Icelandic Variation Projects: An Overview

Höskuldur Þráinsson
University of Iceland
hoski@hi.is
Outline of the talk

A brief description of projects and the methods:
• Björn Guðfinnsson’s study in the 1940s
• RÍN in the 1980s
• Ásta Svavarsdóttir’s study 1981, Tolli and Jóhannes 2001+ etc.
• Sigríður Sigurjónsdóttir and Joan Maling 1999+
• IceDiaSyn and FarDiaSyn 2004+
• RAUN 2010+

Our experience with the methods used:
• Methodological points about the phonological studies
• Methodological points about the syntactic studies

Some similarities and differences:
• On “natural data” and other data
• The pervasiveness of intra-speaker variation
BG’s study in the 1940s

**Topic:** Phonological variation

**Subjects:** 6520 12 year olds

**Method:** Reading of special texts

**Coverage:** All parts of the country

**Emphasis:** Regional variants

---

The areas visited by BG

The way BG’s data are catalogued
Main results of BG’s study:
• There are regional phonological variants in Icelandic
• Some of these variants are restricted to relatively small geographical areas and characterize the speech of small minorities of the population
• In general, however, these variants could (with some simplification) be grouped into:
  a. Northern and North-Eastern variants
  b. South-Eastern variants
  c. Southern variants
  d. North-Western variants

Reykjavík is a melting pot but gradually emerging as the home of “the majority dialect”
RÍN 1980s: Kristján & Höskuldur

**Topic:** Phonological variation

**Subjects:** 2800+ subjects, all age groups

**Methods:** Interviews centering around pictures
Reading of texts

**Coverage:** All parts of the country

**Emphasis:** Comparison with BG (400 of his subjects reinterviewed); regional and social variation; development of the variants

---

**RÍN-pictures**

*Form used in analysis of the RÍN-data*
RÍN, 2

Some results from RÍN (see also a special poster):
• Most of the regional variants can still be found
• Some of the variants are fading fast, others gaining ground
• Hence, there is often clear correlation with age ("apparent time")
• Hardly any evidence for correlation with education
• Female speakers are sometimes leading the way in changes

Dissemination of results and accessibility of data:
• Various articles (cf. references on handout by poster)
• Recordings are being digitized and some are already accessible on the Internet (cf. presentation on Saturday)

Some remaining questions and tasks:
• How do the variants spread?
• More detailed study of change in real time needed (= RAUN)
Ásta’s study of “Dative Sickness”

**Topic:** Variation in subject case  
**Subjects:** Some 200 11 year olds  
**Method:** Filling blanks in a text  
**Coverage:** Selected areas (11 schools)  
**Emphasis:** Linguistic and social variation

Ásta’s fill-in method:

Á morgun á að fara í skólaferðalag. Sif ætlar auðvitað með. Sif þykir gaman að ferðast og Sif hlakkar óskaplegra til. Þó kvíðir Sif fyrir einu. Sif verður oft svo skelfilega bilveik og Sif langar sko ekki að fara að æla í rútunni. Sif leidist ósköp þessi bilveiki en vonar þó að allt gangi vel.

Á morgun á að fara í skólaferðalag. Sif ætlar auðvitað með. __ þykir gaman að ferðast og __ hlakkar óskaplegra til. Þó kvíðir __ fyrir einu. __ verður oft svo skelfilega bilveik og __ langar sko ekki að fara að æla í rútunni. __ leidist ósköp þessi bilveiki en vonar þó að allt gangi vel.
Some results from Ásta’s study:

- The method apparently worked well.
- Dative Substitution was found in all the areas – but it was relatively less frequent in Reykjavík than in the rural areas.
- There were considerable lexical differences.
- There was some indication that the use of the “correct” case (Acc., sometimes Nom.) correlated with:
  a. Social class (lower social class = more Dat.)
  b. Academic performance as evaluated by the teachers (lower acad. ability = more Dat.)
  c. Gender (boys relatively more likely to use Dat.)

