Gradient Grammaticality as an Effect of Selective Constraint Re-ranking

Frank Keller
University of Edinburgh

1 Introduction

The validity of a grammatical framework can be verified in at least the three ways: (a) by showing its applicability to wide range of linguistic phenomena, (b) by demonstrating the soundness of its formal foundations, and (c) by verifying its compatibility with experimental evidence. As for Optimality Theory (OT; Prince and Smolensky 1993), option (a) has been pursued extensively in the recent phonological literature (and to a lesser extend in the syntactic literature). Also option (b) has been the the topic of some research (e.g., Ellison 1994; Karttunen 1998). However, no attempts have been made so far to test the concepts and mechanisms assumed in OT against experimental evidence.

The present paper attempts to fill this gap by testing OT against evidence from what is probably the most natural empirical domain for a linguistic framework: grammaticality judgments. More specifically, we focus on the phenomenon of gradience in linguistic data. We argue that gradient data can serve as a tool for evaluating the status of suboptimal candidates in OT, an approach that allows to scrutinize OT's concepts of constraint ranking and constraint interaction. The experimental data we present show that constraint violations are cumulative, and that two types of constraints have to be distinguished: hard and soft ones. These results lend limited support to the notion of constraint ranking assumed in OT, and seem compatible with OT's concept of strict domination of constraints.

The second part of this paper deals with the theoretical issues arising from an attempt to model gradient linguistic data in OT. We show that a naive model that equates relative grammaticality with relative optimality is not tenable, and propose an alternative approach based on the concept of selective constraint re-ranking. This approach, which is grounded in OT learnability theory, predicts the cumulativity of constraint violations, and allows to model the distinction between hard and soft constraints, thus accounting for the experimental findings.

2 Concepts and Mechanisms of OT

2.1 The OT Framework

Grimshaw (1997) lists the following as the basic assumptions underlying the OT framework:

(1) *Basic Assumptions of Optimality Theory*
- a. Constraints are universal.
- b. Constraints can be violated.
- c. Grammars are rankings of constraints.

M. Catherine Gruber, Derrick Higgins, Kenneth S. Olson, and Tamra Wysocki (eds.).
Papers from the 34th Meeting of the Chicago Linguistic Society. Vol. 2: The Panels,
95–109. Chicago, 1998.

d. The optimal form is grammatical; all nonoptimal candidates are ungrammatical. (Grimshaw 1997: 373)

OT specifies a generation function *Gen* which generates a set of candidate structures (the reference set) for a given input representation. In Grimshaw's (1997) account, the input representation is the numeration that is shared by all candidate structures. An output structure is assigned to the input I as the result of an optimization process over the candidate structures for I. More precisely, the output S_{opt} for an input I is the optimal structure in the reference set $R = Gen(I)$. Grimshaw (1997) gives this definition of optimality:

(2) *Optimality*
An optimal output form for a given input is selected from among the class of competitors in the following way: a form that, for every pairwise competition involving it, best satisfies the highest-ranking constraint on which the competitors conflict, is *optimal*. (Grimshaw 1997: 373)

An OT grammar for a given language L has to be constructed such that, for every input I, the output structure S_{opt} in $Gen(I)$ is the grammatical realization of I in L. To achieve this, an OT grammar specifies a set of universal grammatical constraints along with a set of language-specific constraint rankings. Note that OT differs from more traditional grammar frameworks in that the grammaticality of a structure is not determined by its inherent properties, but by the set of structures it competes with.

2.2 Testable Claims

In the following, we examine the assumptions in (1) and (2) and identify the ones that can be verified experimentally.

Universality of constraints: OT assumes that constraints are universal, both in terms of being highly general in their formulation, and in terms of being valid across languages. This assumption is central to OT's approach to crosslinguistic variation: constraints state generalizations across languages, while constraint rankings state language-specific facts.

