

Anton Karl Ingason

9 Rhythmic preferences in morphosyntactic variation and the theory of loser candidates

1 Introduction

This chapter considers cases where rhythmic well-formedness arbitrates the output of variable morphosyntax.¹ In simple intuitive terms: When syntax does not care, rhythmic preferences influence decision making. Two case studies from Icelandic and English are used to show how a generalized and restrictive typology of frequency distributions neatly falls out from a proposed Extended Rank-Ordered Model of EVAL (building on Coetzee, 2004, 2006).

Derivational approaches to variation in the variable-rules (VR) tradition (Labov, 1969, and subsequent work) have proven to be quite successful in accounting for variable outputs. However, such analyses are unconstrained in the sense that any frequency distribution can be described; a preference for the less-harmonic is as easily captured as harmonic optimization. If our theory of typological generalizations is to include possible/impossible frequency distributions, a probabilistic rules notation overgenerates typology. The current study shows how a comparison-based system of ranked constraints rules out analysis of the offending frequency distributions. This result echoes the success of comparison-based models in the context of work that seeks a tight relationship between formal analysis and predictions about typology.

Our data come from two sources, both of which involve an interaction between rhythm and variable morphosyntax. First, we look at an alternation between dative and accusative subject case in Icelandic. Two independent experiments confirm that acceptability of candidates is constrained by their relative rhythmic well-formedness. Independent evidence suggests that trochaic word forms are rhythmically more harmonic than other forms due to lapse/clash effects

¹ I would like to thank Kristján Árnason, Eiríkur Rögnvaldsson and Höskuldur Þráinsson for discussions on various aspects of this work. I thank Einar Freyr Sigurðsson and Joel C. Wallenberg for discussions and their annotation work in the IcePaHC corpus project. I also thank Laurel MacKenzie for valuable discussions on related joint work and participants at the DGfS workshop *Rhythm beyond the word* for insightful comments. Last but not least, I would like to thank Ralf Vogel and Ruben van de Vijver for organizing the DGfS workshop and this volume. All mistakes are mine and no one I have talked to necessarily agrees with me on anything.

and related phenomena. Note that Icelandic has word-initial stress. This rhythmic preference has consequences for Icelandic inflection classes where the dative (DAT) contributes an extra syllable over the accusative (ACC) as demonstrated by the masculine proper names in (1).

- (1) a. Jón-ACC-σ; Jóni-DAT-σσ
b. Guðmund-ACC-σσ; Guðmundi-DAT-σσσ

Our comparison-sensitive grammar will predict a rhythmic preference towards Jóni-DAT and Guðmund-ACC in contexts of ACC/DAT variability. The conditions for the preference only refer to units that are specific to linguistic competence, rather than social context or imperfections in production.

Second, we look at extraposition of relative clauses in Icelandic and English, showing a gradient effect of increased extraposition as an increased length of the relative clause makes the rhythmic structure less harmonic. Representative examples are given in (2) where it is much more natural to extrapose the longer relative clause than the short one.

- (2) a. People [who sing] are cool.
b. People are cool [who sing].
c. People [who sing about the paradise city where the grass is green and the girls are pretty] are cool.
d. People are cool [who sing about the paradise city where the grass is green and the girls are pretty].

Both patterns favor a model of variability that is sensitive to an operation that compares candidates based on how harmonic they are rather than applying a rule with an acquired probability. In the greater context of architectural dynamics we propose a *derive-and-compare* analysis, where ranked constraints replace the acquired probabilities, but not the rules, of variable-rules-style grammars, thus maintaining the locality enforcement of stepwise derivation (see Embick, 2010, on issues with purely global approaches) and its computational efficiency. Such a design has broad implications for the interaction between rhythm and morphosyntax since it predicts that every case of variability in syntax is sharply constrained by rhythmic factors and it categorically rules out frequency distributions in production that favor less-harmonic candidates.

The chapter is organized as follows: In Section 2 we discuss the grammatical architecture proposed, first by outlining in Section 2.1 why comparisons are

needed and second by introducing in Section 2.2 a formal framework for comparison, the Extended Rank Ordering Model of EVAL. In Sections 3–4 we present two case studies. Section 3 shows how prosodic preferences constrain variable case marking in Icelandic and Section 4 adds support to our *derive-and-compare* approach by looking at the relationship between heaviness and relative clause extraposition in Icelandic and English. Section 5 concludes.

2 Derive-and-compare

2.1 Comparison-based decisions

Let us outline our assumptions about the *derive-and-compare* architecture. First, we assume that structure building exists and that the structures built are not infinite. We also assume that multiple structure building grammars are needed for each individual speaker to account for regular variability as motivated by work on sociolinguistics (Labov, 1969, 2006), diachronic syntax (Kroch, 1989) and learning mechanisms in language acquisition (Yang, 2002). The *derive-and-compare* architecture departs from such approaches to variable output by proposing that the mechanism that selects between multiple options is sensitive to system-internal comparison rather than acquired probabilities.