(cf. Ásta Svavarsdóttir 1982, Ásta, Gísli & Þórólfur 1984)
Tolli and Jóhannes:

Subject Case in Icelandic and Faroese

Two (or more) projects (cf. overview on poster):

**Topic:** Variation in subject case **in Icelandic**

**Subjects:** Some 850 11 year olds (for comparative purp.)

**Method:** Fill-in similar to Ásta’s

**Coverage:** Selected areas (20 schools)

**Emphasis:** Linguistic and social variation

**Topic:** Case variation **in Faroese**

**Subjects:** 290 teenagers and 280 adults

**Method:** Fill-in similar to Ásta’s

**Coverage:** Mostly Tórshavn

**Emphasis:** Linguistic variation
Tolli and Jóhannes, 2

Some results from Tolli and Jóhannes’s Icelandic study:
• Dative substitution had apparently increased since Ásta’s study
• Correlation with education of the s’s mothers (less ed. = more D.)
• Dative substitution more prevalent w. boys
• Dat. subst. relatively less common in Reykjavík than elsewhere – and in Reykjavík it was more common in the newer parts

... and their Faroese study (together with the Icel. one):
• Dat. and Nom. subst. shows extensive intra-speaker variation
• This variation is arguably “grammar-internal” optionality rather than competition between grammars
• There is no evidence for influence from one dialect (or language) on another here – the development is towards more regularity

cf. e.g. Jóhannes & Tolli’s joint papers 2003, 2005 and several other papers of theirs (vs. Kroch 1989, 2001)
**Sigga and Joan Maling 1999-2000**

**Topic:** The New Passive/New Impersonal  
**Subjects:** 1700 teenagers + 200 adults  
**Methods:** Questionnaire (asking for judgments)  
**Coverage:** All parts of the country  
**Emphasis:** Linguistic and social variation

The form of the questionnaire used by Sigga and Joan:

```
Leiðbeiningar: Settu X í viðeigandi dálk.

Já = Svona getur maður sagt!  
Nei = Svona getur maður ekkki sagt!
```