It follows that (1a) has to be regarded as a prescriptive statement (it states how constraints in OT have to be formulated), not as an empirical one. It is therefore not amenable to empirical verification, and will be disregarded in the present study.

Violability of constraints: as constraints in OT are formulated in a highly general fashion, they will inherently conflict, i.e., the satisfaction of a given constraint will lead to the violation of another. As a consequence, constraints in OT cannot be absolute, but have to be violable. This entails that even a grammatical structure may incur constraint violations.

Again, this is not an assumption that can be empirically challenged: (1b) is true by definition; it is a stipulation that has to be made once we have adopted (1a), the universality of constraints.

Constraint ranking: constraints in OT are inherently conflicting and may be violated even in a grammatical structure. This means that a mechanism for resolving constraint conflicts is necessary to differentiate between grammatical and ungrammatical structures. OT's assumption is that constraints are ranked (assumption (1c)): the violation of a higher ranking constraint is more serious than the violation of a lower ranking one.

(1c) is not a necessary assumption: other mechanisms apart from constraint hierarchies could be used to resolve constraint conflicts. It is therefore legitimate to ask whether there is any empirical support for constraint hierarchies. For instance, we can investigate if the fact that some constraint violations are more serious than others is reflected in behavioral data, i.e., in data on how humans produce, process, or judge linguistic structures. In the present study, we focus on evidence from linguistic judgments.

Optimal structures are grammatical: a grammar has to distinguish grammatical structures (forms) from ungrammatical ones. Conventional linguistic frameworks assume a trivial mechanism for achieving this: a structure is grammatical if it does not violate any constraints. As constraints in OT are universal, violable, and ranked, a more complicated mechanism is required: OT's assumption is that a structure is grammatical if it is optimal in that it incurs the least serious constraint violations compared to a set of competitors.

The assumption that optimal structures are grammatical is again one that is true by definition. We have to assume something like (1d) if we stipulate that constraints can be violated, even in grammatical structures. It follows that (1d) cannot be subject to empirical investigation, but rather is a precondition for it.

Constraint interaction: assumption (2) specifies how constraints rankings are assessed in the computation of optimal structures. It asserts the principle of *strict domination*, which states that the highest ranking constraint on which two structures conflict is crucial for deciding which of the structures is more optimal. Strict domination entails that the violation of a constraint C cannot be compensated by any number of violation of constraints that are lower ranking than C.

Many other forms of constraint interaction are conceivable as alternatives to strict domination,[1] hence it seems worthwhile to test the plausibility of this assumption against experimental findings. Again, the present studies attempts this by using evidence from grammaticality judgment experiments.

2.3 Suboptimal Candidates

This study uses evidence from grammaticality judgments to investigate the plausibility of OT's assumptions about constraint ranking (definition (1c)) and constraint interaction (definition (2)). Data on suboptimal candidates and the associated gradient grammaticality judgments play a particularly crucial role here.

Note that the standard case in the (non-experimental) OT literature is to investigate optimal candidates, rather than suboptimal ones. More specifically, a linguist working in the OT framework determines relative constraint rankings by comparing an optimal candidate S_{opt} and a suboptimal candidate S_{sub}. Such a pair of candidates is informative if S_{opt} and S_{sub} have the same violation profile, with the exceptions of the constraints C_{sub}, which is only violated by S_{sub}, and C_{opt}, which is only violated by S_{opt}. Under the assumption that S_{opt} is grammatical, while S_{sub} is ungrammatical, it follows that the correct constraint ranking is $C_{sub} \gg C_{opt}$.

In such a setting, the suboptimal candidate in itself is not of interest, it plays a role only in relation to the corresponding optimal candidate. By testing informative pairs of optimal and suboptimal candidates, a linguist can infer a set of constraint rankings for a given phenomenon. Naturally, this approach presupposes the specific mechanisms in (1) and (2) that determine the optimal candidate in OT. These mechanisms can be tested in an indirect fashion by developing plausible OT analyses for a set of linguistic phenomena (strategy (a), section 1).