By comparison we mean a function that answers the question: Which of two is more? The proposal here is that the grammar ranks the outputs of multiple grammars using comparison and that the grammar is sensitive to the size of a difference so that it restricts not only the structure of one-input-to-many-outputs but also the structure of many-inputs-to-many-outputs. Optimality Theory (OT) is a set of frameworks that all apply comparison in one way or another to select candidates from a set of options. For this reason it is reasonable to draw on OT for implementing a *derive-and-compare* grammar. Most flavors of OT, including the original proposal (Prince & Smolensky, 1993/2004), assume that the result of the comparison is only one candidate, the optimal one which is output. Classic OT produces simply one output and OT models of variability generally involve the probability of a particular constraint ranking (= grammar) rather than making full use of the properties of comparison, cf. partial ranking that results in multiple grammars (Anttila, 1997) and Stochastic OT which injects noise into the system to make multiple grammars possible (Boersma & Hayes, 2001).

The emphasis on a single output in OT is irregular. The architectural dynamics that come with comparison have potential to restrict the distribution of output in ways that are hard to build into weighted rules if the system is really sensitive

to the question: Which one of two is more? If the empirical object of comparison is harmony between preservation of contrast and minimal effort, then variable weighted rules are destined to overgenerate typology because they cannot exclude a preference for a less harmonic candidate in variable contexts. Therefore, a multiple output OT-style system is worth considering.

A notable exception to single output OT is the Rank-Ordering Model of EVAL (ROE) (Coetzee, 2004, 2006). The model is motivated by the fact that whether we like it or not an OT grammar predicts the rank of all candidates, not just which one is optimal. If we remove the optimal candidate from the set, we can calculate which one is the second best, third best, etc. Rather than taking this as a flaw and claiming that the grammar is overgenerating meaningless information, Coetzee proposes that the rank-ordering imposed by the EVAL function arbitrates the structure of variable output. We will adopt and extend this proposal in Section 2.2.

The proposal will aim to capture the following types of facts:

- If a single input can surface as two different outputs, the more harmonic output is more frequent. (Coetzee's ROE model)
- The relative size of a win in a comparison determines the relative size of an effect. Therefore, all else being equal, a candidate which is preferred by a big margin is more frequent in production than a candidate which is preferred by a small margin. (Extended ROE)

Such facts are not easy to capture within a system of rules with acquired probabilities because a system which relies on probabilities and is not sensitive to comparison does not inherently rule out a higher probability for a less harmonic form. The only solution is to blame the non-grammar part of the brain. If the facts however do fall out from systematic relationships between the units that are specific to linguistic competence – then a model that captures the facts in a unified way is superior to a model that does not.

2.2 Extended Rank-Ordering Model of Eval

Whether we like it or not, we can read the relative rank of candidates from an OT tableau. Under Classic OT assumptions about a single output this fact is a strong argument against OT since it implies great (or even infinite) amounts of meaningless overgeneration. The Rank-Ordering Model of Eval (ROE) (Coetzee, 2004, 2006) does away with the single output requirement and claims that the extra output is in fact meaningful. In ROE the relative rank imposed by EVAL makes predictions about the structure of variation; the model makes use of the dynamics of comparison. A ROE tableau is demonstrated in (3).

(3) ROE tableau

	Input	Cnstrnt1	Cnstrnt2	Cnstrnt3
\mathbb{S}_1	candidate 1			*
\mathbb{S}_2	candidate 2		*	
	candidate 3	*!		

ROE proposes a cut-off point between highly ranked (more important) and less highly ranked (less important) constraints. The cut-off is indicated by a thick line in the tableau. A candidate that is only disfavored by constraints below cut-off is not ungrammatical and it is a possible output. It is however less frequent in production than the more harmonic candidates. The frequency distribution falls out from a single grammar and does not involve re-ranking the constraints. In tableau (3), candidates 1 and 2 are both grammatical, but candidate 1 is more frequent than candidate 2. Candidate 3 is disfavored by a constraint above cut-off and is therefore ungrammatical.

ROE has pleasing properties when it comes to the usual variability problem of mapping one input to two or more outputs: The structure of the gradience is predicted by the grammar. An approach in this spirit would however be even more feasible if it could also capture gradience across inputs. The most simple approach is to take the outputs of distinct inputs and apply the rank-ordering EVAL function as in (4).

(4) Simple cross-input comparison in ROE (to be rejected)

		Cnstrnt1	Cnstrnt2
\mathbb{S}_1	Output from input 1		*
\mathbb{S}_2	Output from input 2	*	

An analysis in this spirit has been applied to the [sC,VC_i]-restriction in English (Coetzee, 2008) but it has drawbacks as pointed out by Coetzee (2010). First, such an analysis can result in an ungrammatical candidate being relatively more well-formed than a grammatical candidate. This makes little sense if the model is supposed to reflect real relationships. Second, there is no such thing as the size of an effect when ROE is applied this way and therefore the model cannot distinguish a small preference from a larger preference, which is a serious flaw if the main goal is to make predictions about frequency distributions in variable contexts.

