

# The Icelandic (Pilot) Project in ScanDiaSyn

Höskuldur Thráinsson, Ásgrímur Angantýsson, Ásta Svavarsdóttir, Thórhallur Eythórsson, Jóhannes Gísli Jónsson<sup>1</sup>  
*University of Iceland*

## Abstract:

In this paper we outline the Icelandic research plans in the Scandinavian Dialect Syntax project and explain why we have made these plans the way we have. We begin by reporting on a pilot project that was conducted in Iceland 2004-2005, explain its nature and describe the resulting plans. As will be seen, our research project includes the collection and analysis of spoken language corpora (“spontaneous speech” of different kinds), collection of syntactic material by using different elicitation techniques (including written questionnaires and interviews), and the comparison of this material. The spoken language corpora are listed and described in the second section of the paper. In the third section we describe how our present (and future) work relates to some previous work done on syntactic variation in Icelandic (and Faroese) and offer some thoughts on the nature of syntactic variation in general.

## 1. The pilot project: Purpose, methods and main results

### *1.1 What was known about syntactic variation in Icelandic*

When the Icelandic research team represented at the grand meeting in Leikanger became a member of the ScanDiaSyn network, relatively little was known about syntactic variation in Icelandic. The situation (“state of the art”) can be summarized as in (1):

---

<sup>1</sup> Our research projects have been supported by the Icelandic Centre for Research and the Research Fund of the University of Iceland. This support is hereby gratefully acknowledged, as is the general support that the ScanDiaSyn Network has received. In this report Höskuldur and Ásgrímur are mainly responsible for section 1, Ásta for section 2, Thórhallur for sections 3.1. and 3.2, and Jóhannes for section 3.3. We blame all errors on each other.

- (1) a. Virtually nothing was known about regional syntactic variation.  
 b. Some evidence for “syntactic change in progress” had been found.  
 c. Some syntactic “innovations” were believed to have social correlations (mainly depending on age groups but possibly also on level of education).  
 d. No evidence for syntactic dialects correlating with morphological (or morpho-phonemic) differences had ever been found.

Points (1b, c) mainly have to do with case marking of the subject (more specifically the so-called Nominative Shift (ND) and Dative Shift (DS)) on the one hand and the New Passive (or New Impersonal) on the other. These phenomena had already been investigated in some detail as will be discussed in section 3.1 below so there is no need to dwell on them here.

Points (1a) and (1d) should be emphasized, however, since Icelandic is rather different from Mainland Scandinavian in this respect. Although syntactic variation has been much less extensively studied than, say, phonological and lexical variation all over Scandinavia, linguists have for a long time known of the existence of certain regional variation in the syntax of Danish, Norwegian and Swedish: some of this variation has been described as can be seen in bibliographies on Scandinavian dialect syntax accessible from the ScanDiaSyn home page (see <http://uit.no/scandiasyn/publications/>). In Iceland, on the other hand, nobody, not even linguists, could tell you about a regional syntactic variant (cf. (1a), possibly with one minuscule exception: It has been noticed that people in a couple of towns in Northern Iceland normally say things like (2a) while (2b, c) are the general rule (with (2c) somewhat more formal):

- (2) a. Petta er bíll-inn Jóns.  
*this is car-the John.G*  
 b. Petta er bíll-inn hans Jóns.  
*this is car-the his.G John.G*  
 c. Petta er bíll Jóns.  
*this is car John.G*  
 ‘This is John’s car.’

As can be seen here, this variation has to do with the interaction of the definite article and the possessive genitive of proper nouns like *Jón* (the same holds for kinship terms such as *mamma* ‘mom’ and *pabbi* ‘dad,’ for instance). This phenomenon has never been studied in any detail but some people knew about

it anyway before we did our pilot study. But it cannot be said to be common knowledge, and most speakers of Icelandic would undoubtedly tell you that there is no regional syntactic variation in Icelandic.

Because of this, it almost goes without saying (but needs to be said here anyway!) that when we did our pilot study, nobody knew of any syntactic variation in Icelandic correlating with morphological or morphophonemic differences (cf. (1d)). There were thus no reports, not even rumors, that certain speakers had, say, reduced verbal or adjectival or nominal inflection that could thus possibly correlate with some syntactic phenomena, nor a different gender system or anything of that nature. This, too, is quite different from the situation in Mainland Scandinavia, as is well known.

As the statements in (1b, c) suggest, on the other hand, we knew of some syntactic variation in Icelandic and we had reasons to believe that it might correlate with the age of the speakers and possibly also to some extent with their level of education (or perhaps the level of the education of their parents, cf. section 3.1 below). We also suspected that there might be some regional variation in the syntax of Icelandic, although it had not been noticed or described. We felt, however, that it was necessary for us to get a better idea of the nature and extent of syntactic variation in Icelandic before embarking on a major study of the type planned in ScanDiaSyn. It was clear to us, for instance, that if there turned out to be only really minor regional differences in Icelandic syntax but some age-related variation, we might need to use elicitation methods that were partially different from those that had been used for example in the study of regional syntactic dialects in Northern Italy or in the Netherlands, although these seemed to make sense for Mainland Scandinavia. At the same time, we wanted our main study (“the real project”) to be comparable to other ScanDiaSyn studies, to the extent that this is possible.

### *1.2 The purpose and nature of the pilot study*

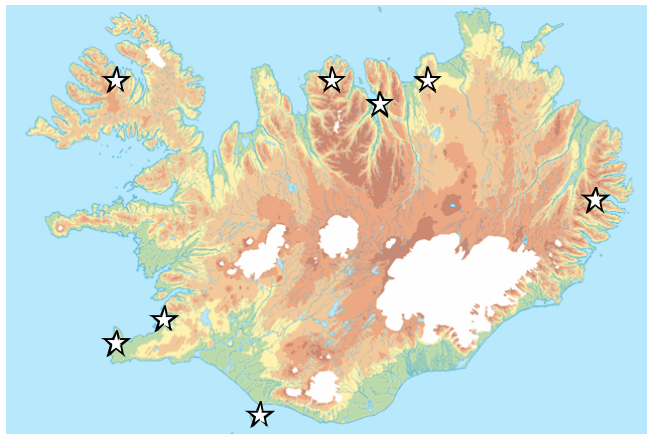
For the reasons described in section 1.1, we designed a pilot study that had the following main goals:

- (3) a. To look for interesting variation in Icelandic syntax to investigate in more detail in the “real project.”
- b. To compare some methods of data collection with the “real project” in mind.

Although we only knew for certain of a handful of phenomena showing syntactic variation in Icelandic (mainly believed to be age-related), we

suspected (and hoped) that there might be more syntactic variation in the language than commonly believed. We even hoped to be able to find evidence for regional variation that people had not noticed. Our first aim was thus to look for such variation (cf. (3a)). We did this by designing a written questionnaire that included a relatively large number of syntactic constructions of various kinds, about 30 in all. For a few of these we had reasons to expect some variation depending on age groups at least, but for others it was more of a shot in the dark. The questionnaire we presented to a total of 187 speakers from 8 different locations in Iceland, relatively evenly distributed w.r.t. gender, age and social class. The subjects came from three age groups (20-25, 40-45, 65-70) and the 8 locations can be seen on this map of Iceland:

(4)



*Map of Iceland, showing the locations included in the Icelandic pilot study.*

By having a relatively large geographical spread and including subjects from three different age groups we hoped to be able to catch both geographical and age-related variation. As will be shown in the following subsection, about half of the constructions included showed evidence for some variation.

In the pilot project, we systematically compared the following elicitation methods:

- (5) a. written questionnaires involving
1. acceptability judgments (slightly different types, cf. below)
  2. fill-ins (described below)
- b. interviews centering around a subset of the constructions included in the questionnaires

All 187 subjects answered the written questionnaires and roughly 50% of them were also interviewed. Research assistants in each of the 7 locations outside Reykjavík were in charge of picking out the subjects in their area. They were instructed to select an equal number of male and female participants for each age group (4 males and 4 females) and asked to try to avoid a skewed spread with respect to such social categories as occupation, education etc. All the subjects were then asked to show up at the same time at a given location, usually a school, and there a member of the research group would hand out the questionnaires and give instructions.<sup>2</sup> The subjects would then answer the questionnaires on location and the interviews were typically conducted the same day, usually after a break.

The front page of the questionnaire contained a couple of background questions (gender, age, occupation, education, place where the subject grew up etc.) and brief instructions that the researcher read aloud before the beginning of the session. They read as follows (roughly translated):

- (6) Icelandic is more varied than we normally realize. Modern Icelandic is different from Old Icelandic, people in different parts of the country pronounce words in different ways, teenagers do not speak the way their grandparents speak, etc. In addition, there is often a choice between variants. Thus most speakers can either say *She put all the stuff **down*** or *She put **down** all the stuff*. In other cases only one variant is acceptable. Thus all speakers would presumably say *She put it **down*** and nobody *\*She put **down** it*.

The purpose of this study is to investigate a number of syntactic variants to see if they vary from one region to another or between age groups. In this connection it is important to keep in mind that we are NOT interested in our subjects' beliefs about **good (or correct) and bad (or wrong) usage** or what they have been **told** about such differences but only what they **think they themselves would normally say or could**

---

<sup>2</sup>The graduate students Ásgrímur Angantýsson and Halldóra Björt Ewen were in charge of this.