The aim of the present study, in contrast, is to investigate OT's concepts and mechanisms directly by testing them against experimental evidence (strategy (c), section 1). A way of achieving this is by investigating suboptimal candidates, rather than optimal ones. By definition, an optimal candidate is grammatical in OT, and hence the effects of the constraint violations it incurs are not directly observable. In suboptimal candidates, on the other hand, at least some violations have an observable effect, as they lead to ungrammaticality. Now the crucial assumption is that we can learn something about the constraints violated by a suboptimal candidate by investigating the degree of ungrammaticality that it exhibits. The present study uses this assumption as the key to an investigation of constraint ranking and constraint interaction in OT.

Note that this approach makes use of a different sort of informative pair: we compare two candidates S and S' (both may be suboptimal) which have the same violation profile, apart from the constraint C, which is violated only by S. Now we try to learn about the nature of the violation of C by determining how the grammaticality of S differs from the grammaticality of S'. This method allows us to investigate constraint ranking (by varying C) and constraint interaction (by determining the effect of multiple violations).

2.4 Empirical Issues

In section 3, we will review a set of experimental sstudies that investigate the relative acceptability of suboptimal linguistic structures, thus providing data that bear on the OT notions of constraint ranking and constraint interaction:

- **Constraint ranking:** is there evidence that some constraint violations lead to a higher degree of ungrammaticality than others?
- **Constraint Interaction:**
  - Are constraint violations cumulative, i.e., does the degree of unacceptability increase with the number of constraints violated?
  - How does constraint ranking interact with the cumulativity of constraints? If a constraint C outranks C', does this mean that even multiple violations of C' cannot compensate for a single violation of C (strict domination)?

These questions can be studied by measuring the degree of unacceptability caused by different types of constraint violations, and by determining the effect of multiple constraint violations.

3 Experimental Evidence

3.1 Magnitude Estimation

The approach pursued in this study uses gradient grammaticality judgment data as a means of testing assumptions from linguistic theory. Our interest in gradient judgments implies that data need to be elicited experimentally, since the informal elicitation technique traditionally used in linguistics is unlikely to yield reliable data for degrees of grammaticality (Schütze 1996). A suitable experimental paradigm is magnitude estimation (ME), a technique standardly applied in psychophysics to measure judgements of sensory stimuli (Stevens 1975). The ME procedure requires subjects to estimate the magnitude of physical stimuli by assigning numerical values proportional to the stimulus magnitude they perceive. Highly reliable judgements can be achieved for a whole range of sensory modalities, such as brightness, loudness, or tactile stimulation.

The ME paradigm has been extended successfully to the psychosocial domain (Lodge 1981) and recently Bard et al. (1996) and Cowart (1997) showed that linguistic judgements can be elicited in the same way as judgements of sensory or social stimuli. In contrast to the 5- or 7-point scale conventionally used to measure human intuitions, ME employs a continuous numerical scale, thus providing fine-grained measurements of linguistic acceptability, which are robust enough to yield statistically significant results, while being highly replicable both within and across speakers.

ME requires subjects to assign numbers to a series of linguistic stimuli proportional to the acceptability they perceive. First, subjects are exposed to a modulus item, which they assign an arbitrary number. Then, all other stimuli are rated proportional to the modulus, i.e., if a sentence is three times as acceptable as the modulus, it gets three times the modulus number, etc.

3.2 Gradient Judgments Data

Keller (1996a,b) presents the results of an ME study investigating gradience in extraction from picture NPs. This study identifies definiteness as a factor that influences acceptability: for examples like (3), the experimental data indicate that extraction from indefinite NPs is more acceptable than extraction from definite NPs (a similar effect is reported by Cowart (1997: ch. 1) and Neville et al. (1991)):

(3) a. Who did you see *a* picture of?
b. Who did you see *the* picture of?