A crucial feature of a model that is able to capture cross-input comparisons is the ability to capture the size of an effect. Pater (2007), followed by Coetzee & Pater (2008) and Coetzee (2010), addresses this problem by (I) introducing a com-

petition between the optimal candidate and the best alternative and (II) resorting to weighted constraints in an Harmonic Grammar framework. The acceptability of an output form is then calculated by subtracting the harmony score of the best alternative from each candidate's harmony score. We will adopt the first observation that comparison with the best alternative is needed in a theory of cross-input comparison but we will argue that once we do that the second step of abandoning ranked constraints is not needed and should be avoided. As Pater (2007) notes: "Much remains to be done in terms of further developing and evaluating HG as a theory of typology and of gradient acceptability, as well as better understanding the nature of cumulativity in different domains." As for gradient acceptability across inputs, there are not many frameworks that claim to have a mature solution. However, the ranked constraints of OT do provide important and well developed insights into typology and this progress should not be dispensed with in the absence of an alternative that matches it or does better.

The reasoning for our Extended ROE is exactly the same as the original motivation for the ROE approach. If we can read out of an OT tableau the relative size of an effect this could be interpreted as meaningless overgeneration but we will take it to be a meaningful predictor of frequency distributions. To describe this we need to introduce the concept of a Ranking Span:

(5) **Ranking Span of a win in Extended ROE:**

The span between the rank of the highest ranked constraint that favors the winner and the rank of the highest ranked constraint that favors the best alternative.

The tableaux in (6) and (7) use a shaded background to indicate the Ranking Span which reflects the size of the win in each case.

(6) **Ranking Span – small effect:**

	Input 1	CONSTRNT1	CONSTRNT2	CONSTRNT3
res_1	Output 1.1			*
res_2	Output 1.2		*	

(7) **Ranking Span – larger effect:**

	Input 2	CONSTRNT1	CONSTRNT2	CONSTRNT3
res_1	Output 2.1			*
res_2	Output 2.2	*		

Within each of the tableaux we get exactly the same dynamics as in Standard ROE (henceforth SROE). Additionally, by observing the different Ranking Spans we get the benefits of comparison with the best alternative without resorting to weighted constraints. Without adding any extra elements to the SROE grammar the Extended ROE Grammar (henceforth EROE) predicts that the frequency difference in (7) is greater than in (6) because the smaller Ranking Span is a subset of the larger Ranking Span. The benefit of the EROE approach is that if we have independent evidence for a particular preference in the grammar, we can now explain constraints on frequency distributions whose conditions are internal to the grammar by comparing Ranking Spans which are in a subset relationship.

If the EROE approach proves to make the correct predictions it is clearly superior to an alternative which has to overgenerate typology (I) by allowing less harmonic options to be more frequent or (II) by allowing smaller differences in harmony to result in larger effects. It would not save a VR grammar to blame the rest of the brain for those patterns because the patterns are formulated in terms of units that are only meaningful to computation withing the linguistic competence and are therefore scope-wise internal to the grammar. An explanation which fails to capture such predictions about frequency distributions also fails to capture a close relationship between categorical and gradient output patterns. In the following sections we will look at two case studies which demonstrate the advantages of the EROE approach.

3 Rhythmic constraints on variable morphosyntax

3.1 Variable subject case

Our first case study involves prosodic constraints on variable subject case in Icelandic. After a brief overview of the data in question we will show the categorical typology of variable output that is captured by our system. Icelandic is generally assumed to have oblique (non-nominative) subjects (see Andrews, 1976; Thráinsson, 1979) in addition to the canonical nominative subject case. A full description of this phenomenon is beyond the scope of this chapter but we will outline the facts that are most important for our discussion, primarily the Elsewhere Condition (Kiparsky, 1973) hierarchy of experiencer subjects. For a more detailed discussion of the semantic patterns, see Jónsson (2003). Experiencers are generally dative, thus blocking the canonical nominative case, and small subsets (and some individual verbs) of experiencers, such as experiencers of physical discomfort or lacking, get accusative case in standard Icelandic, thus blocking the

more general dative case of experiencers. Agents are always nominative. Representative examples are given in (8).

- (8) a. Pál þyrstir í kaffi.
Paul-ACC thirsts in coffee
'Paul has thirst for coffee.'
- b. Páli líkar kaffi.
Paul-DAT likes coffee
'Paul likes coffee.'
- c. Páll drekkur kaffi.
Paul-NOM drinks coffee
'Paul drinks coffee.'

The hierarchy can be formally expressed as follows:

- (9) Subject case hierarchy in Old Icelandic (Ingason, 2010)
IF [+experiencer, +lacking]
THEN apply ACC
ELSE IF [+experiencer]
THEN apply DAT
ELSE
apply NOM (default)

This pattern is quite stable in the historical record but since the 19th century the contexts for accusative subjects (physical discomfort/lacking), have lost ground. Thus, many speakers of Modern Icelandic alternate freely between dative and accusative subject case with verbs like *þyrsta* 'thirst', which results in *Paul.DAT thirsts for coffee*, cf. (8a). The development is called Dative Substitution (henceforth DS) or, by prescriptivists who eagerly oppose the change, Dative Sickness (Ice. *þágufallssýki*). The change has been thoroughly documented (e.g. Barðdal, 2011; Eyþórsson, 2000; Halldórsson, 1982; Ingason, 2010; Jónsson & Eythórsson, 2005; Svavarsdóttir, 1982; Viðarsson, 2009).