<table>
<thead>
<tr>
<th></th>
<th>Já</th>
<th>Nei</th>
</tr>
</thead>
<tbody>
<tr>
<td>Páð var rekið Ólaf úr skólanum.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hún var skilin eftir hjá ömmu og afá.</td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>
```
Sigga and Joan, 2

Some results from Sigga and Joan’s study

• The construction was widespread throughout Iceland
• It was accepted less frequently in “inner Reykjavík” than elsewhere
• There was no correlation with gender but strong correlation with the parents’ education (less education = higher accept.)
• The construction was typically rejected by the adult subjects

Table from Sigga and Joan: Acceptance rate by different groups:

<table>
<thead>
<tr>
<th></th>
<th>Adults</th>
<th>Teenagers in inner R</th>
<th>Teenagers in outer R and elsewhere</th>
</tr>
</thead>
<tbody>
<tr>
<td>Það var beðið mig að vaska upp</td>
<td>7%</td>
<td>47%</td>
<td>73%</td>
</tr>
<tr>
<td>there was asked me(A) to wash up</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Það var sagt mér að taka til</td>
<td>2%</td>
<td>34%</td>
<td>62%</td>
</tr>
<tr>
<td>there was told me(D) to clean up</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

cf. Sigga and Joan 2001, Joan and Sigga 2002, etc.
The dialectal surveys providing the empirical data

IceDiaSyn

- Part of ScanDiaSyn (Scandinavian Dialect Syntax)
- Connected to NORMS (Nordic Center of Excellence in Microcomparative Syntax, http://norms.uit.no/)

Three large scale surveys (7–800 participants in each, different age groups, different parts of the country ...), using questionnaires (different tasks: judgments, fill-ins ...); also interviews, comparison with corpora ...

FarDiaSyn

Faroese sister project, two large scale surveys (220–320 participants in each ...), using comparable methods, also interviews, etc...

RÍN

Large scale survey of phonological dialects of Icelandic. Interviews with some 3000 speakers (some being reinterviewed).
On the methodology of the syntactic surveys

Important points:

To get reliable results using written questionnaires it is important to take certain methodological precautions (cf. e.g. Schütze 1996 (ch. 5), Cornips and Poletto 2005):

• make sure everybody get the same instructions (read them)
• explain the grading scale by giving illustrative examples
• vary the order of the test sentences (e.g., reverse for half of the subjects)
• test different constructions in each overview and include fillers
• vary the tasks (absolute judgments, relative judgments, fill-ins …)
• include a break in long sessions to prevent excessive fatigue and boredom
• include context sentences to get all subjects thinking of similar contexts
• try to use natural sounding examples (short, plausible, lexically neutral …)
• test multiple examples of each construction to minimize unwanted effects
• try to make the contrasting variants maximally close to minimal pairs
• test different types of speakers (age groups, locations …)
• throw out data from “unreliable speakers” (e.g. “language specialists”)
• get speakers to report on their own intuition (cf. Henry 1995, 2005a,b)
Even more important points:
We have shown that it is possible to get reliable results by using written questionnaires of the kind used in the syntactic surveys:

• The variation is systematic (differences between age groups and (in a few cases) regions, etc.) and not random (Or: “The proof is in the pudding” – and the pudding tastes good).
• All generations seem reliable (e.g., it’s not the case that the youngest generation “accepts everything”).
• The subjects answer honestly in general and don’t seem worried by any kind of prescriptivism or the like (there is very little awareness of most of the variants anyway, cf. below).
• Comparison of different tasks confirms reliability of judgments.
• Comparison with corpora confirms reliability of judgments.
• Comparison with interviews confirms reliability of judgments.
• Comparison with results of the phonological interviews shows interesting parallels strengthening the conclusion.
On the reliability of the data

Systematic variation and generational differences:

Acceptance of Topicalization in complement clauses by different age groups. Correlation with age: $r = 0.466$. Statistical significance: $p < 0.001$.

Acceptance of Dative Subjects with typical “Dative Sickness” verbs by different age groups. Correlation with age: $r = 0.511$. Statistical significance: $p < 0.001$. 