**say**. Sometimes we also ask what the participants in this study think that **other speakers say** or what they have **heard**. – Note in particular that we are not checking people’s knowledge of what is commonly considered proper style or appropriate in the written language, such as when people are instructed that one should write X rather than Y [here actual (and familiar) examples of X and Y were given]. We are mainly interested in **the spoken language**. The participants can, however, point it out in a separate comment if they believe a particular variant mainly occurs in the written language.

Finally note that all the results will be treated anonymously, but the subjects are asked to give their names so that they can be contacted during the processing of the data if questions should arise.

The **written questionnaire** was typically some 12 pages. It usually took the subjects some 30 minutes to complete the questionnaire and it involved the three different types of tasks listed in (7). The different types will be described in some detail below:

- (7) a. Straight grammaticality judgments of individual sentences (about 130 in all, divided into three groups A, B, C).
- b. Choice between two or more alternatives (some 15 sets)
- c. Fill-ins (three types).

For the first type of task we tried out the following three variants in most locations:<sup>3</sup>

---

<sup>3</sup>Some studies (see Sigurjónsdóttir and Maling 2001, Maling and Sigurjónsdóttir 2002) have used questionnaires giving only two alternatives, acceptable and unacceptable. The motivation for this would be to try to force the subjects to make up their mind and not be too tempted to check the “questionable alternatives” too often. The problem is, however, that some sentences are in fact “questionable” rather than perfect or impossible. Hence we opted for more than two alternatives and the subjects were in general not overly inclined to pick the middle one(s).

(8) Variant I:

Put an X in the appropriate column:

**Yes** = A natural sentence. This is a sentence that I could easily say.

**?** = A doubtful sentence. This is a sentence that I could hardly say.

**No** = An unacceptable sentence. I could **not** say this.

You can also add a short comment or explanation in the column **Comments** if you feel that this is necessary. Remember that we are interested in what you **really think yourself** – this is not a test.

		Yes	?	No	Comments
1	I think that this is my car				

(9) Variant II:

This variant was exactly the same as variant I except that in the instructions we substituted ‘one’ for ‘I,’ i.e. (emphasis added here):

**Yes** = A natural sentence. This is a sentence that **one** could easily say.

etc.

(10) Variant III:

Here the subjects were asked to chose between four types of judgments instead of three and the alternatives were described as follows:

**1** = A very natural sentence. I could **easily** say this.

**2** = A rather natural sentence. I could **probably** say this.

**3** = A rather unnatural sentence. I could **probably not** say this.

**4** = An unacceptable sentence. I could **definitely not** say this.

		1	2	3	4	Comments
1	I think that this is my car					

As described above, variant I is slightly more “personal” than variant II, using phrases like ‘I could easily say this’ as opposed to ‘**One** could easily say this.’ In a recent study (Sigurjónsdóttir and Maling 2001, Maling and Sigurjónsdóttir 2002), some of the subjects commented that they felt that the more impersonal question could be interpreted as a question about what others could say, not just the subjects themselves. Since we were mainly interested in the subjects’ own intuitions, we wanted to make it as clear as possible that this was what we were after (cf. also the general introduction read to all the subjects and the explanation of the meaning of the different alternatives offered in

the acceptability judgments). We wanted to make sure, however, that this more direct way of asking for the judgments ('I could ...') did not put the subjects off in any way so we also included the more indirect one ('One could ...') for comparative purposes. The results did not reveal any problems with the more direct way so that is how we will ask the questions in the actual project.

The difference between variants I and III involves just one question mark as the "doubtful" category vs. two such categories, i.e. 'probably' and 'probably not.' The results of our pilot study do not indicate so far that the added complication involved in variant III yields any interesting results over variant I. Hence we will presumably just include three options in our questionnaire in the real study, i.e. *yes*, *?*, *no*, although we are still doing a more detailed comparison of these variants.

The second type of task on the written questionnaires involved the choice between two or more alternatives. One variant is exemplified in (11), where the only difference between the alternatives is the case marking of the subject (Npl, Dpl, Apl, respectively) but the sentences mean 'The kids want ice cream':

- (11) Check the variant of the following sentences that **you would use** in normal speech:
- a.  Krakkarnir langa í ís.
  - b.  Krökkunum langar í ís.
  - c.  Krakkana langar í ís.

The subjects were also presented with the same kinds of alternatives and asked which alternative they thought **people their age** would be most likely to use in normal speech and or which variant **they felt was most natural** in normal speech. While the subjects dutifully picked one variant in these tasks in most instances, their answers sometimes raised questions that we would have liked to be able to ask them directly. The main reason was the fact that we did not really know what the subjects thought of the alternatives that they did not pick or why they preferred one variant over the other when asked to choose, etc. Thus our general conclusion was that this type of task (choice between alternatives, ranking of alternatives) would be better suited for the interviews than for the written questionnaires. We will take this into account in the planning of the real project.

The fill-ins were of three different types. The first type is illustrated in (12):



- (12) The order of words in Icelandic sentences may vary. Below you will be given a word in parentheses and your task is to put it in the position (gap) in the following sentence where it most naturally belongs in your opinion. If you think that it could occur in either place, write it in both places and underline the position that you find more natural:

(*never*) This is the only show that I \_\_\_\_\_ have \_\_\_\_\_ watched.

Although the subjects were given two choices in tasks of this kind, and also the opportunity to rank them, they usually picked just one. Other than that, these fill-ins worked pretty well and we will presumably use similar tasks to break up the direct judgment tasks in the written questionnaires of the real study.

A second type of fill-in task is illustrated in (13) (to the extent that this can be done appropriately in English):

- (13) In the following sentences you are asked to select one of the two words or word forms enclosed in parentheses and put it in the gaps. If you feel that either one would do, write them both in the relevant position and underline the one you prefer:

(*him/himself*) I showed John a picture of \_\_\_\_\_ (*him/himself* = John)

(*seem/seems*) There \_\_\_\_\_ to be mice in the bathtub.

Although here again the subjects picked both alternatives less frequently than we had expected, fill-ins of this kind worked pretty well in general and in some instances provided interesting results for comparison with those of the direct judgment tasks. Hence, we expect to use fill-ins of this kind too to break up otherwise monotonous judgment tasks in the real study and thus avoid potential repetition effects that have been reported in connection with plain judgment tasks with long lists of examples (see e.g. the discussion in Cornips and Poletto 2005).

The third type of fill-ins was first used for linguistic purposes in Svavarsdóttir's study (1982) of Dative Substitution in Icelandic and it has been used several times since then for similar purposes in studies of subject case marking in Icelandic and Faroese (see Jónsson and Eythórsson 2005 and references cited there). Here the basic idea is to present the subject with a story involving one main character as the agent/experiencer of most of the predicates involved, leave the relevant syntactic position blank in most cases and ask the subjects to fill it with the relevant form as briefly illustrated in (14) (this is to be understood as "Icelandic with English words"):

- (14) Here is a short narrative about **Rachel**. Her name is only mentioned a few times in the story but in the rest we have blanks where it would be more natural to use pronouns like *she/her* etc. Write these words in the blanks as appropriate.

**Rachel** is 11 years old and lives by the sea. \_\_\_\_\_ is going on a fishing trip with her father. \_\_\_\_\_ looks very much forward to this because \_\_\_\_\_ likes to go on boat trips in her father's boat ...

The Icelandic predicates are selected in such a way that sometimes a nominative subject (*hún* 'she') would be expected, sometimes an accusative one (*hana* 'her.A') and sometimes dative (*henni* 'her.D'). This is one of the few areas where we knew of some variation in Icelandic and it had been studied in some detail before (see the discussion of Dative Substitution and Nominative Substitution in section 3.1 below and references cited there).

The kind of fill-in task described here has worked quite well in these studies, but it is probably fair to say that it is best suited for studies involving young speakers such as school children, as the studies referred to typically did. One of the reasons it has been so popular is that the so-called Dative Substitution (or Dative Sickness) in Icelandic has been frowned upon in schools and thus stigmatized to some extent. As a result it is not clear that all speakers would answer honestly when asked directly to judge sentences involving predicates prone to Dative Substitution, whereas they might reveal their actual usage better in tasks of this sort. Interestingly, however, we found that our subjects were not in general too preoccupied with normative judgments.

Nevertheless, we will probably continue to complement direct judgment tasks with other tasks for comparative purposes when there is reason to suspect interference of normative judgments. In such instances fill-ins are sometimes a possible alternative – or supplement, rather, since in isolation they normally fail to tell you how the subjects evaluate the alternatives which they do not insert.<sup>4</sup>

Fortunately, there are relatively few areas of Icelandic syntax where we need to worry about normative judgments or the feeling that a given variant is a non-standard dialect variant as opposed to an officially accepted standard.

---

<sup>4</sup> As mentioned above, some types of fill-ins are perhaps better suited for younger speakers than for adults, e.g. the last type described here. This may turn out to be particularly useful since there is some reason to suspect that young subjects (e.g. 12-14 years old) give less reliable answers than older speakers when presented with written questionnaires asking for grammaticality or acceptability judgments (see Stefánsdóttir 2005).

This should make straight judgment tasks simpler in studies of Icelandic syntax than in some other languages. We will return to this issue at the end of section 1.3.

As mentioned above, the **interviews** centered mostly around questions involving a subset of the constructions investigated in the written questionnaire. While most of the tasks in the written questionnaire involved judgments of single sentences rather than a comparison of alternatives, we used the opportunity to ask the subjects to compare alternatives in the interviews although we also asked for judgments of a few single sentences in the interviews. Thus a typical question in an interview would be like this:

- (15) Which of the following do you find more natural:
- a. Jón missir vinnuna ef hann ekki hættir að drekka.  
*John loses job.the if he not stops to drink*
  - b. Jón missir vinnuna ef hann hættir ekki að drekka.  
*John loses job.the if he stops not to drink*  
'John loses his job if he doesn't stop drinking.'