A closely related development is the modern tendency to replace standard nominative on two experiencer verbs, *hlakka* 'look forward to' and *kviða* 'be anxious about', with dative or accusative case, resulting in intra-speaker alternations between nominative, accusative and dative case with the two verbs. This effect has been dubbed Oblique Sickness (Ice. *aukafallssýki*) (Eyþórsson, 2000). Previous studies on the nature of variable subject case in Icelandic indicate that it is not the case that different speakers acquire different grammars. On the contrary, studies

that allow for detection of intra-speaker variation all suggest that a single speaker will frequently accept and produce more than one subject case with the same verb (Búadóttir, 2007; Ingason, 2010). Therefore, this seems to be a genuine case where morphosyntax can allow for more than one option.

Búadóttir (2007) used a combination of experimentation techniques to study DS in an interview study with 30 women of age 25–30. Her results confirm that the same speaker will frequently accept both accusative and dative case with the same verb. For all the five DAT-ACC alternating verbs tested, over 50% of the participants accepted both cases in a judgment task. In the elicited production part of her experiment, the same speaker would sometimes produce alternating subject cases with the same verb. Crucially, for the current discussion, the forced choice part of the study was a clear outlier. In this part, there was a stronger tendency to choose accusative subjects for both ACC-DAT alternating verbs and NOM-ACC-DAT alternating verbs, than there was in other parts of the same study.

One might suspect that this increased preference for the accusative is an effect of prescription since DS is considered to be incorrect from a prescriptive point of view. This explanation is highly suspicious due to the fact that accusative case is the most frequently selected case for both the ACC-DAT alternating verbs, where it is consistent with the accusative prescription, and the NOM-ACC-DAT verbs, where it goes against the nominative prescription. One can of course speculate that an experimental technique that emphasizes a contrast between minimal pairs that are subject to prescription will result in an unpredictable bias of some sort, but in the absence of an explanation for the opposite behavior of the two classes of verbs, a more straightforward causal account is desirable.

Here, we propose that the bias comes from an interaction with rhythm rather than a contradictory interaction with prescription. Multiple phenomena demonstrate a preference for trochaic rhythm in Icelandic (e.g. Ingason, 2008). Prosodically improving input to output mappings frequently result in a monosyllabic form gaining an extra syllable or multi-syllable forms losing syllables resulting in two syllable outputs. What does this mean for case? For some inflection classes there is no prosodic contrast between accusative and dative, cf. the case syncretism seen in the proper names in (10), but for others the dative contributes an extra syllable, cf. (11). It turns out that unlike other parts of Búadóttir's study, the forced choice experiment included only subjects of type (11b), where the accusative is the more eurhythmic option.

- (10) a. Önnu-ACC-σσ; Önnu-DAT-σσ
b. Árna-ACC-σσ; Árna-DAT-σσ

(11) a. Jón-ACC-σ; Jóni-DAT-σσ

b. Guðmund-ACC-σσ; Guðmundi-DAT-σσ

To complement Búadóttir's study we designed an experiment where the main methodological goal was to disguise the pattern being tested. To this effect, we used variable sentence types and a high number of fillers to avoid the prescriptive effect that might have been present in Búadóttir's study. 54 participants of age 16 gave their acceptability judgments.

To test this prediction we constructed 4 sentences where the verb *vanta* 'lack, need' was presented with each of the four forms above as the subject.

(12) 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-σσ needs new jacket.
'Guðmundur needs a new jacket.'c. Það er ljóst að Jóni vantar betri hugmynd.
it is clear that John-DAT-σσ needs better idea
'It is clear that John needs a better idea.'d. Vantar ekki Guðmundi bara stærri jeppa?
needs not Guðmundur-DAT-σσ just bigger jeep
'Doesn't Guðmundur just need a bigger jeep.'

All of the sentences are perfectly normal Icelandic apart from the possibility of variable subject case marking preferences. Participants were allowed to comment on their judgments and all the comments involved the case marking of the subject in question.

First, by looking at the general acceptability pattern of *Jón* vs. *Guðmundur*, the predictions are borne out as 50% (27/54) accepted the sentence with *Jón*-ACC-σ as the subject whereas 61% (33/54) accepted the sentence with *Jóni*-DAT-σσ, a slight effect of a preference for the bisyllabic form. The effect for the other name is larger as predicted: 74% (40/54) accepted the form with *Guðmund*-ACC-σσ as the subject but only 20% (11/54) accepted the three syllable form *Guðmundi*-DAT-σσ.

A standard account of case would predict that the selection of lexical case is only sensitive to the semantics of the verb of lacking but not to the form of the subject. Even if we allow for variability, a pure-morphosyntax account would therefore expect the same rate of accusative acceptance with each name and

the same goes for dative acceptance. Moreover, under standard assumptions, we would not expect individual speakers to give different acceptability judgments on the same subject case with the same verb. However, we found that 18 speakers were inconsistent in whether they accepted or rejected accusative subjects with our verb of lacking and 29 speakers were inconsistent in whether they accepted or rejected the dative forms. Looking only at those inconsistent speakers, 12 judged the accusative *Guðmund*-σσ/**Jón*-σ but only 6 gave the judgments **Guðmund*-σσ/*Jón*-σ, which falls in the direction of the prediction. More strikingly, there was a categorical effect for the datives, where all of the 29 inconsistent participants gave the judgment **Jóni*-σσ/*Guðmundi*-σσσ and the pattern *Jóni*-σσ/**Guðmundi*-σσσ was unattested.