Acceptability also depends on the semantics of the matrix verb. Aspectual class seem to be a main factor here: state verbs are more acceptable than activity verb (example 4a)), while for achievements and accomplishments, a verb of creation is more acceptable than a verb of destruction (example (4b,c)). Keller (1996a,b) reports significant acceptability differences for all pairs in (4):

(4) a. Who did you *have/analyze* a picture of?
b. Who did you *take/destroy* a picture of?
c. Who did you *find/lose* a picture of?

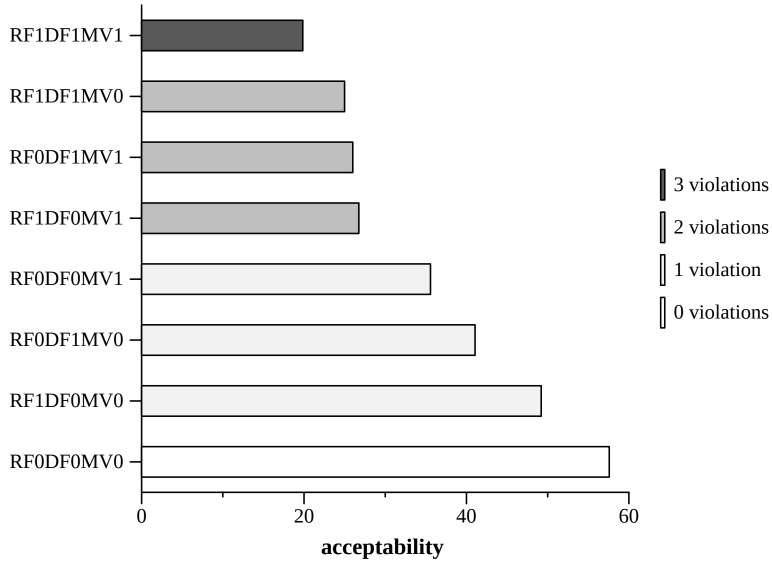

Figure 1: Cumulativity effect for violations of constraints on extraction from picture NPs

The third significant factor is the referentiality of the extracted NP. Here, the experimental data reveals an acceptability hierarchy, with decreasing acceptability from (5a) to (5d):

(5) a. *Who* did you take a picture of?
b. *Which man* did you take a picture of?
c. *What* did you take a picture of?
d. *How many men* did you take a picture of?

However, only one point on this hierarchy (*how many N*) is significantly different from the others. In this sense, this experiments fails to yield firm evidence that some violations of the referentiality constraint are more serious than others.

To investigate constraint ranking and interaction, we reanalyzed the data in Keller 1996a,b data and checked for an effect of multiple violations. Figure 1 shows the relative acceptability for the factors RF (referentiality), DF (definiteness), and MV (main verb),[2] revealing a cumulativity effect: a single constraint violation is significantly more acceptable than a double violation, which in turn is significantly more acceptable than a triple violation. Note also that single violations of different types of constraints cause varying degrees of grammaticality, but these are only non-significant tendencies.[3]

Keller (1997b) reports further experiments to confirm the cumulativity effect for a different domain. He tested violations of phrase structure (PS), number agreement (NA), and subcategorization (SC) constraints. The study was conducted with native speakers of German, using stimuli such as the following:

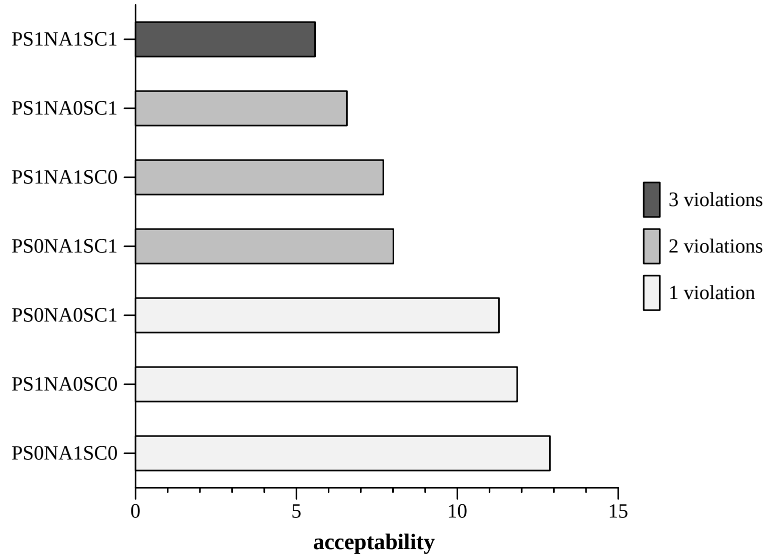

Figure 2: Cumulativity effect for violations of constraints on phrase structure, agreement, and subcategorization

(6) a. Der Mann liest den sehr lustigen Brief von seiner Frau.

Der	Mann	liest	den	sehr	lustigen	Brief	von	seiner	Frau.
the	man	reads	the	very	funny	letter	by	his	wife.

b. Der Mann liest den lustigen sehr Brief von seiner Frau.

Der	Mann	liest	den	lustigen	sehr	Brief	von	seiner	Frau.
the	man	reads	the	funny	very	letter	by	his	wife.

c. Der Mann lesen den sehr lustigen Brief von seiner Frau.

Der	Mann	lesen	den	sehr	lustigen	Brief	von	seiner	Frau.
the	man	read	the	very	funny	letter	by	his	wife.

d. Der Mann lacht den sehr lustigen Brief von seiner Frau.

Der	Mann	lacht	den	sehr	lustigen	Brief	von	seiner	Frau.
the	man	laughs	the	very	funny	letter	by	his	wife.

Each experimental item contained zero to three constraint violations, yielding the mean ME ratings in figure 2. These results confirm the cumulativity effect for a different set of constraints. Note that there was no significant differences between different types of constraint violations; only the number of violations yielded a significant effect.

Keller (1997b) also reports a validation study that replicated the cumulativity effect for English. This study included violations of constraints on NP extraction, as well as violations of PS, NA, and SC. The results (figure 3) suggest that two types of constraints can be distinguished: hard constraints, that cause strong ungrammaticality when violated (PS, NA, SC), and soft constraints causing only mild ungrammaticality (RF, DF): violations of hard constraints are significantly less acceptable than violations of soft constraints. Note that this holds even for multiple soft violations; the effect from constraint type (hard or soft) seems to outweigh the cumulativity effect.

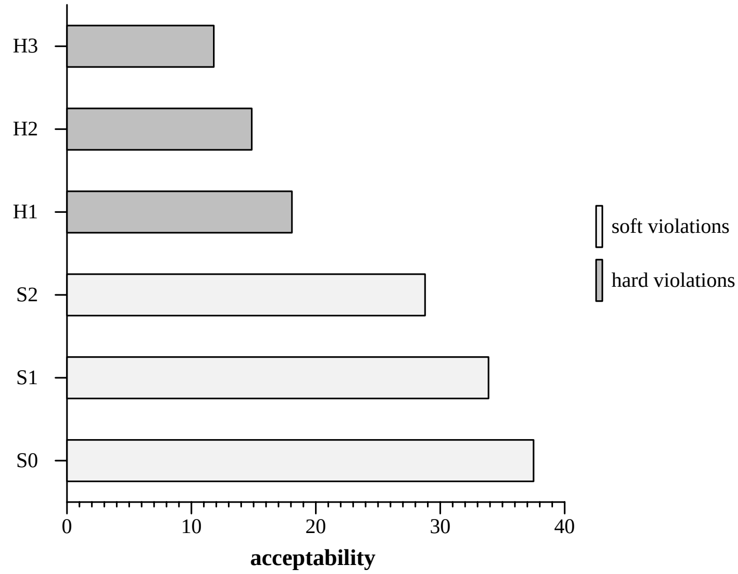

Figure 3: Interaction of violations of soft and hard constraints

3.3 Empirical Results

In the last section, we reviewed a set of experimental findings that bear on the OT notions of constraint ranking and constraint interaction. The results can be summarized as follows:

- **Constraint ranking:** two types of constraints can be distinguished: soft and hard ones. Violations of soft constraints are judged as significantly less serious than violations of hard constraints. This supports a limited form of constraint ranking: soft constraints are ranked higher than hard ones. However, our experiments failed to find significant differences in grammaticality for violations of different types of soft and hard constraints.

- **Constraint Interaction:**

  - There is evidence that constraint violations are cumulative, i.e., that the degree of ungrammaticality increases with the number of constraints violated. This holds both for soft and for hard constraints.
  - The effect from constraint type (hard or soft) is stronger than the cumulativity effect. Even a single hard violation is judged less grammatical than multiple soft violations. This indicates that the interaction of soft and hard constraints conforms to OT's notion of strict domination.

Taken together, the experimental findings reviewed in this section lend some initial plausibility to the grammatical notions stipulated in OT. Based on this result, the next section develops a framework for gradient grammaticality based on concepts from OT. This framework, which explains gradience in terms of selective constraint re-ranking, allows to model gradient data directly, instead of merely being compatible with it (as is the case for standard OT).

4 Modeling Gradience in OT

4.1 A Naive Model

A straightforward model of gradient grammaticality can be obtained by dropping OT's assumption that all suboptimal structures are (equally) ungrammatical (definition (1d)). Instead, we can stipulate that the relative optimality of a structure corresponds to its relative grammaticality: a candidate structure is more grammatical than a given competitor if it is more optimal than the competitor. (Such a model was actually proposed in Keller 1996a, 1997a.)

This approach retains OT's definition of optimality in (2) and uses it to predict a grammaticality ranking over a candidate set, which can then be tested against gradient linguistic judgments. In such a model, constraint rankings can be determined directly by comparing the relative grammaticality of suboptimal candidates, an option that is not available in standard OT (section 2.3). Hence the use of gradient data increases the range of evidence that can be expoited when developing OT analyses.

However, the naive model of gradient grammaticality encounters a number of serious problems when faced with experimental evidence such as the one presented in section 3. One problem is that it predicts grammaticality differences *only* for structures in the same candidate set; relative grammaticality cannot be compared across candidate sets. The experimental findings, however, show that subjects can judge the relative grammaticality of arbitrary sentence pairs, a fact that cannot be accommodated by the naive model.

Another problem is that grammaticality differences are predicted between *all* structures in a candidate set. A typical OT grammar assumes a richly structured constraint hierarchy, therefore all or most structures in a given candidate set will differ in optimality. The naive model predicts that there is a grammaticality difference whenever there is a difference in optimality. The experimental results, however, indicate that the scope of gradience in the grammar much more limited: differences in acceptability were found between violations of soft and hard constraints only; different types of soft and hard constraint violations did not differ in relative acceptability. This means that the naive model seriously overgenerates, in the sense that it predicts many more grammaticality differences than are justified by the experimental data.

The cumulativity of constraint violations poses a third problem for the naive model of gradience. The experimental results demonstrate that the degree of ungrammaticality of a structure increases with the number of constraints it violates, both for soft and hard constraints. This fact is not accounted for by the naive model: it relies on the standard OT notion of optimality, which is defined via strict domination and predicts cumulativity effects only for constraints with the same ranking. The experimental results provide only limited support for strict domination: violations of hard constraints seem to be strictly worse than (even multiple) violations of soft constraints. In the naive model, the only way of accounting for this is by stipulating a grammar with only one ranking, viz., between soft and hard constraints, which of course is highly implausible.[4]

4.2 Learnability in OT

In the next section, we propose an alternative model of gradience based on mechanisms from OT learnability theory. OT captures crosslinguistic variation as variation in the constraint hierarchy. Therefore, grammar learning amounts to determining a language specific hierarchy over a set of universal grammatical constraints.