Here the only difference between the alternatives is the position of the negation *ekki* 'not' in the embedded clause. In the first alternative it precedes the finite verb, as it normally does in embedded clauses in Mainland Scandinavian. In the second variant the negation follows the finite verb, as it typically does in Icelandic embedded clauses as well as in main clauses. When subjects were asked to judge these variants in isolation on the written questionnaire (they did not occur there next to each other), they typically did not like the a-variant (the Mainland Scandinavian order). But if the interviewer varied the stress and intonation in the interview, e.g. by putting an extra stress on the negation in the a-variant in, it turned out that this could influence the judgments and make the subjects more likely than otherwise to accept the a-variant.

Another construction that was evaluated more positively in the interviews than on the written questionnaires was (16b):

- (16) a. Í Bandaríkjunum er fullt af fólki sem á enga peninga.  
*in United-states.the is lots of people that have no money*
- b. Í Bandaríkjunum að þá er fullt af fólki  
*in United-states.the that then is lots of people*  
sem á enga peninga.  
*that have no money*  
'In the US there is lots of people that have no money.'

Here the a-variant is “regular” topicalization of a non-argument whereas the b-variant is a more colloquial variant of the same with an added *að þá* (lit. ‘that then’) between the topicalized element and the rest of the clause. On the written questionnaire most subjects judged the b-variant quite harshly but if it was presented in the interview with an intonation break after the fronted element and before the “extra material” *að þá*, most speakers were willing to accept it.

The two cases just described show that oral interviews are important when asking for judgments of sentences where intonation and stress play a role. In most other cases, however, there turned out to be **only minimal differences** between the judgments we got on the written questionnaire and the judgments found in the oral interviews.

We also experimented with using the oral interviews to ask in more detail about some of the constructions involved, such as about possible semantic differences between variants. This is something that written questionnaires are not particularly useful for. The interview situation makes it possible for the interviewer to react to the subjects’ answers, modify the questions and dig deeper and in that sense they are much more flexible, of course, than the written questionnaires. This is also very useful when subjects are asked to rank different alternatives. But data collection in oral interviews is obviously much more time-consuming than when one is using written questionnaires of the kind we experimented with.

The general conclusion of the methodological part of the pilot study was, therefore, that carefully constructed written questionnaires could be used reliably when one is interested in getting an overview of the situation and doing a relatively large scale comparison between groups of speakers or between geographical areas. As will be shown in the next subsection, our pilot study confirmed certain ideas that we had beforehand about syntactic variation in Icelandic and it also suggested variation and ongoing changes where we did not know about it in advance. If we had, say, simply interviewed two or four speakers in each location, we would not have discovered this. The main reason is that there are no clear syntactic dialects in Iceland. That does not mean, however, that there is no linguistically interesting syntactic variation in Icelandic. Some of the variation we found is very interesting indeed and one of the reasons is that it can teach us something about the nature of syntactic change, how it spreads and sometimes even how it originates (see e.g. the discussion in section 3.1 below).

*1.3 Some examples of the syntactic variation we found*

In (17) we give an overview of the constructions where we found some evidence for variation in our pilot study. The constructions are referred to by names that may not all be familiar to the reader so they will be explained and exemplified below. The variation found is divided into **age-related** variation (variation between the three age groups involved in the study), **areal** variation (variation between geographical areas) and **social** (basically variation having to do with level of education). An x in a given column indicates that **considerable** variation of that kind was found for the construction in question whereas an x in parentheses indicates only **some** variation. It should be emphasized, however, that as the pilot study only involved 187 speakers from 8 different locations (23-24 subjects in each), the variation discovered is probably not statistically significant in our data as of yet. Hence it is possible that some of it may not stand up to scrutiny. Thus at present we treat our results as only suggestive.

(17)	<b>age</b>	<b>area</b>	<b>class</b>
1. Extended progressive	x		
2. Dative (and nominative) substitution	x	(x)	(x)
3. The new passive	x		(x)
4. Agreement with nominative objects	x	x	
5. Prenominal possessive	x		
6. Possessive PPs	x	(x)	
7. Subjunctive in certain embedded clauses	x		(x)
8. Definites in an expletive constructions			
9. Agreement in complex infinitival constructions	x	(x)	
10. Distributive pronouns	x	(x)	
11. Tough-infinitives	x		
12. NP-structure		x	
13. Object case		(x)	(x)
14. Reflexives vs. personal pronouns		(x)	
15. Extra <i>að</i> in subordinate conjunctions		(x)	

As this overview shows, the variation found is mostly age-related. Some readers might suspect that this kind of variation typically reflects a situation where the older generations are preserving some ancient dialectal feature (especially in remote areas perhaps) and the speech of the youngest generation represents a kind of standard language. But this is not at all the case in Icelandic, it seems. In many instances the judgments of the younger generations (or the youngest one) indicate a “change in progress,” i.e. an innovation that is spreading but has not necessarily been accepted by the majority of the speech

community. In many instances the features have not been noticed much and hence they have not become a part of the linguistic discussion in Iceland. Hence it is not problematic to ask for judgments of sentences involving these phenomena – i.e., normative judgments will not interfere with real ones since there is no known or advertised norm. In other instances the innovations have been pointed out and fought against by purists or by teachers and that may create complications for the researchers, but such phenomena are relatively few as already mentioned.

#### *1.4 An illustration of the constructions showing variation*

In this subsection we will illustrate the 15 constructions listed in (17) above and comment briefly on the variation found.

The term **extended progressive** refers to the fact that the progressive construction in Icelandic (auxiliary *vera* ‘be’ + infinitive of the main verb) has been “extended” to classes of verbs that previously did not allow this construction. This includes stative verbs like *skilja* ‘understand’:

- (18) Ég er **ekki að skilja** þetta.  
*I am not to understand.INF this*  
 ‘I don’t understand this.’

This appears to be a relatively recent innovation as most of the older speakers rejected this.

By **Dative Substitution** (DS) we mean substitution of dative case for accusative case on the subject of certain experiencer verbs. Similarly, **Nominative Substitution** (NS) refers to the fact that some speakers use nominative subjects with verbs that used to take accusative subjects. The thematic role of these subjects is theme (cf. the discussion in section 3.1.1 below):

- (19) a. **Mér/Mig** langar í bjór.  
*me.D/A longs for beer*  
 ‘I want beer.’  
 b. **Báturinn/Bátinn** rak á land.  
*boat.the.N/A drifted to land*  
 ‘The boat drifted ashore.’

Here again the variation was mostly age-related – the older speakers were less likely to accept the DS and NS variants. There was also slight variation between the areas considered and some evidence that the level of education of the speakers correlated (inversely) with their willingness to accept the DS and NS variants involving the “substitutions.” As these are among the few variants

that have been stigmatized (especially DS), the constructions call for caution when interpreting the results and here it is also worth trying out different methods of elicitation (as we did, in fact, cf. the discussion above).

The **New Passive** refers to a construction that looks like an expletive (or impersonal) passive of sorts. Consider first the regular Icelandic passive:

- (20) a. Einhver barði litla strákin.  
*somebody hit little.ASG.M boy.the.ASG.M*  
 b. Litli strákurinn var barinn.  
*little.NSG.M boy.the.NSG.M was hit.NSG.M*

As is well known, a nominative subject in the regular passive corresponds to an accusative object in the active. Moreover, if the passive subject is indefinite, an expletive version of the passive is possible and the passive subject can in fact appear in two different positions, keeping its nominative case. This expletive variant is not possible if the passive subject is definite, cf. (21):

- (21) a. Það var lítill strákur/\*litli strákurinn barinn  
*there was little boy/ little boy.the hit.NSG.M*  
 í Kringlunni í gær.  
*in Mall-the yesterday*  
 b. Það var barinn lítill strákur/\*litli strákurinn  
*there was hit.NSG.M little boy/ little boy.the*  
 í Kringlunni í gær.  
*in Mall-the yesterday*  
 ‘A little boy was hit in the Mall yesterday.’

The New Passive is in some respects similar to expletive passives like (21b) but it differs from it in two crucial aspects which a careful study of (22) will reveal (see also the discussion in 3.1.2 below):

- (22) Það var **barið lítinn strá**  
*there was hit.NSG.N little boy.ASG.M*  
 í Kringlunni í gær. (New Passive)  
*in Mall.the yesterday*

First, the noun phrase *lítinn strá* ‘little boy’ is marked for accusative and not nominative as *lítill strákur* ‘little boy’ in the expletive passive in (21b). As a result of this we get the non-agreeing (default) NSG.N form of the participle *barið* in (22) instead of the agreeing NSG.M *barinn* in (21b) (passive participles only agree with nominative subjects). Second, this NP can be definite whereas the passive subject in (21) can only be indefinite.

In our study the younger speakers were most likely to accept the New Passive and there was also some indication that the level of education played a role here too. This is not surprising since the phenomenon is an innovation that has been noticed and commented negatively on in the puristic literature. Accordingly, it is likely that speakers with a higher level of education are more likely than others to pay attention to it and avoid it. But this needs to be investigated further.

There was also some variation with respect to **agreement with nominative objects**. The variants are illustrated in (23):

- (23) Honum **líkuðu**/líkaði ekki skoðanir ráðherrans.  
*him.D liked.PL/SG not opinions.NPL minister.G*  
 ‘He didn’t like the minister’s opinions.’