This pattern would be mysterious without system internal comparisons but it can be easily explained if we assume that two grammars derive the options which are then compared for harmony. The crucial point is not the implementation details of a particular analysis but rather that the typology of frequency distributions is most naturally captured by a comparison operation rather than with a probabilistic application of a rule. Let us demonstrate how this might work for the case alternation under investigation.

For the rhythmic markedness effect we will declare word forms of 1 or 3 syllables a rhythmic violation which result from lapse/clash effects and related phenomena. The OT literature provides analyses that could be substituted for the current descriptive constraints. In simple terms, the effort-minimizing half of an OT-style system seeks eurythmic output of beyond-the-word contexts (effects of this sort are also discussed and modeled by Shih et al. (this volume)).

(13) *3σ: 3 syllable prosodic words are rhythmically bad (lapse effect)

(14) *1σ: monosyllabic prosodic words are rhythmically bad (clash effect)

We assume that in Old Icelandic, prior to the rise of Dative Substitution, there was an effect in place that required semantic features like [+physical discomfort] and [+lacking] to be contrasted with plain dative experiencer subjects.

(15) ACC: Express subjects with the features [+physical discomfort] and [+lacking] with accusative morphology

Nominative is the default case in Icelandic so we assume there to be markedness constraints which disfavor dative and accusative morphology. The accusative morphology is less marked (even if its subject case semantics are more marked)

as can be argued on the basis of syncretism patterns (Rögnvaldsson, 1990) and therefore we assume the ranking *DAT >> *ACC.

- (16) a. *DAT: No dative morphology
b. *ACC: No accusative morphology

In the following tableau we have a SROE grammar that accounts for Búadóttir's results where two bisyllabic accusative is favored when there is a prosodic contrast between the accusative and the dative. We use the * / ∅ notation to indicate that the distinction represented by the constraint (in Old Icelandic) is no longer there. The *1σ constraint is included for easier comparison with the upcoming tableaux.

- (17) Preference for 2σ over 3σ

	[+exp,+lack]	Acc	*3σ	*1σ	*DAT	*ACC
☞ ₁	Guðmund-ACC-σσ	* / ∅				*
☞ ₂	Guðmundi-DAT-σσσ	* / ∅	*		*	

Here, the accusative form is preferred because it is prosodically better than the accusative form. The dative remains a grammatical option since it is only disfavored below cut-off but it is less harmonic, less frequent in production and less acceptable. We will of course have to add more data to the picture for further support. For the time being the effect is intuitively an example of phonology having an opinion when morphosyntax does not care. The preference should be reversed if the accusative is monosyllabic and the trochaic form is dative as for the name *Jon*:

- (18) Preference for 2σ over 1σ

	[+exp,+lack]	Acc	*3σ	*1σ	*DAT	*ACC
☞ ₂	Jón-ACC-σ	* / ∅		*		*
☞ ₁	Jóni-DAT-σσ	* / ∅			*	

Moreover, assuming this ranking, EROE predicts that the difference in relative well-formedness is greater between the forms *Guðmund*/*Guðmundi* than *Jón*/*Jóni* as we see if we add the Ranking Span information to the tableaux because the smaller Ranking Span is a subset of the larger Ranking Span:

- (19) Ranking Span indicates a large preference effect for ACC

	[+exp,+lack]	Acc	*3σ	*1σ	*DAT	*ACC
☞ ₁	Guðmund-ACC-σσ	* / ∅				*
☞ ₂	Guðmundi-DAT-σσσ	* / ∅	*		*	

- (20) Ranking Span indicates a small preference effect for DAT

	[+exp,+lack]	Acc	*3σ	*1σ	*DAT	*ACC
☞ ₂	Jón-ACC-σσ	* / ∅		*		*
☞ ₁	Jóni-DAT-σσσ	* / ∅			*	

Even if these results are promising, we need a better motivated constraint ranking to properly test the EROE. In the following section we will turn to a case study of EROE predictions in an inherently fixed constraint ranking.

4 Heavy relative clause extraposition

As demonstrated in (21) and (22), heaviness makes relative clause extraposition more acceptable:

- (21) a. People [**who sing**] are cool.
b. People are cool [**who sing**].
- (22) a. People [**who sing about the paradise city where the grass is green and the girls are pretty**] are cool.
b. People are cool [**who sing about the paradise city where the grass is green and the girls are pretty**].

It is quite unnatural to extrapose a relative clause which consists only of the verb and an extracted subject. The extraposed variant becomes much better if we add a number of words as in (22). The interpretation of heavy relative clause extraposition (HRE) in terms of constraint interaction is straightforward. Breaking up the matrix clause causes a rhythmic lapse effect: The relative clause splits the subject and the predicate and despite the rhythmic nature of the intervening material, the prominence of the beats is too low and therefore the whole configuration is arhythmic.

