. Blackwell, Oxford.
- Henry, Alison. 1995. *Belfast English and Standard English: Dialect Variation and Parameter Setting*. Oxford University Press, Oxford.
- Henry, Alison. 2005a. Idiolectal variation and syntactic theory. In Leonie Cornips and Karen P. Corrigan (eds.): *Syntax and Variation: Reconciling the Biological and the Social*, pp. 109–122. John Benjamins, Amsterdam.
- Henry, Alison. 2005b. Non-standard dialects and linguistic data. *Lingua* 115:1599–1617.
- Jónsson, Jóhannes Gísli, and Thórhallur Eythórsson. 2003. Breytingar á frumlagsfalli í íslensku. [Changes in subject case in Icelandic.] *Íslenskt mál* 25:7–40.
- Jónsson, Jóhannes Gísli, and Thórhallur Eythórsson. 2005. Variation and change in subject case marking in Insular Scandinavian. *Nordic Journal of Linguistics* 28(2):223–245.
- Kroch, Anthony S. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1:199–244.
- Kroch, Anthony S. 2001. Syntactic change. In Mark Baltin and Chris Collins (eds): *The Handbook of Contemporary Syntactic Theory*, pp. 699–729. Blackwell, Oxford.
- Labov, 1972. *Sociolinguistic Patterns*. University of Pennsylvania Press, Philadelphia.
- Lightfoot, D., 1999. *The Development of Language*. Blackwell, Oxford.
- Maling, Joan. 2002. Það rignir þágufalli á Íslandi: Verbs with dative objects in Icelandic. *Íslenskt mál* 24:31–105.
- Maling, Joan, and Sigríður Sigurjónsdóttir. 2002. The "New Impersonal" construction in Icelandic. *Journal of Comparative Germanic Linguistics* 5:97–142.
- Schütze, Carson T. 1996. *The Empirical Base of Linguistics: Grammatical Judgments and Linguistic Methodology*. University of Chicago Press, Chicago.
- Svavarsdóttir, Ásta. 1982. Þágufallssýki. Breytingar á fallnotkun í frumlagssæti ópersónulegra setninga. [Dative Sickness ...] *Íslenskt mál* 4:19–62.
- Svavarsdóttir, Ásta. 2006. Texti, tal og tilraunir. Um efnivið og aðferðir í tilbrigðarannsóknunum. [Text, spoken language and experiments ...] Paper presented at Hugvísindafélag, University of Iceland, November 3, 2006.
- Svavarsdóttir, Ásta. 2010. Þágufallshneigð í sjón og raun. [Dative sickness in appearance and in reality.] Chapter 4 in Thráinsson (ed.) 2010.
- Svenonius, Peter. 2002. Icelandic Case and the Structure of Events. *Journal of Comparative Germanic Linguistics* 5:197–225.
- Thráinsson, Höskuldur. 2007. *The Syntax of Icelandic*. Cambridge University Press, Cambridge.
- Thráinsson, Höskuldur. 2010. Variation and parametric correlations in Faroese. Ms. University of Iceland, Reykjavík.

- Thráinsson, Höskuldur, Hjalmar P. Petersen, Jógvan í Lon Jacobsen and Zakaris Svabo Hansen. 2004. *Faroese. An Overview and Reference Grammar*. Føroya Fróðskaparfelag, Tórshavn.
- Thráinsson, Höskuldur, and Kristján Árnason. 1992. Phonological variation in 20th century Icelandic. *Íslenskt mál* 14:89–128.
- Thráinsson, Höskuldur (ed.). 2010. Tilbrigði í íslenskri setningagerð. [Variation in Icelandic Syntax.] Ms., University of Iceland.
- Trudgill, Peter. 1986. *Dialects in Contact*. Blackwell, Oxford.
- Woolford, Ellen. 2006. Lexical case, inherent case and argument structure. *Linguistic Inquiry* 37:111–130.
- Yang, Charles. 2000. Internal and external forces in language change. *Language Variation and Change* 12:231–250.
- Yang, Charles. 2004. Universal Grammar, statistics or both? *Trends in Cognitive Sciences* 8, 10:451–456.
- Yang, Charles. 2010. Three factors in language variation. *Lingua* 120, 5:1160–1177.

*Höskuldur Thráinsson
Institute of Linguistics
University of Iceland
hoski@hi.is*