Tesar and Smolensky (1998) propose a learning algorithm that achieves this task efficiently based on positive evidence only. This algorithm is error-driven: when exposed to an example sentence, the learner determines whether this sentence is optimal (and thus grammatical) according to his or her grammar. If the example sentence fails to be optimal, the learner modifies his or her constraint hierarchy so that the example becomes optimal. By iterating this over a sufficient number of example sentences, the learner's grammar is gradually modified to match the adult's grammar by which the examples were generated.

The following is a simplified version of Tesar and Smolensky's (1998) learning algorithm:

(7) *Grammar Learning in Optimality Theory*
To learn an OT grammar, carry out the following steps for each each example sentence E you are exposed to:

a. *Robust Interpretative Parsing:* parse E, i.e., determine its structure S and the underlying input I.
b. *Implicit Negative Evidence:* compute $R = Gen(I)$, the set of competitors of S.
c. *Constraint Demotion:* for each candidate S' in R, and each constraint C violated by S but not by S', demote C to immediately below the highest ranking constraint that is violated by S' but not by S. This results in a constraint hierarchy that makes S the optimal candidate in R.

This algorithm relies on the assumption that parsing is robust, i.e., that the learner can determine the structure S of an example E even if S is not optimal in his or her current grammar (step (7a)). Robust parsing gives the learner access to implicit negative evidence in the form of the competitors of S (step (7b)). Based on the competitor set, the learner computes the re-rankings necessary to make S optimal in his or her grammar (step (7b)). By performing these re-rankings, the learner adapts his or her grammar to the target grammar, i.e., the adult grammar that generated E. (If S is already optimal in the learner's grammar, then no learning is possible—E is an uninformative example.)

Tesar and Smolensky (1998) proof the correctness of this algorithm and show that it is of quadratic data complexity: only $(n-1)n$ informative examples are needed to learn the ranking of a grammar with n constraints. (Note that a grammar with n constraints allows $n!$ possible rankings.)

4.3 A Re-ranking Model

An adequate model of gradience should be able to accommodate the experimental facts reported in section 3, such as the cumulativity effect and the distinction between soft and hard constraints. A naive model that just equates grammaticality and optimality fails to achieve this, as discussed in section 4.1. A more plausible account of gradient grammaticality can be devised based on the concept constraint re-ranking, borrowed from OT learnability theory.

Consider the following algorithm, which is closely related to the OT learning algorithm in (7):

(8) *Gradient Grammaticality in Optimality Theory*

To compute the degree of grammaticality of a sentence E, carry out the following steps:

a. *Robust Interpretative Parsing:* parse E, i.e., determine its structure S and the underlying input I.

b. *Implicit Negative Evidence:* compute $R = Gen(I)$, the set of competitors of S.

c. *Selective Constraint Re-ranking:* use the constraint demotion algorithm (7c) to determine which re-rankings are required to make S the optimal candidate in R.

This algorithm handles ungrammatical input via robust parsing, in the same way as the learning algorithm in (7). It computes which (hypothetical) constraint re-rankings are required to make a given structure optimal. This information can then be used to compare structures with respect to their degree of grammaticality. (Note that the re-rankings are not actually carried out—we are dealing with an adult grammar, which can be assumed to be stable and should not be modified in the face of ungrammatical data.)

To provide an account of gradient grammaticality, we can now postulate that the degree of grammaticality of a structure S depends on the number and type of re-rankings required to make S optimal, as computed by step (8c). Such a re-ranking model offers the necessary flexibility to accommodate the experimental findings on constraint ranking and constraint interaction in OT:

- The re-ranking model allows to determine the relative grammaticality of arbitrary structures by comparing the number and type of re-rankings required to make them optimal. Comparisons of grammaticality are not confined to structures in the same candidate set, which accounts for the fact that subjects can judge the relative grammaticality of arbitrary sentence pairs.