October 8, 2010

Höskuldur Práinsson
NLVN, Reykjavik
On the reliability of the data, 2

Honest answers and stigmatized (?) variants:

A potential problem:
When asked to judge sentences reflecting “Dative Sickness” subjects will not admit that they find them acceptable. If that’s so, then the acceptance rate will be lower than the relative frequency found in masked tests (disguised fill-ins) and in spontaneous speech. But this is not what we find:

Despite the difference between judgments and fill-ins, there is still a strong and significant correlation between the judgments and the selection of dative subjects in fill-ins for typical Dative Sickness verbs: \( r = 0.570, p < 0.001 \)
On the reliability of the data, 3

Comparison between judgments of the New Passive (cf. e.g. Maling and Sigurjónsdóttir 2002) and and its production in interviews: Correlation: \( r = 0.989 \), Significance: \( p = 0.001 \)

**Judgment task:**
\( \text{Það var strítt honum á hverjum degi} \) \( \text{yes} \) \( ? \) \( \text{no} \)
there was teased him every day

**Rephrasing in an interview with a subset of the speakers:**

**Model** (the formation of an expletive sentence):

Sumir köstuðu tómötum í söngvarann > Það var kastað tómötum ...

some threw tomatoes at the singer there were thrown tomatoes ...

**First example** (calls for a normal expletive):

Einhverjir fóru að syngja í rútunni > Það fóru einhverjir að syngja ...

some began to sing in the bus there began some to sing ...

**Later example** (calls for the New Passive):

Krakkarnir hrintu mér í frímínútunum > Það var hrint mér ...
the kids pushed me during intermission there was pushed me ...
Inter-speaker variation

Variation in case marking is arguably more “pure” 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 our purposes.
Inter-speaker variation, 2

(Dative) case alternation is

**Dative subjects in Icelandic**

- Some speakers use **dative** (or accusative) **subjects** with a couple of verbs that used to take nominative subjects. 

  *Acceptance rate (%) in the written survey of IceDiaSyn:*

    | Nom | Acc | Dat |
    |-----|-----|-----|
    | 48,6 | 59,7 | 44,2 |

  - *hlakka til ‘look forward to’*

- Some speakers use **dative subjects** with a few verbs that used to take accusative subjects.

  *Acceptance rate in the written survey of IceDiaSyn:*

    | Acc. | Dat. |
    |------|------|
    | 92,1 | 56,6 |
    | 88,3 | 68,2 |

  - *vanta ‘need, lack’*
  - *langa ‘want, long for’*
Inter-speaker variation, 3

Dative objects in Icelandic

Some speakers use **dative objects** with certain verbs where other speakers use accusative objects.

*Acceptance rate (%)* in the written survey of IceDiaSyn:

<table>
<thead>
<tr>
<th>Verb</th>
<th>Acc.</th>
<th>Dat.</th>
</tr>
</thead>
<tbody>
<tr>
<td>faxa ‘fax’</td>
<td>91,3</td>
<td>23,8</td>
</tr>
<tr>
<td>framlengja ‘extend’</td>
<td>82,7</td>
<td>61,0</td>
</tr>
<tr>
<td>negla ‘nail (a ball into a goal)’</td>
<td>66,5</td>
<td>72,6</td>
</tr>
<tr>
<td>rústa ‘demolish’ (lit. and fig.)</td>
<td>22,3</td>
<td>82,4</td>
</tr>
</tbody>
</table>
Inter-/Intra-speaker variation

Observe:
Adding the percentages for acceptance of (Nom) + Acc + Dat subjects and Acc + Dat objects in the preceding slides gives more than 100% in each instance: Suggests intra-speaker variation.

**Judgments of subject case (% of speakers accepting):**

<table>
<thead>
<tr>
<th></th>
<th>Nom</th>
<th>Acc</th>
<th>Dat</th>
<th>=</th>
</tr>
</thead>
<tbody>
<tr>
<td>hlakka til ‘look forward to’</td>
<td>48,6</td>
<td>59,7</td>
<td>44,2</td>
<td>152,5</td>
</tr>
<tr>
<td>vanta ‘need, lack’</td>
<td>92,1</td>
<td>56,6</td>
<td></td>
<td>148,7</td>
</tr>
<tr>
<td>langa ‘want, long for’</td>
<td>88,3</td>
<td>68,2</td>
<td></td>
<td>156,5</td>
</tr>
</tbody>
</table>

**Judgments of object case:**

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>faxa ‘fax’</td>
<td>91,3</td>
<td>23,8</td>
<td></td>
<td>115,1</td>
</tr>
<tr>
<td>framlengja ‘extend’</td>
<td>82,7</td>
<td>61,0</td>
<td></td>
<td>143,7</td>
</tr>
<tr>
<td>negla ‘nail (a ball into a goal)’</td>
<td>66,5</td>
<td>72,6</td>
<td></td>
<td>139,1</td>
</tr>
<tr>
<td>rústa ‘demolish’ (lit. and fig.)</td>
<td>22,3</td>
<td>82,4</td>
<td></td>
<td>104,7</td>
</tr>
</tbody>
</table>
Intermezzo about a myth

The myth