Extrapositing the relative clause avoids rhythmic lapse, but it also introduces an attachment task and therefore it increases the chance of a misunderstanding (for the relevant psycholinguistic experiments, see MacWhinney & Pleh, 1988). Note that we assume that prosodic complexity is to blame rather than syntactic complexity since a syntactic complexity measure of heaviness incorrectly predicts examples like (23) to be light whereas the prosodic account has no such problems (Ingason & MacKenzie, 2011).

- (23) a. People [who sing lalalalalalalala] are cool.
 b. People are cool [who sing lalalalalalalala].

An important aspect of HRE is that heaviness is a continuous predictor which leads to difficulties (and notational unrestrictiveness) for comparison free approaches to variability. To confirm this we coded relative clauses from two corpora for extraposition and number of words. IcePaHC 0.5 (Wallenberg et al., 2011) is an Icelandic corpus of ca. 632,000 words and PPCEME (Kroch et al., 2004) is an English corpus of ca. 1,794,000 words. The study included an exhaustive sample of relative clauses on subjects and objects in matrix clauses and finite subordinate clauses in the two corpora.

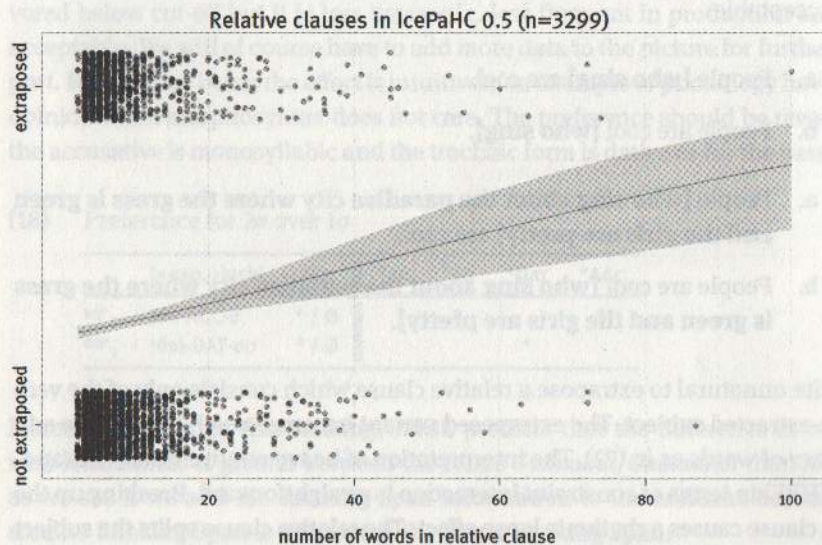


Figure 1: Relative clause extraposition in IcePaHC 0.5 (Wallenberg et al., 2011)

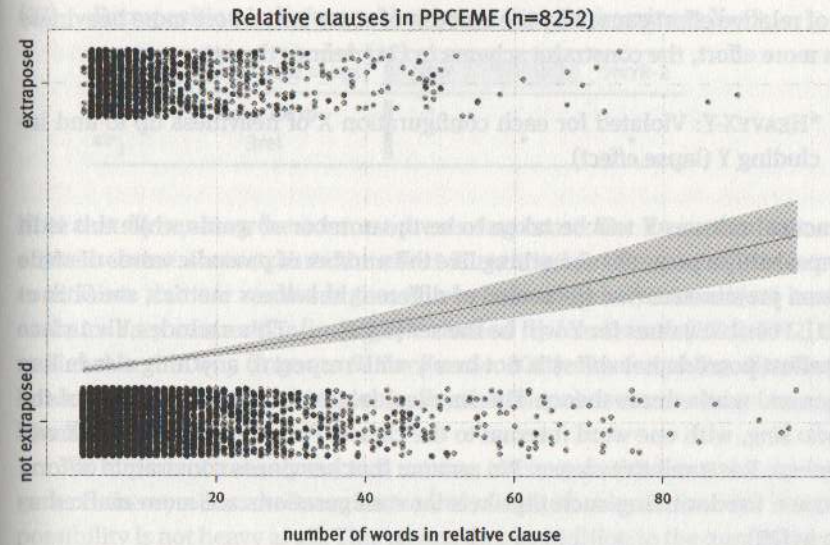


Figure 2: Relative clause extraposition in PPCEME (Kroch et al., 2004)

The results are shown in Figures 1 and 2 (drawn using R (R Development Core Team, 2011) and ggplot2 (Wickham, 2009)). Each dot represents one clause. A dot at the top is extraposed and one at the bottom is not. The x-axis shows the number of words in the clause. The logistic regression line shows the same gradual effect in both languages. The data are obviously sparse for the extra-long cases but the gradience is convincingly robust up to over 40 words.

If we wanted to build this effect into a variable-rules style grammar we would need to multiply an acquired probability with a continuous heaviness value – which would greatly complicate the probability matching of a learner pursuing a VR architecture. Moreover, if we are not allowed to compare harmony, the grammar would have to introduce a counting mechanism to access the heaviness value to be used for multiplication. Even if we accept that, the weighted-rules approach fails in its design to capture the fact that HRE does in fact result from comparing options for harmony. Therefore the most logical option is to have comparison-based evaluation in the grammar. If we assume a *derive-and-compare* architecture, this means that we have one structure building machine that derives each word order and then a comparison based evaluation component that compares the two options and is sensitive to the size of any preference there is for one of them.