Nominative objects only occur with a handful of verbs in Icelandic and they all take dative subjects. Here it seemed that the younger speakers were less likely to have the finite verb agree in number with the nominative object than the older speakers were. In addition, this kind of agreement seemed more common in one of the locations visited than in others. If this holds up to further scrutiny, we have here an instance of geographical variation in Icelandic syntax that has not been previously described.

By **prenominal possessives** we here refer to the pattern typically found in Danish, Swedish, and English, for instance, whereas postnominal possessives are clearly more common in Icelandic than in the Mainland Scandinavian varieties mentioned. Examples are given in (24):

- (24) Hún segir **sína** skoðun/skoðun sína.  
*she expresses her opinion/opinion her*  
 ‘She expresses her opinion.’

Prenominal possessives have often been said to be typically licensed by some sort of contrastive comparison (even contrastive stress) in Icelandic, but in our study it turned out that the younger speaker were more likely than the older ones to accept them without reservation in our judgment tasks (and also in fill-in tasks).

In some languages, certain prepositional phrases can alternate with possessives or genitives. In Faroese, for instance, the preposition *hjá* (originally a locative ‘at’) is commonly used to indicate “possession” in a general sense:



- (25) Bilurin hjá Jógvani      kettan hjá mær      húsini      hjá Turið  
*car.the at Jogvan      cat.the at me      houses.the at Turid*  
 ‘Jogvan’s car’      ‘my cat’      ‘Turid’s house’

While Icelandic is still very different from Faroese in this respect, there is some evidence that **possessive PPs** are gaining ground in Icelandic – and this interestingly involves the same preposition as in Faroese, namely *hjá*:

- (26) Stingdu þessu í vasann      **hjá þér.**  
*put      this in pocket.the at you*  
 ‘Put this in your pocket.’

Examples of this sort were mainly accepted by the youngest generation, suggesting that this is an innovation. Nobody has so far suggested any Faroese influence, though.

There is also some evidence for variation in the use of **subjunctive in embedded clauses**. Since the subjunctive is a grammatical category that has been lost as a productive category both in MSc and in Faroese (and in English for that matter), one might have expected it to be generally on the way out in Icelandic – and claims to that effect have sometimes been heard. Interestingly, however, it is used more by the younger generations in certain contexts, such as in *if*-clauses:

- (27) Við förum ekki út ef það sé/      er      rigning.  
*we go not out if it be.SUBJUNCT /is.INDICAT rain*  
 ‘We will not go out if it is raining.’

This may involve some sort of semantic generalization of subjunctive usage. This is one feature that speakers have been aware of for a while and it probably figures in “grammar corrections” in the schools. Hence it is not surprising that there is some evidence for a negative correlation between the level of education of the subjects and their use (in fill-ins) and acceptance of this variant.

While definite subjects are as a rule disallowed in expletive constructions, **definite subjects may occur in a particular type of expletive constructions**. This is illustrated in (28):

- (28) Það er búin      **mjólkinn.**  
*there is finished milk.the*  
 ‘The milk is finished.’

Interestingly, there is some evidence that the older speakers are more likely to accept this than the younger ones. This suggests that this has no relation to the apparent “violation of the definiteness constraint” found in the New Passive

(cf. above) since the New Passive is accepted more generally by the younger generations.

While subject-verb agreement is completely general in Icelandic (agreement between a **nominative** subject and a finite verb, that is), there is some variation w.r.t. agreement with nominative objects as mentioned above. Another type of this kind is **agreement with a nominative object in complex infinitival constructions**, mainly in so-called raising constructions. The difference between this kind of agreement and the type discussed above (#4 in the list in (17)) is the fact that in this case there is a clause-boundary of sorts between the finite verb and the agreement triggering object (the object being in an infinitival clause and the finite verb in the matrix clause):

- (29) Jóni **virðast**/virðist alltaf hafa líkað hestarnir.  
*John.D seem.PL/seems.SG always have liked horses.the.NPL*  
 ‘John always seems to have liked the horses.’

For this construction there is some evidence that the older speakers are more likely to favor the agreeing variant and there is also slight evidence that some geographical variation is involved too.

The varying use of the so-called **distributive possessive pronouns in binominal constructions** is partly illustrated in (30):

- (30) Strákarnir fengu  
*boys.MPL got*  
**sithvort** eplið / hvor sitt epli(ð).  
*their.each.N apple.the.N / each.M his.N apple.the.N*  
 ‘The boys got an apple each.’

In this case it is difficult to give a word-for-word gloss that reflects what is going on. Briefly, it seems that instead of the complex construction *hvor sitt epli(ð)* ‘each his apple’ with *hvor* each agreeing in gender with the subject *stákarnir* ‘boys’ and *sitt* ‘his’ agreeing in gender with the object *epli* ‘apple,’ a new lexical item *sithvort* (lit. “their-each”) has emerged and it agrees in gender with the object (cf. Vangsnes 2002 for a study of this construction in varieties of Norwegian). This is not the whole story (the facts, especially the agreement facts, are more complicated than this), but the relevant point in this connection is that the younger speakers are much more likely to use this new lexical item and disprefer the more complex construction. There is also slight evidence for geographical variation here.

So-called **tough-infinitives** are generally believed to be quite restricted in Icelandic. In our study, slightly more speakers of the older generations accepted sentences like the following one:

- (31) Jóhannes er **erfiður að tefla** við.  
*Johannes is tough to play chess with*

The aspect of **NP-structure** referred to in #12 in (17) has to do with the interplay between possessives and definite nouns, as explained in (2) above. As mentioned there, this was pretty much the only aspect of syntactic structure in Icelandic where we believed we knew of a geographically restricted variant. In brief, this variant had reportedly been heard in two towns in Northern Iceland. We visited one of these towns and the results were as expected. Here the speakers generally accepted sentences like (32a) whereas speakers elsewhere rejected it and only accepted (32b) or (32c) (the last variant being somewhat more formal):

- (32) a. Petta er **bíllinn Jóns**.  
*this is car.the John.GEN*  
 b. Petta er bíllinn hans Jóns.  
*this is car.the his John.GEN*  
 c. Petta er bíll Jóns.  
*this is car John.GEN*  
 ‘This is John’s car.’

While variation of subject case marking had previously been noted and studied, as already mentioned, no systematic survey had been done of **object case marking**. We knew of some anecdotal reports on geographical variation of case marking of individual verbs (t.d. *flýta sér* lit. ‘hurry oneself.DAT’ vs. *flýta sín* lit. ‘hurry oneself.GEN,’ *keyra einhvern/einhverjum* ‘drive somebody.ACC/DAT’ i.e. ‘give somebody a lift’) but we were especially interested in checking dative objects. The reason is that certain regularities involving the thematic role of dative objects have recently been described (see e.g. Barðdal 2001, Maling 2002, Jónsson 2005). One verb we checked was the verb *rústa* ‘destroy’:

- (33) Þau rústuðu **íbúðinni**/íbúðina.  
*they destroyed apartment.the.DAT/ACC*

To our surprise, there was some evidence for geographical (and perhaps even social) variation here so a more systematic investigation of a selected class of dative objects might be in order.

As is well known, **reflexive pronouns** in Icelandic are quite complex. To make a long story short, there is sometimes a choice between reflexive and non-reflexive pronouns in two contexts, as illustrated in (34):

- (34) a. Ég sendi Jóni bækurnar **sínar**/hans.  
*I sent John books.the his.REFL/NON-REFL*
- b. Jón segir að þú hatir **sig**/hann.  
*John says that you hate him.REFL/NON-REFL*

In the a-example we have a non-subject antecedent (the indirect object *Jóni* ‘John’) and in such instances there is some “optionality” in the choice between a reflexive and a non-reflexive pronoun (see e.g. Thráinsson 1991, Reuland and Sigurjónsdóttir 1997 and references cited there). Interestingly, there appeared to be some geographical variation in the preferences for reflexive pronouns in these contexts.

Finally, it has long been possible to stick in **að ‘that’ as an extra element in many subordinations** in Icelandic. The interrogative complementizer *hvort* ‘whether’ is one of these. This is illustrated in (35):

- (35) Ég veit ekki hvort (að) hún kemur.  
*I know not whether (that) she comes*  
 ‘I don’t know whether she is coming.’

Here we found slight geographical variation.

This concludes our survey of the constructions showing variation in our pilot study.

### *1.5 What are our current plans how are they related to the pilot study?*

The purpose of the pilot study was outlined in (3) above, repeated here for convenience:

- (3) a. To look for interesting variation in Icelandic syntax to investigate in more detail in the “real project.”
- b. To compare some methods of data collection with the “real project” in mind.

In the preceding subsections we have described the variation we found. Obviously, it is not all equally interesting. While we have not made a final decision as to which constructions to investigate, the following points will figure in the decision:

- (36) a. How likely is it that a study of the existing variation will help us to a better understanding of the phenomenon?
- b. Do we have any hypotheses that we can test by studying the variation?
- c. Is the construction in question interesting from a comparative Scandinavian point of view – does it e.g. show variation within Scandinavia?

To explain the point in (36a), we can take variation like the reported variation in object case marking of individual verbs like *flýta* ‘hurry’ mentioned above. Since we have no idea what this variation between dative and genitive case marking of the object might be related to, it is not very likely that detailed information about the geographical distribution of each type will be very revealing. Conversely, we believe that a detailed study of the spreading of the New Passive may tell us something about its nature and about how syntactic innovations may spread. Point (36b) hardly need further explanation, but theoretical issues concerning variation will be discussed in section 3.1. With respect to point (36c) we will just remind the reader of the fact that a list of topics to study within ScanDiaSyn is being worked out (and was extensively discussed in Leikanger).