It would be best if we could rely on “natural data” in syntactic studies (“real examples” found in spontaneous speech) but unfortunately we cannot because the constructions we are interested in are too infrequent.

The truth
“Natural data” can never tell the whole story about the internal grammar of the speakers. An illustration (cf. preceding slide):

<table>
<thead>
<tr>
<th></th>
<th>Dative selected</th>
<th>Dative accepted</th>
</tr>
</thead>
<tbody>
<tr>
<td>Object case with <em>rústa</em></td>
<td>88,1%</td>
<td>83,6%</td>
</tr>
<tr>
<td>Subject case with <em>langa</em></td>
<td>19,2%</td>
<td>68,2%</td>
</tr>
</tbody>
</table>

Clear and important difference here which would not show up in “natural data”.
Intermezzo about intra-speaker variation and case in general

• The intra-speaker variation just illustrated does not reflect a general case-confusion. For most verbs there is no variation in subject or object case. Another “case” in point: Case-transfer from a relative clause to the NP it modifies:

Listaverkin eru ómetanleg. Listaverkunum var stolið.
the pieces-of-art(Nom) are invaluable the pieces-of-art(Dat) were stolen

Listaverkin/Listaverkunum [sem var stolið e ] eru ...
the pieces-of-art Nom/Dat that were stolen (Dat) are ...

Acceptance by all subjects (N 710) Nom.  Dat. =
by oldest group (N 160) 30,0% 57,4% 87,4%
47,0% 50,0% 97,0%
Intra-speaker variation

For most **Faroese verbs** there is no variation in subject or object case but for a few there is (there are virtually no Acc. subjects left in Faroese, cf. Thráinsson et al. 2004, Jónsson & Eythórsson 2005).

### Examples of subject case variation in Faroese:

<table>
<thead>
<tr>
<th>Verb</th>
<th>Nom.</th>
<th>Dat.</th>
<th>Either (= intra-sp.var.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>tørva ‘need’</td>
<td>21.6%</td>
<td>73.9%</td>
<td>0?</td>
</tr>
<tr>
<td>mangla ‘lack’</td>
<td>73.5%</td>
<td>58.4%</td>
<td>&gt;32%</td>
</tr>
<tr>
<td>nýtast ‘need’</td>
<td>60.6%</td>
<td>74.3%</td>
<td>&gt;35%</td>
</tr>
<tr>
<td>dáma ‘like’</td>
<td>54.5%</td>
<td>86.6%</td>
<td>&gt;41%</td>
</tr>
</tbody>
</table>

### Examples of subject case variation in Faroese:

<table>
<thead>
<tr>
<th>Verb</th>
<th>Acc.</th>
<th>Dat.</th>
<th>Either</th>
</tr>
</thead>
<tbody>
<tr>
<td>stýra ‘steer’</td>
<td>37.1%</td>
<td>77.4%</td>
<td>&gt;14%</td>
</tr>
<tr>
<td>náa ‘get, reach’</td>
<td>75.9%</td>
<td>49.8%</td>
<td>&gt;26%</td>
</tr>
</tbody>
</table>
Intra-speaker variation, 5

Possible objection:
The extensive intra-speaker variation illustrated here is somehow an artifact of the research method (evaluation of sentences rather than production).

Replies:
• The judgments are based on the speakers’ intuition and they are not a less important reflection of grammatical knowledge than production is (cf. slide 16 above, see also Schütze 1996:50–52, who refers to Pateman 1987 on the difference between intuition and introspection).
• Production also shows intra-speaker variation in the use of Dat subjects (Ásta Svavarsdóttir 2010): Subject case of langa ‘want’ in a spoken language corpus:

<table>
<thead>
<tr>
<th></th>
<th>Nom</th>
<th>Acc</th>
<th>Dat</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sp. 1</td>
<td>1</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Sp. 2</td>
<td>0</td>
<td>19</td>
<td>1</td>
</tr>
<tr>
<td>Sp. 3</td>
<td>1</td>
<td>5</td>
<td>0</td>
</tr>
</tbody>
</table>
Intra-speaker variation, 8

The observed intra-speaker variation is **not random**:

• Speakers exhibit it to **different degrees** (cf. the histograms).

• Even where it is extensive, it is often possible to find interesting **correlations** between judgments that “should” go together according to certain theoretical proposals (parameters). Some Faroese examples of this:
  
  **Correlation between the acceptance of V-movement in emb. clauses**
  
  **and Stylistic Fronting**
  
  **Correlation between the acceptance of Expletive Associates in High Pos.**
  
  **and Transitive Expletives**
Intra-speaker variation:

- **Is common and pervasive** in those aspects of grammar that are undergoing change.
- Can be observed in (spontaneous) **speech production** but sometimes even more clearly in (syntactic) **judgments**.
- **Is not random**, may be affected by extra-linguistic features (style, situation...) but not necessarily so, and may show interesting correlations.
- **Needs to be taken seriously** in models of grammar.
References


References, 3