We have to define a set of constraints to handle heaviness comparison in the proposed framework. As we hypothesize that heaviness effects in general are the

result of relative effort caused by the amount of some unit where more heaviness causes more effort, the constraint schema in (24) defines the set:

- (24) *HEAVYX-Y: Violated for each configuration X of heaviness up to and including Y (lapse effect)

For practical reasons Y will be taken to be the number of words while this is in reality probably a proxy for something like the number of prosodic words of some minimum prominence (for discussion of different heaviness metrics, see Shih et al. 2011). Possible values for Y will be the set {2,3,4,...}. This excludes Y = 1 since the smallest possible heaviness is not heavy with respect to anything else in line with general markedness theory. The smallest relative clause is therefore of the type *who sing*, with one word internal to the clause. For our study of HRE, X will be R, where R is a relative clause. We assume that heaviness constraints of form (24) have a fixed ranking such that heavier configurations are more marked as shown in (25).

- (25) Fixed ranking of heaviness hierarchies

A hierarchy of heaviness constraints has a fixed ranking such that *HEAVYX-Y₁ always outranks *HEAVYX-Y₂ if Y₁ > Y₂.

Since most theories of variability focus on the trivial case where one input can be mapped to more than one output, they do not necessarily have much to say about a distinction like *who sing* vs. *who sing songs* – those are two different inputs. EROE uses the concepts of comparison with the best alternative and the Ranking Span to measure effect size. If we label the contrast preserving constraint STAY (do not extrapose) and rank this constraint above the heaviness hierarchy, the grammar predicts the observed preference for not extraposing short relative clauses, it allows for extraposition and the extraposition grammar is a relatively better option if the clause is heavier as seen in the following tableaux where the smaller Ranking Span is a subset of the larger Ranking Span:

- (26) Extraposition is disfavored but possible

	who [sings songs]	STAY	*HVYR-3	*HVYR-2
RS ₂	2rel + extrp	*		
RS ₁	2rel			*

- (27) Extraposition is disfavored but better than for heaviness Y = 3

	who [sings nice songs]	STAY	*HVYR-3	*HVYR-2
RS ₂	3rel + extrp	*		
RS ₁	3rel		*	*

Note that by ranking the contrast preserving constraint lower than some or all of the heaviness constraints the effect is reversed for the part of the heaviness constraints that are ranked higher. The relationship between subject heaviness and auxiliary contraction in English studied by MacKenzie (2011, 2012) can be modeled in this way (Ingason & MacKenzie, 2011). The frequency distribution is however always restricted in such a way that more heaviness always creates less harmonic output.

An important consequence of moving to comparative heaviness based on a general theory of markedness instead of weighted rules is that the least heavy possibility is not heavy at all. This means that in addition to the quantitative effect, relative clauses of the type *who sings* are qualitatively different from all longer types. This prediction is very clearly borne out in the corpus study, cf. Figures 1 and 2. One word relative clauses are practically never extraposed while two word relative clauses do so frequently and fall in a more general pattern. Building this special condition into a weighted rules system would be very difficult without adding a decision making component that compares candidates. In our approach the prediction comes for free from the independently motivated theoretical observation that the least marked type does not violate effort constraints.

5 Conclusion

The main goal of this paper is to present case studies which demonstrate the need for comparison-based evaluation of candidates in patterns that span the morphosyntax-phonology interface. The conclusion is that all cases of variable syntax will be constrained by the grammar in categorical as well as in gradient ways and that a grammatical architecture which supports cross-input comparison does not have to resort to weighted constraints because the ranking mechanisms of OT-style systems can be extended along the lines of the Rank Ordering Model of EVAL to capture the relevant generalizations and restrict the typology of variability. We do not suggest a system where comparison is somehow internal to the structure building process but rather that multiple grammars, which we need anyway to account for historical change, produce the candidates to be compared.

The case study on variable subject case in Icelandic shows that prosodic constraints on variable morphosyntax can be quite robust. The most striking result is that inconsistent judgement patterns about dative subjects with the same lexically case assigning verb were categorically in the direction of a prosodic preference where all 29 speakers in this group made the judgment * $\sigma\sigma\sigma/\sigma\sigma$ whereas the pattern $\sigma\sigma\sigma/*\sigma\sigma$ was unattested. The *derive-and-compare* approach was further supported by the rhythmic effect on heavy relative clause extraposition as the gradient pattern encountered in both Icelandic and English points to a system that is sensitive to which of two candidates is more harmonic.