The comparison of the different methods of elicitation was very important for us. Its results have obviously influenced our choice of elicitation methods for the real project. Our current plan is as follows:

- (37) a. To do a rather extensive survey of syntactic variation in Icelandic using written questionnaires. The questionnaires will involve direct evaluation of grammaticality (acceptability), asking the speakers if they themselves can say the relevant sentence and probably offering three alternatives as explained above (*yes*, *?*, *no*), plus a possibility for comments. – In addition, the questionnaire will contain some fill-in tasks.
- b. The written overview will be conducted in some 30 different locations in Iceland and it will involve 4 age groups (15, 20-25, 40-45, 65-70) of 8 subjects each (4 males, 4 females) – a total of 960 subjects. To cover all the constructions we want to include in the overview, we will probably have to do two or three such surveys or sessions.
- c. After we have studied the results of the written overview, we will interview a subset of the subjects. The interviews will be taped and

they will center around questions similar to those included in the overview. We will, however, try to dig deeper into selected issues and frequently ask the speakers to compare variants, both with respect to their semantics and acceptability.

- d. Given the lack of “real” dialects in Icelandic, the taped interviews are not particularly likely to produce very interesting material for a spoken language data bank. We will therefore collect different types of spoken language corpora, both for comparative purposes (comparison with the results from the interviews and written overviews) and in order to get authentic good examples of at least some of the interesting constructions on tape (for our data bank).

Readers familiar with the ScanDiaSyn-related plans elsewhere will notice some differences between these and our plans. The main differences are listed and explained below.

- (38) a. The written survey has a more prominent role in our project than elsewhere. The main reason is the fact that when the results of the written questionnaires were compared to those of the interviews in our pilot study, there were only minimal differences in most instances. Hence we feel that the written questionnaires are in most instances reliable enough to yield interesting and true results. Another reason is the (almost total) lack of syntactic dialects in the usual sense. Hence a common written form of a sentence does not *per se* carry with it any particular notion of standard vs. dialect. There simply is no variation in Icelandic syntax that would correlate with a phonological or morphological feature that would need to be represented in a particular (non-standard) way in the written form. – It also follows from this that there is very little need for any kind of special transcription of the material. The standard written form will do just fine in most instances. If there is a special need to transcribe something, such as a contracted form of a verb and a following unstressed pronoun (which is a general colloquial way of speaking, not a particular dialect), then this can easily be added. This situation is obviously very different from what we find in many places in Mainland Scandinavia – or in the Netherlands and Northern Italy for that matter.
- b. Because the variation we have found evidence for is mostly age-related rather than geographically conditioned, we want to include

more age groups than what is planned for Mainland Scandinavia. So we have added a younger group (the 15 year olds) and a middle group (40-45) to the two that are planned for most other countries. By including a higher number of speakers in the groups we also hope to get statistically significant results in many instances showing us for example in which direction the development is going, i.e. which variants are gaining ground and which ones are on the retreat. This way we should be able to get data that bear on the question of how syntactic change spreads. Since we will also include subjects from different social groups, we might be able to find correlations between certain variants and social variables such as education.

- c. Given this, we can use the interviews for digging deeper into certain questions, testing the relevance of stress and intonation where appropriate (cf. the discussion of the different methods above), etc. The lack of real dialects also makes it unnecessary for us to train “local interviewers” who speak the local dialect and can then conduct the interviews in the appropriate dialect. In fact, this is not just unnecessary for us but totally impossible since there is no such thing as a clear enough local (syntactic) dialect.

The results of your pilot study and the experiments we have done since then have also led to certain modifications of the way we pose our questions. One of the most important modification is that in the real study we will typically give a “context sentence” before each “test sentence.” This helps make it more natural and it also ensures that the subjects are thinking of more or less the same context when judging the sentence. This is illustrated in (39):

(39) Context: Mér gekk illa á prófinu.

‘I did badly on the exam.’

Test sentence: Ég var ekki að skilja neitt í spurningunum.

‘I was not understanding the questions at all.’

Here the real test sentence (involving the “extended progressive” explained above) is preceded by a sentence providing a context.

Another thing we learned from the pilot study is the importance of breaking up standard judgment tasks by other kinds of tasks, such as fill-ins. In the real study we will probably also reverse the order of the test sentences for half of the subjects in each age group.

In general, we feel it is safe to say that we learned several very important lessons from our pilot study and they will help us in the real project.

Finally, it should be mentioned here that since the linguistic situation in the Faroes is in many respects similar to the Icelandic one, the Faroese sub-project of ScanDiaSyn will presumably use very similar methodology as the one we are developing. It also makes sense from a comparative point of view since comparison of Faroese with Icelandic will presumably be an important part of ScanDiaSyn.

## **2. Compilation of spoken language text material**

This section describes the spoken language corpora that we plan to include in our project.

### *2.1 Purposes and origin of material*

One of the aims of the Icelandic project is to compile authentic text material, especially of spoken language in various contexts. The purpose of this material within the project is twofold: First, it will serve as a model for examples and a source of syntactic constructions in designing questionnaires and preparing interviews in the main survey. Second, it will give us an opportunity to compare the results from the survey with constructions that occur naturally in speech and writing.

Whereas written texts are relatively easy to acquire, samples of spoken language are not as readily obtainable. The speech events must be recorded, and the recordings must be transcribed in order to make the material accessible for analysis and search. As the transcription of speech is a very time-consuming task, collaboration between projects and sharing of material is desirable, and we have already reached agreement with a number of projects and institutions that have already collected some spoken language data and transcribed it to a greater or a lesser extent, to use the material for our purposes. Most of this material needs some further preparation, and in exchange for the access we will share any additions in transcriptions and analysis done within our project with the other projects.

### *2.2 Description of the material*

#### *2.2.1 Recent recordings of natural speech*

In 1999–2001, a group of researchers joined forces to establish a corpus of Icelandic spoken language, ISTAL, both for the benefit of spoken language research and of language technology (see Eiríkur Rögnvaldsson 2002). With a grant from the Icelandic Centre for Research, 31 spontaneous conversations



were recorded in various parts of Iceland. They took place in natural and informal settings, each with 2–6 adult participants, male and female. The material consists of approximately 20 hours of digitally recorded speech, together with full transcriptions of the entire material.

The transcription conventions used in the ISTAL-project, is based on normal orthography. Discourse features, such as overlapping, simultaneous speech, interruptions, etc., are coded in a regular way, and deviations in pronunciation, etc., are commented on outside the text proper. Information on the recording (date, length, situation, participants, etc.) and the informants (age, occupation, etc.) is registered at the beginning of each transcription file. This material has been made available for the syntactic variation project, which in return will take care of the final necessary tasks before a general access can be given to the material, for example the elimination of names and other personal features. Furthermore, it was decided to adopt the ISTAL transcription mode for other spoken language material prepared for our project.

Besides ISTAL, a fully normalized transcription of two sets of spoken language material has been completed. The first of these consists of eight group-interviews on language attitudes, made in connection with the Nordic project *Moderne importord i sprogene i Norden* (MIN) (cf. Sandøy, this volume). The interviews represent a structured discussion on lexical borrowing and English influence in Icelandic, between 3 participants plus the interviewer. The participants are adults, 27–36 years old, all living in Reykjavík. The interviews took place in 2003, in rather formal settings, and were digitally recorded. The total amount of speech is about 9 hours, and the interviews had already been roughly transcribed by Hanna Óladóttir for the MIN-project (cf. Óladóttir 2005: 66–70). When the permission had been granted to use the material in a wider context than originally intended, the transcriptions were completed and normalized for inclusion in the corpus.

The other set of data consists of recordings from Alþingi, the Icelandic legislative assembly. They represent a formal register of the spoken language, and as our aim was to get samples of free speech, we chose recordings of unprepared discussion sessions rather than of regular meetings where speakers are more likely to present written speeches. In selecting the material, we also sought to get recordings of sessions where the participants were both male and female, young and old, representing different age groups and both experienced and less experienced speakers. Our material should therefore be relatively representative for free speech within the chosen register. Altogether 52 members of parliament participate in one or more sessions, 39 men and 13 women

(born 1938 to 1979). Meetings of Alþingi are regularly recorded, transcribed, and edited for publication by the parliament staff, and we were provided with digital recordings of extracts from 11 meetings together with the first unedited version of the transcriptions. These were checked against the recordings and normalized by the same method as the other two sets of texts.

Table 1 gives an overview of the three sets of spoken language material, recorded in recent years, that has been completed as described above. Together it consists of approximately 50 hours of digital recordings, fully transcribed in a normalized format. The transcriptions are close to 440 thousand running words (heads, codes and comments included).

Table 1: Three different types of spoken language material to be included.

<i>Material</i>	<i>Type</i>	<i>Formality and involvement</i>	<i>Length of recordings</i> (appr. in hours)	<i>Length of transcriptions</i> (appr. in running words)
ISTAL	Spontaneous conversations	Informal/personal	20	180,705
MISN	Group interviews	Formal/personal	9	86,216
Alþingi	Unprepared speeches	Formal/impersonal	21	169,315

Besides this data there exist other fully or partly transcribed texts that might be added to our corpus. Among these are samples of free speech collected by Finnur Friðriksson for his Ph.d.-project (cf. Friðriksson 2004:171), and some telephone conversations from a call-in radio program, recorded and transcribed for the Institute of Linguistics in 1996, at the initiative of Halldór Ármann Sigurðsson (see e.g. Wide 2002: 73). Furthermore, Hrafnhildur Ragnarsdóttir has a collection of adults' and adolescents' narratives, in some cases both a spoken and a written version from the same informant that would be of great interest for our project (see e.g. Ragnarsdóttir and Strömqvist 2004; Berman, Ragnarsdóttir, and Strömqvist 2002). This material is already fully transcribed and partially coded for morphosyntactic features in CLAN-format, and we have made agreements on access to these texts.

