

Can phonology make categorical predictions about syntactic variation?
On the interaction between prosody and subject case marking in Icelandic

Anton Karl Ingason
University of Iceland
anton.karl.ingason@gmail.com

March 4, 2009
Workshop: Rhythm Beyond The Word
University of Osnabrück – DGfS 2009

1 Introduction

- (1) If phonology can decide between two candidates when there is no syntactic preference (Golston, 1995) – phonology should be able to predict the structure of syntactic variation.
- (2) This talk will discuss variation in subject case in Icelandic and the goal will be to say something about the structure of this phenomenon – by referring to prosodic preferences.
- (3) By observing variation in subject case marking across the syntax-phonology interface it may be possible to account for categorical effects which appear only as unexplained frequencies if we only look at the syntax.

2 Variation Across the Interface

- (4) Types of variation:
 - a. Different systems: In Optimality-Theoretic-terms (Prince and Smolensky, 1993/2004) each system has its own constraint ranking but within a system the algorithm produces the same output given the same input. Makes sense when describing regional dialects, social registers and other situations where the systems are indeed different. (typological difference)
 - b. Variation within a “fixed” system: The algorithm allows for the possibility of different outputs even if the input stays the same. Makes sense if the system produces different outputs when there seems to be no motivation for changing the system. (optionality)
- (5) Ways to predict variation within a fixed system (introduce the possibility of “don’t care” situations):
 - a. Partial ranking (Anttila, 1997)
 - b. Add some noise at EVAL time (Boersma and Hayes, 2001)
 - c. ... and some more ...
 - d. Allow the speaker to access non-optimal candidates (Coetzee, 2004, 2006).
- (6) A Rank Ordering Model of EVAL (ROE) (Coetzee, 2004, 2006) makes use of the observation that an OT grammar predicts, whether we like it or not, not only the optimal candidate but also the second best, the third best etc. Instead of holding this against OT (meaningless predictions) the model assumes that the predicted relative grammaticality reflects what the speaker knows about variation.

| | Input | Cnstrnt 1 | Cnstrnt 2 | Cnstrnt 3 |
|----------------|-------------|-----------|-----------|-----------|
| ☞ ₁ | candidate 1 | | | * |
| ☞ ₂ | candidate 2 | | * | |
| ☞ ₃ | candidate 3 | * | | |

- (7) Predictions about frequency / tendencies:
- The best (most grammatical) candidates are most frequently observed
 - The worse ones are less frequent
 - There will be a cut-off point where everything below is too bad to ever occur as output. Finding the cut-off point is an empirical question.
- (8) A ROE-like model can also be used for categorical predictions about variation. As long as the grammar in (6) is really “fixed” it should allow for some variation patterns while others are predicted never to occur.
- Possible patterns: Cand1/*Cand2/*Cand3 (no variation); Cand1/Cand2/*Cand3 ; Cand1/Cand2/Cand3
 - Impossible patterns: Cand1/*Cand2/Cand3 ; *Cand1/Cand2/Cand3 ; *Cand1/Cand2/*Cand3

In other words: Variation patterns which violate the relative order of grammaticality are not possible.

- (9) Golston (1995) proposes that the syntax-phonology interface is reflected in the manner most natural to OT, through constraint ranking. More specifically he proposes that syntax outranks phonology. If all syntactic constraints are ranked above all phonological constraints we may not have a syntax which happens first but in fact such a precedence is implicit in the constraint ranking.

| Input | SYNTAX | PHONOLOGY |
|---------------|--------|-----------|
| ☞ candidate 1 | * / Ø | |
| candidate 2 | * / Ø | *! |

- (10) If there is no syntactic preference between two candidates, phonology will decide.
- (11) If we apply a ROE-like model of variation across the syntax-phonology interface with (10) in mind some patterns in syntactic variation should appear.

3 Case Variation in Icelandic

- (12) The phenomenon in Icelandic referred to as dative substitution (a.k.a. “dative sickness”) (Svavarsdóttir, 1982; Jónsson, 2003; Jónsson and Eyþórsson, 2003; Jónsson and Eythórsson, 2005) seems to be a promising place to look for “syntax doesn’t care” situations.
- (13) Dative substitution: A strong tendency in Modern Icelandic to replace lexically assigned accusative case on subjects with dative case. There seems to be no difference in meaning if the case changes. (There is more to say about dative substitution but this is the main thing.)
- (14)
- Hana vantar nýja skó.*
She-ACC needs new shoes.
‘She needs new shoes’
 - Henni vantar nýja skó.*
She-DAT needs new shoes.
‘She needs new shoes’
- (15) A lot of variation is observed in natural and experimental data. For a given verb some speakers will prefer accusative case, others will prefer dative and many will alternate between the two cases (an apparent “don’t care” situation).
- (16) Let us use the label SYNTAX for the effect that assigns lexical case.
- (17) And let us also define a prosodic effect which could influence the outcome of the algorithm:
- Prosodic words want to have two syllables (trochee sickness)
 - Two syllables are better than three
 - (Two syllables are better than one which is better than three?)