References

- Andrews, Avery. 1976. The VP-Complement Analysis in Modern Icelandic. *Proceedings of NELS* 6. 1–21.
- Anttila, Arto. 1997. Deriving variation from grammar. In Frans Hinskens, Roeland van Hout, & W. Leo Wetzels (eds.), *Variation, Change, and Phonological Theory*, 35–68. Amsterdam: John Benjamins.
- Barðdal, Jóhanna. 2011. The Rise of Dative Substitution in the History of Icelandic. *Lingua* 121. 60–79. Special volume on Case Variation. Eds. Klaus von Heusinger and Helen de Hoop.
- Boersma, Paul & Bruce Hayes. 2001. Empirical tests of the gradual learning algorithm. *Linguistic Inquiry* 32. 45–86. Available on Rutgers Optimality Archive, <http://ruccs.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. 2010. Gradient well-formedness in Harmonic Grammar: Phonological performance as a window on phonological competence. *Journal of the Phonetic Society of Japan* 14. 13–23.
- Coetzee, Andries W. 2006. Variation as accessing 'non-optimal' candidates. *Phonology* 23. 337–385.
- Coetzee, Andries W. 2008. Grammaticality and ungrammaticality in phonology. *Language* 84. 218–257.
- Coetzee, Andries W. & Joe Pater. 2008. Weighted constraints and gradient restrictions on place co-occurrence in Muna and Arabic. *Natural Language & Linguistic Theory* 26. 289–337.
- Embick, David. 2010. *Localism versus Globalism in Morphology and Phonology*. Cambridge, MA: The MIT Press.
- Eyþórsson, Þórhallur. 2000. Fall á fallanda fæti? Um breytingar á frumlagsfalli í íslensku. *Íslenskt mál* 22. 185–204.
- Halldórsson, Halldór. 1982. Um méranir. Drög að samtímalegri og sögulegri athugun. *Íslenskt mál* 4. 159–189.
- Ingason, Anton Karl. 2008. Hrynkerfi íslensku í bestunarkenningu. [Icelandic prosody in Optimality Theory.]. B.A. Thesis. University of Iceland.
- Ingason, Anton Karl. 2010. Productivity of Non-Default Case. *Working papers in Scandinavian syntax* 85. 65–117.
- Ingason, Anton Karl & Laurel MacKenzie. 2011. Heaviness as evidence for a derive-and-compare grammar. Poster presented at the Manchester Phonology Meeting.
- Jónsson, Jóhannes G. 2003. Not so Quirky: On Subject Case in Icelandic. In *New Perspectives on Case Theory*, eds. Ellen Brandner & Heike Zinsmeister, 127–163. CLSI Publications.
- Jónsson, Jóhannes G. & Thórhallur Eythórsson. 2005. Variation in subject case marking in Insular Scandinavian. *Nordic Journal of Linguistics* 28.2. 223–245.
- Kiparsky, Paul. 1973. 'Elsewhere' in Phonology. In *A Festschrift for Morris Halle*, 93–106. New York: Holt, Rinehart & Winston.
- Kroch, Anthony S. 1989. Reflexes of Grammar in Patterns of Language Change. *Language Variation and Change* 1. 199–244.
- Kroch, Anthony S., Beatrice Santorini & Lauren Delfs. 2004. Penn-Helsinki Parsed Corpus of Early Modern English. CD-ROM. First Edition. Size: 1.8 million words.
- Labov, William. 1969. Contraction, Deletion, and Inherent Variability of the English Copula. *Language* 45. 715–762.
- Labov, William. 2006. *The social stratification of English in New York city*. Cambridge Univ Pr.
- MacKenzie, Laurel. 2011. Quantitative data as a clue to auxiliary contraction processes. Paper presented at the 2011 annual meeting of the Linguistic Society of America.
- MacKenzie, Laurel. 2012. Locating variation above the phonology. Ph.D. Dissertation. University of Pennsylvania.
- MacWhinney, Brian & Csaba Pleh. 1988. The processing of restrictive relative clauses in Hungarian. *Cognition* 29. 95–141.
- Pater, Joe. 2007. Cumulative Ill-Formedness in Typological and Experimental Data. Handout from the workshop on Experimental Approaches to Optimality Theory, University of Michigan, Ann Arbor.
- Prince, Alan & 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.
- R Development Core Team. 2011. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <http://www.R-project.org>, ISBN 3-900051-07-0.
- Rögnvaldsson, Eiríkur. 1990. *Íslensk orðhlutafræði. [Icelandic morphology]*. Reykjavík: Institute of Linguistics, University of Iceland.
- Svavarsdóttir, Ásta. 1982. "Þágufallssýki" ['Dative Sickness']. *Íslenskt mál og almenn málfræði* 4. 19–62.
- Thráinsson, Höskuldur. 1979. *On Complementation in Icelandic*. New York: Garland.
- Viðarsson, Heimir Freyr. 2009. Tilbrigði í fallmörkun aukafallsfrumlaga: Þágufallshneigð í forníslensku. *Íslenskt mál* 31. 15–66.
- Wallenberg, Joel C., Anton Karl Ingason, Einar Freyr Sigurðsson & Eiríkur Rögnvaldsson. 2011. Icelandic Parsed Historical Corpus (IcePaHC). Version 0.5. Size: 632 thousand words. URL http://www.linguist.is/icelandic_treebank.
- Wickham, Hadley. 2009. *ggplot2: elegant graphics for data analysis*. Springer-Verlag New York Inc.
- Yang, Charles. 2002. *Knowledge and Learning in Natural Language*. Oxford: Oxford University Press.