### 2.2.2 Recordings from AMI

For several decades the Árni Magnússon Institute in Iceland (AMI) has made various recordings for their folkloric collection. Among other kinds of material, they include interviews, narratives and descriptions of folk customs. The informants come from various parts of Iceland, as well as from the Icelandic settlements in North America, and they are typically older people, many of them born in the 19th century. This material is thus an interesting complement to the material described above. The interviews have all been transcribed, but

most of the transcriptions only exist in typewritten manuscripts, and only a fraction of the original tape-recordings have been digitalized. To make this material readily accessible for search, which is necessary for the purposes of our project, we have offered to digitalize a selection of transcriptions, sharing the outcome with the AMI.

The material worked on so far are recordings of American Icelandic speech, made in 1972 and 1982. The older set of data consists mainly of narratives, and the younger one of interviews on daily life, language use, etc. An overview is given in table 2.

Table 2: Recordings of American Icelandic.

<i>Material</i>	<i>Recorded by</i>	<i>Type</i>	<i>Length of transcriptions</i> (estimated in running words)
American Icelandic 1972	Hallfreður Örn Eiríksson	Narratives	6-700,000
American Icelandic 1982	Gísli Sigurðsson	Interviews	116,500

The transcription manuscripts have been keyboarded directly, without comparing them to the recordings. In the transcription of this data, phonological and morphological deviation from the standard language are imitated directly in the orthography, and they have not been changed with respect to the conventions used for the other recordings, i.e. normal orthography with comments on deviant pronunciation and word forms outside the transcription text. This makes search in the AMI(D)a more difficult, as the standardized word-form usually applied in searching, may leave out some interesting occurrences of the word or construction in question. This must therefore be amended at a later stage, if the data is to fully serve our purposes, either by normalizing the transcriptions or by other methods, e.g. by listing the forms of each word that occur in the texts, and linking them together.

If American Icelandic proves to be an interesting variant of Icelandic, we might add more such material to our database.

### 2.3 Application

ISTAL was granted support to collect spoken language material that would serve various theoretical and practical purposes. The material added to the corpus within the Icelandic project on syntactic variation should be multifunctional in a similar way so that it not only serves the needs of our project and ScanDiaSyn in general, but also be accessible and useful for other linguistic and language technological projects concerning Icelandic. We are already collaborating with an ongoing project at the Institute of Lexicography, where a

balanced and tagged corpus of modern Icelandic, MÍM (*Mörkuð íslensk málheild* ‘Tagged Icelandic Corpus’; cf. Helgadóttir 2004), is under construction. The corpus should preferably contain some spoken language besides written texts, though the project itself cannot finance the transcription of such material. The agreement is that our project provides the corpus with transcriptions of the spoken language material, which we collect and meets the definitions of the MÍM-corpus, and in exchange we get the material tagged and lemmatized and converted to a TEI conformant XML-format. The tagging is done by a morphosyntactic tagger developed at the Institute of Lexicography (Helgadóttir 2005). Besides that, we will be able to use the written language texts in the corpus for comparison with the spoken language, and to make use of the tools for analysis and searching developed by the corpus project.

### **3. Syntactic variation in Icelandic and Faroese: previous studies and tasks for the future**

This section first describes a couple of studies that have been done on variation in Icelandic and Faroese and then offers some general thoughts on variation.

#### *3.1 Variation in Icelandic*

As mentioned at the beginning of this paper, two phenomena that had been studied before the pilot study are (1) variation in subject case marking and (2) the so-called New Passive. The main issues dealt with in these studies will now be briefly described.

##### 3.1.1 Subject case marking

Looking first at variation in subject case marking, it is principally of two kinds: Nominative Substitution (NS) and Dative Substitution (DS). NS involves substitution of nominative case for oblique case, mostly with theme/patient subjects.

- (40) a.     Bátinn    rak    að landi.  
           *boat.the.A drifted to land*  
       b.     Bátturinn rak    að landi.  
           *boat.the.N drifted to land*  
           ‘The boat drifted to the shore.’

DS involves substitution of dative case for accusative, and occurs exclusively with experiencer subjects.

- (41) a. Krakkana vantar mat.  
*kids.the.A needs food*
- b. Krökkunum vantar mat  
*kids.the.D needs food*  
 ‘The children need food.’

Reports on previous surveys of variation in subject case marking in Icelandic are found in the following works: Svavarsdóttir (1982), Halldórsson (1982), Gíslason (2001) and Jónsson and Eythórsson (2003, 2005).

The survey conducted in 2001 (Jónsson and Eythórsson 2003) tested 900 pupils (age 11-12) in 20 schools throughout Iceland. This survey was in the form of a written test where the participants were asked to fill in the relevant forms; it was modelled on the survey by Svavarsdóttir (1982) for ease of comparison. The results indicate that there has been about 25% increase in DS in Icelandic since 1982. The results show that about 90% of 11-year-old children use dative rather than the original accusative with the subject of some verbs (*gruna* ‘suspect,’ *dreyma* ‘dream,’ *langa* ‘want,’ *minna* ‘(seem to) remember,’ *vanta* ‘lack, need,’ *kitla* ‘tickle,’ *svíða* ‘smart, sting’ and *svima* ‘be dizzy’). There is considerable variation in the range of DS with different verbs (from 24.9% with *minna* to 60.4% with *svima*). It also emerged that NS, which is very common with theme subjects, also occurs to some degree with subjects of experiencer verbs.

As for the geographic distribution of the changes, it was established that DS is common in all parts of Iceland but least common in “Inner Reykjavík” (i.e. the central/western part of the city). The results for NS are less indicative with respect to its geographic distribution. It seems to be a reasonable assumption that the fact that DS is least common in Inner Reykjavík relates to the fact that this is where the level of education is the highest in the country. In fact, in other respects as well, the study showed that there are certain sociolinguistic factors connected to the use of case. In particular, there is an inverse relationship between DS and the education of the mother (the lower the education of the mother, the higher the instances of DS). This is perhaps a surprising result in view of the common assumption that Iceland is a very homogenous society with very little sociolinguistic variation. There was also found to be a certain correlation with gender, notably the fact that the innovative DS is more common among boys than girls.

### 3.1.2 The New Passive

The New Passive (a.k.a. the “new impersonal construction” as it has been called by Maling and Sigríður Sigurjónsdóttir) in Icelandic contains an auxiliary *vera* ‘be’ plus a past participle which can assign accusative case to the postverbal NP (the object). The postverbal NP can be definite:

- (42) *Það var barið litla stelpu / litlu stelpuna / mig.*  
*there was hit little girl.A / little girl.the.A / me.A*  
 ‘A little girl / the little girl / I was hit.’

As pointed out above, the New Passive is in many ways similar to the canonical expletive passive (or *það*-passive) in Icelandic, in which, however, the NP must generally be indefinite and is never in accusative case, but rather in the nominative:

- (43) *Það var barin lítil stelpa / \*litla stelpa / \*ég.*  
*there was hit.NF little girl.N / little girl.the.N / I*  
 ‘A little girl was hit.’ (the other variants are ungrammatical)

Moreover, in the canonical *það*-passive the indefinite NP can occur to the left of the non-finite verb, i.e. in a subject position, but this does not seem to be possible in the new passive.

- (44) *Það var lítil stelpa barin.*  
*there was little girl.N hit.NF*  
 ‘A little girl was hit’

A survey on the new passive was conducted in 1999-2000 in the form of a written test (Sigurjónsdóttir and Maling 2001, Maling and Sigurjónsdóttir 2002). The subjects in this survey included about 1700 students, age 15-16, in 65 schools throughout Iceland, and 200 adults.

The main results for the new passive show that it is common everywhere, except in “Inner” Reykjavík (as is the case with DS). There are also socio-linguistic factors involved, in particular correlation with the education of the mother. Moreover, the survey indicates that the new passive co-occurs with the canonical passive, i.e. there do not seem to be any speakers who exclusively use the new passive and not the canonical passive.

As to the status of the innovative construction in Icelandic, Sigurjónsdóttir and Maling (2001, cf. also Maling and Sigurjónsdóttir 2002 and Maling 2005) argue that, despite its passive morphology, it is in fact an impersonal active construction, containing a null subject which must be [+human]. This analysis is challenged in Eythórsson (2005) arguing that the construction really is

passive (cf. also Barðdal and Molnár 2003). The fact that the structural case of the object NP is “preserved” is shown to have parallels in passive constructions in various other languages, including the Mainland Scandinavian (in particular, Norwegian) *det*-passive (cf. Áfarli 1992).

### 3.2 Variation in Faroese

So far surveys have been conducted in two main areas: (1) variation in subject case marking (mainly Nominative Substitution, NS) and (2) word order (verb placement) in embedded clauses, i.e. whether the finite verb precedes or follows an S-adverb (“V-to-I movement”).

#### 3.2.1 Subject case marking

In Modern Faroese oblique subjects only occur with experiencer verbs, most of which are rather uncommon in the spoken language (Thráinsson et al. 2004). Increasingly, nominative is being substituted for oblique case with subjects. As mentioned above, however, in Icelandic NS mostly affects themes but experiencers to a more limited degree (these are mainly subjected to DS).