- d. $(\sigma\sigma) \gg (\sigma) \gg (\sigma\sigma\sigma)$
- e. The label for this is PHONOLOGY
- f. Can be implemented in various ways using well known constraints (binary feet, parsing syllables, initial feet, no clash, etc.) but the important thing here is that there is an active force defined in terms of phonology.
- g. Yes, this is an oversimplification, but it will do for the purposes of this talk.

(18) Hypothesis: When it comes to lexically assigned dative we will get three patterns of grammaticality judgements

- a. those who accept dative subjects consistently
- b. those who reject dative subjects consistently
- c. those who prefer the prosodically better case
- d. (if there is no prosodic difference the variation is unpredictable, as far as categorical predictions go)

(19) No speaker will be able to accept a prosodically bad dative subject with a particular verb while rejecting a prosodically better subject which is supposed to get the same kind of lexical case assignment (the same SYNTAX effect).

(20) Combining what we have said, the system which predicts the structure of the variation looks like this.

| Input | SYNTAX | PHONOLOGY |
|-----------------------------------|-----------------|-----------|
| $\text{☞}_1 (\sigma\sigma)$ | * / \emptyset | |
| $\text{☞}_2 (\sigma\sigma\sigma)$ | * / \emptyset | * |

And we are looking for patterns of the sort $(\sigma\sigma) / *(\sigma\sigma\sigma)$ while we don't expect to find $*(\sigma\sigma) / (\sigma\sigma\sigma)$.

4 Experiment

(21) A few words on methodology issues:

- a. Dative sickness is by far the best known example of prescriptive grammar in Icelandic
- b. Every child who goes through the Icelandic school system will get training in avoiding dative sickness – and to many people a linguist is someone who punishes people for being dative sick.
- c. Sometimes they will be able to avoid it, sometimes they won't, and some of them don't need to avoid it because they have the accusative lexical case in their grammar.
- d. Presenting a group of Icelanders with a judgement test with a handful of syntactic minimal pairs all involving dative sickness will cause a very unpredictable observer effect to add noise to the data.
- e. We don't want that so we make the controversial choice of ignoring the idea of syntactic minimal pairs to disguise the experiment (plus we include a ton of filler sentences).
- f. Further defence for this choice is provided below.
- g. And for the non-categorical (frequency) predictions the results are complemented with comparison to another recent study on dative sickness.

(22) 54 speakers, age 16, students of a commercial school in Reykjavik were asked to judge the grammaticality of the following sentences.

a. *Hvers vegna vantar Jón þessa nagla?*
Why needs John-ACC- σ those nails?

'Why does John need those nails?'

b. *Guðmund vantar nýjan jakka.*
Guðmundur-ACC- $\sigma\sigma$ needs new jacket.

'Guðmundur needs a new jacket.'

c. *Það er ljóst að Jóni vantar betri hugmynd.*
It is clear that John-DAT- $\sigma\sigma$ needs better idea.

idea.

'It is clear that John needs a better idea'

- d. *Vantar ekki Guðmundi bara stærri jeppa?*
Needs not Guðmundur-DAT- $\sigma\sigma\sigma$ just bigger jeep?
'Doesn't Guðmundur just need a bigger jeep.'

(23) About the sentences:

- a. But those are not even remotely syntactic minimal pairs! That's right, it is part of the disguise. In the current methodology approach AVOIDOBSERVEREFFECT outranks SYNTACTICMINIMALPAIRS.
- b. Still, all the boldfaced NPs should get lexically assigned case from the verb *vanta* 'need'.
- c. The only obvious reason for rejecting any of the sentences in (22) is the case of the boldfaced NP.
- d. Participants were able to give comments on the sentences and all comments on those sentences were about "wrong" case of the NP in question. (Comments on the numerous filler sentences revealed diverse opinions on all kinds of things.)
- e. Current literature on Icelandic syntax does not predict other reasons why the sentences should be ungrammatical to anyone.
- f. All sentences are really perfectly normal Icelandic unless you disagree with the case on the NP.

5 Results

(24) Frequency predictions about the sentences in (22):

- a. Speakers will have a tendency to prefer the ($\sigma\sigma$) subjects
- b. In judgements about accusative subjects more speakers will accept the ($\sigma\sigma$) *Guðmund* than the (σ) *Jón*.
- c. In judgements about dative subjects more speakers will accept the ($\sigma\sigma$) *Jóni* than the ($\sigma\sigma\sigma$) *Guðmundi*.
- d. N.B. Remember that the traditional view is that lexical case is assigned by the verb and thus the null hypothesis is that

acceptability will correlate with case (accusative or dative) and not with the form of the name (*Jón* or *Guðmundur*).

(25) Frequency results (54 speakers judge (22)):

| Structure | Subject | Positive judgements | Null hypothesis |
|------------------------------|----------|---------------------|-----------------|
| ACC-(σ) | Jón | 27 | 30 |
| ACC-($\sigma\sigma$) | Guðmund | 33 | 30 |
| DAT-($\sigma\sigma$) | Jóni | 40 | 25.5 |
| DAT-($\sigma\sigma\sigma$) | Guðmundi | 11 | 25.5 |

(26) The results confirm that there is a significant effect from prosody on judgements with dative subjects. The traditional hypothesis cannot account for this significant difference for the same case between word forms. Although the difference in ACC word forms goes in the right direction, it is not statistically significant.