- (45) a. Mær dámar væl hasa bókina.  
*me.D likes well this book.the.A*
- b. Eg dámi væl hasa bókina.  
*I like.1SG well this book.the.A*  
 ‘I like that book’

Several surveys on variation in subject case marking in Faroese have been undertaken recently. In 2002 a survey was conducted which was in the form of a written test and modelled on the Icelandic test discussed above. 344 pupils (age 11-12) throughout the Faroe Islands were tested, and the answers of 286 were evaluated. Some adults were tested as well, as a control group (Eythórs-son and Jónsson 2003). The main results show that NS is very common; about 90% of the participants had some instances of NS with verbs taking dative subjects. Moreover, accusative case with subjects has become virtually extinct. As in Iceland, there is a correlation with gender; in the Faroes the innovative nominative was more common among the boys than among the girls.

In the spring of 2004 a survey testing 277 adult speakers was carried out, in which the same written test was used as in the study of the Faroese children (Jónsson and Eythórs-son 2005). The results were similar to the results for the children with respect to the difference between dative and nominative with

individual verbs. It is in particular worth noting that the original accusative is virtually non-existent among adults just as is the case among the children.

A third survey, conducted in December 2004, was designed to test intra-speaker variation in subject case marking. 66 speakers were asked to rate sentence pairs of various kinds. The results reveal considerable intra-speaker variation: for example, 25.8% of the participants accepted both nominative and dative with *dáma* ‘like,’ and 54.5% showed only a slight preference for one case over the other. Only 19.7% of the participants in this survey made a sharp distinction by accepting one case and rejecting the other.

### 3.2.2 V-Adv / Adv-V in embedded clauses

In Faroese there is variation with respect to the position of the finite verb in embedded clauses. The verb either precedes S-adverbs (V-Adv), as in Icelandic, or follows S-adverbs (Adv-V), as in Mainland Scandinavian. In other words, there is variation with respect to “V-to-I movement” (Höskuldur Thráinsson et al. 2004).

- (46) a. Eg haldi at Jógvan hefur **aldri** lisið bókina.  
*I believe that Jógvan has never read book.the.A*
- b. Eg haldi at Jógvan **aldri** hefur lisið bókina.  
*I believe that Jógvan never has read book.the.A*  
 ‘I believe that Jógvan never has read the book’

So far two, relatively small-scale, surveys have been made into this variation (Petersen 2000, and Thráinsson 2001, 2003). Petersen (2000) tested 18 students (age 20) throughout the Faroes. Most of these speakers rejected the V-Adv order in embedded clauses. Thráinsson (2001, 2003) tested 14 high-school students and two linguists, and found more variation in verb placement in embedded clauses than Petersen (2000). In addition, Thráinsson (2001, 2003) reports on his survey of written texts from the 19th and 20th centuries. In this survey V-Adv order was found in 80% of the cases with bridge verbs, but in only 37% of the cases with non-bridge verbs.

## 3.3 Further studies

### 3.3.1 What is variation?

At this point it is useful to consider what intra-language variation really is. Often variation is defined as something that involves two or more formally similar ways of expressing the same meaning or discourse function. What we have in mind here is a strong formal similarity that only allows for minor



differences, e.g. in morphosyntax, the presence or absence of “dummy” elements or the use of reflexives vs. pronominals:

- (47) a. Hana / Henni langar heim.  
*she.A/ she.D wants home*  
 ‘She wants to go home.’
- b. Fáir vona að (það) rigni á morgun.  
*few hope that (it) rains tomorrow*  
 ‘Few people hope that it will rain tomorrow.’
- c. Ég hitti Jón hjá frænku sinni / hans.  
*I met John at aunt self/ his*  
 ‘I met John at his aunt’s place.’

Of course, it is quite possible that the same speaker may accept both variants, for instance in (47b), in which case we would have optionality. However, in the absence of a large-scale investigation it is difficult to know whether a particular phenomena shows variation between speakers or mere optionality or both (see Jónsson and Eythórsson 2005 on the last possibility).

Variation in a broad sense also includes differences in acceptability. This can be exemplified by the new passive (48a) vs. the regular passive (48b) in Icelandic:

- (48) a. Það var lamið mig í gær.  
*there was hit me.A yesterday*  
 ‘I was hit yesterday.’
- b. Ég var laminn í gær.  
*I.N was hit yesterday*

The new passive differs from the regular passive in that the logical subject remains in the object position and is assigned case by the passive verb. As a result, the clause-initial subject position is filled by an expletive and there is no agreement between the logical subject and the finite verb and passive participle. By contrast, the nominative subject in (48b) triggers 3rd person singular agreement on the finite verb and masculine singular nominative on the passive participle (if the speaker is a man). Since all speakers who accept the new passive also accept the regular passive, there is no “variation” between the new and the regular passive. However, there is variation in that some speakers (mostly children and teenagers) accept the new passive but other speakers do not (see Sigurjónsdóttir and Maling 2002 for further discussion).

### 3.3.2 The domain of investigation

The limitations of time and other resources make it impossible to investigate all the syntactic phenomena that may be of interest for our study of syntactic variation in Icelandic and some arbitrary decisions will have to be made to limit the domain of investigation. Still, various criteria can be applied to narrow down the list of possible topics. We believe for instance that phenomena of the following kind should be included in the Icelandic survey:

- (49)
- a. phenomena believed to exhibit variation between individuals
  - b. phenomena of special interest in Icelandic
  - c. phenomena of interest in a Scandinavian comparative perspective
  - d. phenomena of special theoretical interest (e.g. with respect to clustering of variation)

Of these criteria, (49a) is probably the most widely applicable and we will have to use the results of the pilot study to exclude phenomena that do not exhibit any significant variation between individuals. The criteria in (49b) is e.g. much more narrow and it is primarily intended to cover phenomena that are undergoing a change in progress and may not have any obvious parallels in the other Scandinavian languages (e.g. the new or extended progressive).

### 3.3.3 Methodology

Despite the obvious usefulness of the Icelandic corpora described above, the data that can be extracted from these corpora are not sufficient for our purposes. This is not only due to the general limitations of samples from real language use (e.g. the fact that they do not really show what is impossible in a language), but also because the Icelandic corpora are too small or homogenous to genuinely reflect sociological or dialectal variation in syntax. Thus, additional data must be solicited from native speakers through various grammaticality tasks. We will use two basic methods of data solicitation, written questionnaires and oral interviews. As we will discuss below, data elicitation brings many problems of its own (see e.g. Schütze 1996 and Cornips and Poletto 2005) and we will try to tackle these problems as best as we can.

Written questionnaires are a convenient way of obtaining a lot of data in a short amount of time and they are also a perfectly viable option in a country like Iceland where dialectal differences are very limited (see Árnason and Thráinsson 2003) and all regional variants have the same written form. Moreover, the results of the pilot study suggest that written questionnaires are just as reliable as oral interviews, at least in cases where phonological factors do

not play a significant role. On the negative side, written questionnaires are not suitable for soliciting judgments on syntactic phenomena that are characteristic of the spoken language, such as the use of expletive *það* ‘there’ in Icelandic. Another problem is that written language tends to be rather formal and conservative, thus favoring judgments that are consistent with the prescriptive norms of the language rather than the intuitions of the speaker. This problem is particularly acute for syntactic phenomena that have been stigmatized in Iceland, e.g. Dative Substitution or the new progressive, but may also be relevant for many other phenomena as well.

The main drawback of oral interviews is that they are very time consuming so only a limited number of the participants can be tested in that way. On the other hand, there are many advantages to oral interviews that make them a highly valuable tool in our survey. First, they provide a more natural setting than written questionnaires as they involve interaction between the investigator and the participants in the survey. Thus, the investigator can react to the answers of the participants, for example by asking follow-up questions or questions of clarification. Second, oral interviews make it much easier to test syntactic phenomena that are sensitive to intonation or stress as these factors can be controlled for. Third, it is easier to test syntactic phenomena that are sensitive to semantics (e.g. Object Shift) in oral interviews as the investigator can more easily ask the participant to judge a particular example in different scenarios. Fourth, the fact that oral interviews are recorded makes it possible to hear whether the participant answers the questions spontaneously or takes some time to think about the answers. This is important because the latter case suggests that the participant may be accommodating to prescriptive pressure.

#### **4. Conclusion: intended outputs and benefits**

This concludes our report of the Icelandic pilot project and the ensuing plans. It is expected that our survey will bring about many outputs and benefits. Among the more tangible things we can mention doctoral dissertations, an extensive database on syntactic variation in Icelandic, published articles and workshops to be held in Iceland. The main benefit of the survey is greater awareness of syntactic variation in Icelandic and a better understanding of its range and limits. This will help linguists not only to make sense of all the conflicting claims about Icelandic syntax in the literature, but also to shed light on the validity of the Principles and Parameters framework of generative grammar. In that framework, syntactic variation between languages reduces to a choice between different values of binary parameters. However, it is an open

question to what extent inter-language variation can be analyzed in the same way and some authors have in fact argued for the existence of non-parametric variation within the same language (see e.g. Jónsson and Eythórsson 2005 on case marking in Insular Scandinavian and Barbiers 2005 on verb clusters in Dutch).

### References

- Áfarli, Tor A. 1992. *The Syntax of the Norwegian Passive Constructions*. John Benjamins, Amsterdam.
- AMI: The Árni Magnússon Institute in Iceland; <http://www.am.hi.is/>
- Árnason, Kristján, and Höskuldur Thráinsson. 2003. 'Fonologiske dialekttræk på Island. Generationer og geografiske områder,' in Gunstein Akselberg, Anne Marit Bødal, and Helge Sandøy (eds) *Nordisk dialektologi*. Novus forlag, Oslo, 151-196.
- Barbiers, Sjef. 2005. 'Word order variation in three-verb clusters and the division of labour between generative linguistics and sociolinguistics,' in Leonie Cornips and Karen P. Corrigan (eds.) *Syntax and Variation: Reconciling the Biological and the Social*. John Benjamins, Amsterdam, 233-264.
- Barðdal, Jóhanna. 2001. *Case in Icelandic. A Synchronic, Diachronic and Comparative Approach*. Doctoral dissertation, University of Lund, Lund.
- Barðdal, Jóhanna, and Valéria Molnár. 2003. 'Passive in Icelandic – compared to Mainland Scandinavian,' in Jorunn Hetland and Valeria Molnár (eds.) *Structures of Focus and Grammatical Relations*. Niemeyer, Tübingen, 231-260.
- Berman, Ruth A., Hrafnhildur Ragnarsdóttir, and Sven Strömquist. 2002. 'Discourse Stance: Written and spoken language,' in Ruth A. Berman and Ludo Verhoeven (eds.) *Cross-Linguistic Perspectives on the Development of Text-Production Abilities in Speech and Writing*. Part 2: Special issue of *Written Language and Literacy* 5:2. John Benjamins, Amsterdam/Philadelphia, 255-289.
- Cornips, Leonie, and Cecilia Poletto. 2005. 'On standardising syntactic elicitation techniques (part 1),' *Lingua*, 939-957.
- Eythórsson, Thórhallur, and Jóhannes Gísli Jónsson. 2003. 'The Case of Subject in Faroese,' *Working Papers in Scandinavian Syntax* 72, 207-232.
- Eythórsson, Thórhallur. To appear. 'The new passive in Icelandic: Merge over Move,' in Thórhallur Eythórsson (ed.) *Grammatical Change and Linguistic Theory: The Rosendal Papers*. John Benjamins, Amsterdam.
- Friðriksson, Finnur. 2004. 'Real vs. Imagined Change. The case of modern Icelandic,' in Britt-Louise Gunnarsson, Lena Bergström, Gerd Eklund, Staffan Fridell, Lise H. Hansen, Angela Karstadt, Bengt Nordberg, Eva Sundgren, and Mats Thelander (eds.) *Language Variation in Europe*, Papers from ICLaVE 2. Uppsala universitet, Uppsala, 168-180.
- Gíslason, Björn. 2001. *Að verjast föllum: Um þágufallssýki og nefnifallssýki*, [Defending against Cases: On Dative and Nominative Substitution]. BA-thesis, University of Iceland, Reykjavík.

- Halldórsson, Halldór. 1982. 'Um méranir: Drög að samtímalegri og sögulegri athugun,' [On Dative Substitution: A draft of a Synchronic and Diachronic Study], *Íslenskt mál og almenn málfræði* 4, 159-189.
- Helgadóttir, Sigrún. 2004. 'Mörkuð íslensk málheild,' in *Samspil tungu og tækni. Afrakstur tungutækniverkefnis menntamálaráðuneytisins*. Menntamálaráðuneytið, Reykjavík, 65-71.
- Helgadóttir, Sigrún. 2005. 'Testing Data-Driven Learning Algorithms for PoS Tagging of Icelandic,' in Henrik Holmboe (ed.) *Nordisk sprogteknologi 2004. Årbog for Nordisk Sprogteknologisk Forskningsprogram 2000–2004*. Museum Tusulanums Forlag/Københavns universitet, København, 257-265.
- ISTAL: <http://www.hi.is/~eirikur/istal>
- Jónsson, Jóhannes Gísli. 2005. 'Merkingarhlutverk, rökliðir og fallmörkun,' in Höskuldur Thráinsson (ed.) *Setningar*. Almenna bókafélagið, Reykjavík, 350-409.
- Jónsson, Jóhannes Gísli, and Thórhallur Eythórsson. 2003. 'Breytingar á frumlagsfalli í íslensku,' [Changes in Subject Case in Icelandic], *Íslenskt mál og almenn málfræði* 25, 7-40.
- Jónsson, Jóhannes Gísli, and Thórhallur Eythórsson. 2005. 'Variation in subject case marking in Insular Scandinavian,' *Nordic Journal of Linguistics* 28, 223-245.
- Kayne, Richard S. 1996. 'Microparametric Syntax: Some introductory remarks,' in James R. Black and Virginia Motapanyane (eds) *Microparametric Syntax and Dialect Variation*, John Benjamins, Amsterdam, i-xviii.
- Maling, Joan. 2002. 'Pað rignir þágufalli á Íslandi. Verbs with Dative Objects in Icelandic,' *Íslenskt mál* 24, 31-105.
- Maling, Joan. 2006. 'From Passive to Active: syntactic change in progress in Icelandic,' in Torgrim Solstad and Benjamin Lyngfelt (eds.) *Demoting the Agent: Passive and Other Voice Related Phenomena*. John Benjamins, Amsterdam.
- Maling, Joan, and Sigríður Sigurjónsdóttir. 2002. 'The "New Impersonal" Construction in Icelandic,' *Journal of Comparative Germanic Linguistics* 5, 97-142.
- MÍM: <http://www.lexis.hi.is/malheild.htm>
- MIN: <http://www.hf.uib.no/moderne/default.html>
- Óladóttir, Hanna. 2005. *Pizza eða flatbaka? Viðhorf 24 Íslendinga til erlendra máláhrifa í íslensku*. Unpublished MA-thesis, Háskóli Íslands, Hugvísindastofnun, Reykjavík.
- Petersen, Hjalmar P. 2002. 'Quirky case in Faroese,' *Fróðskaparrit* 50, 63-76.
- Ragnarsdóttir, Hrafnhildur, and Sven Strömquist. 2004. 'Time, space and manner: Aspects of narrative construction in two closely related languages,' in Sten Strömquist and Ludo Verhoeven (eds.) *Relating Events in a Narrative: Typological and Contextual Perspectives*. Lawrence Erlbaum Associates, Mahwah/NJ.
- Reuland, Eric, and Sigríður Sigurjónsdóttir. 1997. 'Long Distance "Binding" in Icelandic: Syntax or Discourse?,' in Hans Bennis, Pierre Pica and Johan Rooryck (eds.) *Atomism and Binding*. Foris, Dordrecht, 323-340.
- Rögnvaldsson, Eiríkur. 1997. 'Frumlag og fall að fornu,' *Íslenskt mál* 18, 37-69.

- Rögnvaldsson, Eiríkur. 2002. ‘Corpus of Spoken Icelandic (ÍS-TAL),’ in Henrik Holmboe (ed.) *Nordisk Sprogteknologi 2002. Årbog for Nordisk Sprogteknologisk Forskningsprogram 2000–2004*. Museum Tusulanums Forlag/Københavns universitet, København, 43-44.
- Sandøy, Helge. This volume. ‘Experiences from Nordic research collaboration in linguistics.’
- Schütze, Carson. 1996. *The Empirical Base of Linguistics. Grammaticality Judgments and Linguistic Methodology*. The University of Chicago Press, Chicago.
- Sigurjónsdóttir, Sigríður, and Joan Maling. 2001. “‘Það var hrint mér á leiðinni í skólann’: Polmynd eða ekki þolmynd?” [“There was pushed me on the way to school”: passive or not passive?], *Íslenskt mál og almenn málfræði* 23, 123-180.
- Stefánsdóttir, Vala. 2005. ‘Fallmörkun frumlaga og nýja þolmyndin,’ term paper, University of Iceland, Reykjavík.
- Svavarsdóttir, Ásta. 1982. ‘Þágufallssýki,’ [Dative Sickness] *Íslenskt mál og almenn málfræði* 4, 19-62.
- Thráinsson, Höskuldur. 1991. ‘Long-distance Reflexives and the Typology of NPs,’ in Jan Koster and Eric Reuland (eds.) *Long-Distance Anaphora*. Cambridge University Press, Cambridge, 49-75.
- 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. 2001. ‘Um sagnbeygingu, sagnfærslu og setningagerð í færeysku og fleiri málum,’ [On verbal inflection, verb movement and syntactic structure in Faroese and some other languages.], *Íslenskt mál og almenn málfræði* 23, 7-70.
- Thráinsson, Höskuldur. 2003. ‘Syntactic variation, historical development and Minimalism,’ in Randall Hendrick (ed.) *Minimalist Syntax*. Blackwell, Oxford, 152-191.
- Thráinsson, Höskuldur. 2005. *Setningar*, editor and main author Höskuldur Thráinsson, co-authors Jóhannes Gísli Jónsson, Eiríkur Rögnvaldsson, Sigríður Sigurjónsdóttir, Sigríður Magnúsdóttir, and Thórunn Blöndal, vol. III of *Íslensk tunga*. Almenna bókafélagið, Reykjavík.
- Vangsnes, Øystein. 2002. ‘Distributiv possessiv – en binominal konstruksjon,’ in Inger Moen (et al.) (eds.) *Mons 9: Utvalgte artikler fra Det niende møtet om norsk språk i Oslo 2001*. Novus, Oslo, 230-243.
- Wide, Camilla. 2002. *Perfect in Dialogue. Form and functional potential of the vera búinn að + inf. construction in contemporary Icelandic*, PIC Mongraphs 3, Department of Scanavian Languages and Literature, University of Helsinki, Helsinki.