(27) To further support this it is worthwhile comparing this result with a recent study on dative substitution (Búadóttir, 2007).

- a. In this interview study a few different types of tests gave the usual frequencies and tendencies we see in dative sickness studies.
- b. However, in one part of the study, the popularity of accusative subjects (forced choice, syntactic minimal pairs) was unusually high compared to other studies (such as Svavarsdóttir, 1982; Jónsson and Eyþórsson, 2003) and other parts of the same study.
- c. It turns out that all the subjects in this part (and only this part) of the study are names like *Guðmundur* which have two syllables in accusative case as opposed to three in the dative.
- d. This provides important support to the results in (25).

(28) But we want more than frequencies, so let's look at judgement patterns of individual speakers.

(29) 29 of the 54 participants were inconsistent in their judgements about grammaticality of dative subjects.

- a. All of those 29 speakers had the pattern *Jóni*/**Guðmundi*, no one had the pattern **Jóni*/*Guðmundi*.
- b. Current literature on Icelandic syntax does not predict one pattern to be more common than the other and it certainly doesn't predict one pattern to be impossible.
- c. If this really is a categorical effect this is a significant finding.
- d. Given a null hypothesis of equal distribution between the patterns (14.5/14.5), statistical significance is beyond any reasonable doubt ($p < 0.0001$).

(30) 18 of the 54 participants were inconsistent in their judgements about grammaticality of accusative subjects.

- a. We would like those 18 to have the judgement pattern **Jón*/*Guðmund*
- b. Actually, only 12 of them have that pattern
- c. 6 have the pattern *Jón*/**Guðmund*
- d. Maybe the assumption $(\sigma\sigma) \gg (\sigma)$ doesn't really hold?
- e. Or maybe the difference in relative grammaticality just isn't big enough for a categorical effect?
- f. Or maybe this has something to do with the original ("proper" in prescriptive terms) accusative case being somehow strong if present in the speaker's grammar (and then phonology has less of a determinative effect?)
- g. At least the pattern falls in the expected direction, 66.6% of the data fit the prediction.

6 Conclusions and Future Work

- (31) a. There definitely is a pattern in the data (25), (29) and (30), one that is not predicted by traditional accounts of Icelandic syntax.
- b. The idea that phonology will decide if syntax doesn't care seems to be relevant to what is going on.

- c. The proposed approach makes correct predictions about the frequencies in (25) and the categorical prediction about dative subjects in (29) holds. The categorical prediction about accusative subjects in (30) does not hold, but the data still fall in the direction of the general argument.
- d. This approach also explains results of another recent study on dative substitution (Búadóttir, 2007).
- e. However, more work needs to be done. ROE has been proposed as a variation model, for now, but as details of the analysis will be worked out, other approaches may prove to be more accurate.

(32) Answer to question in title: Maybe. (29) in particular seems to support this idea. Still, there is more work to be done.

References

- Anttila, Arto. 1997. Deriving variation from grammar. In *Variation, change, and phonological theory*, ed. Frans Hinskens, Roeland van Hout, and W. Leo Wetzels, 35–68. Amsterdam: John Benjamins.
- Boersma, Paul, and Bruce Hayes. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32:45–86. Available on Rutgers Optimality Archive, <http://rucss.rutgers.edu/roa.html>.
- Búadóttir. 2007. Þágufallshneigð. [Dative substitution.] B.A. thesis. University of Iceland.
- Coetzee, Andries. 2004. What it means to be a loser: Non-optimal candidates in Optimality Theory. Ph. D dissertation, UMass Amherst.
- Coetzee, Andries. 2006. Variation as accessing 'non-optimal' candidates. *Phonology* 23:337–385.
- Golston, Chris. 1995. Syntax outranks phonology: evidence from Ancient Greek. *Phonology* 12:343–368.
- Jónsson, Jóhannes G. 2003. Not so quirky: On subject case in Icelandic. In *New Perspectives on Case and Case Theory*. CLSI.

- Jónsson, Jóhannes G., and Thórhallur Eythórsson. 2005. Variation in subject case marking in Insular Scandinavian. *Nordic Journal of Linguistics* 28(2):223–245.
- Jónsson, Jóhannes G., and Þórhallur Eyþórsson. 2003. Breytingar á frumlagsfalli í íslensku. [Changes in subject case marking in Icelandic]. *Íslenskt mál og almenn málfræði* 25:7–40.
- Prince, Alan, and Paul Smolensky. 1993/2004. Optimality Theory: Constraint Interaction in Generative Grammar. RuCCS Technical Report 2, Rutgers University, Piscataway, NJ: Rutgers University Center for Cognitive Science. Revised version published 2004 by Blackwell. Page references to the 2004 version.
- Svavarsdóttir, Ásta. 1982. “Þágufallssýki” [‘Dative sickness’]. *Íslenskt mál og almenn málfræði* 4:19–62.