- It seems plausible to assume that some constraint re-rankings are more serious than others, and hence cause a higher degree of ungrammaticality in the target structure. This assumption allows to model the experimental findings that some constraint violations lead to a higher degree of ungrammaticality than others. The experimental data justify two types of of re-rankings, corresponding to the soft and hard constraint violations discussed in section 3.

- Another assumption is that the degree of grammaticality of a structure depends on the number of re-rankings necessary to make it optimal: the more re-rankings a structure requires, the more ungrammatical it becomes. This predicts the cumulativity of violations that was found experimentally both for soft and for hard constraints.

5 Discussion

5.1 Alternative Approaches

One alternative model of gradience, the naive model, was already discussed in section 4.1. Müller (1998) develops a variant of the naive model that confines gradience to constraint subhierarchies. While this approach avoids the prediction of bogus grammaticality distinctions, it fails to account for the cumulativity of constraint violation. Furthermore, Müller's (1998) account is based purely on intuitive judgments, which is likely to be an unreliable way of obtaining gradient data (Bard et al. 1996; Cowart 1997; Schütze 1996).

Hayes (1999) puts forward another approach to gradience in OT based on the concept of strictness bands. In this framework, a constraint may carry a strictness bands that specifies possible variation in the ranking of the constraint, and the loss of grammaticality that is associated with this variation.[5] Hayes (1999) provides an algorithm that assigns a degree of grammaticality to a given structure based on which strictness bands (re-rankings) were involved in generating it.

The distinction between hard and soft constraints can be accounted for straightforwardly in Hayes's (1999) model: soft constraints carry strictness bands, while hard ones do not. Thus a soft violation will lead to an intermediate degree of grammaticality as specified by its strictness band, while a hard violation will result in maximal ungrammaticality (no re-ranking is possible in the absence of a strictness band), However, the cumulativity of constraint violations cannot readily be accommodated by Hayes's (1999) approach. We could extend his approach so that the computation of the grammaticality of a given structure takes into account how many re-rankings are required for generating it. This, however, captures the cumulativity of soft constraint violations only; hard violations do not lead to re-rankings, and hence their cumulativity remains unexplained.

5.2 Open Questions

The re-ranking model offers a general way of dealing with degrees of grammaticality in OT, based on concepts that are independently motivated in OT learnability theory. However, a number of open questions remain.

An obvious problem concerns the cumulativity effect: if we assume that the degree of grammaticality of a given structure depends on the number of re-rankings it requires, then this naturally predicts that constraint violations are cumulative. However, this only holds for multiple violations of different constraints (requiring different re-rankings that are counted separately). Multiple violations of the same constraint, however, can be dealt with by a single re-ranking, and hence we fail to predict a cumulativity effect here. This does not seem to be in accordance with the experimental facts: Chapman (1974) found a cumulativity effect for multiple violations of the same constraint (note 3).

Another problem concerns the case of unmarked competitors. Algorithm (8) demotes the constraints violated by S below the ones violated by a given competitor S', so that S becomes optimal (step (7c)). The degree of grammaticality of S depends the type and number of re-rankings required in this demotion process. Constraint demontions is impossible, however, if the competitor S' is completely unmarked, i.e., if it incurs no constraint violations at all.

5.3 Future Directions

On the positive side, a re-ranking model provides a natural way of accounting for free variation, viz., by complementing the existing types of re-rankings (hard and soft) with a third type: free re-rankings, which we assume not to affect grammaticality. A structure that requires only free re-rankings to become optimal is fully grammatical, and hence is an optional realization of the input. (In other words, all candidate structures that require no or only free re-rankings are in free variation with each other).

Another interesting point is that the concept of constraint re-ranking can also be used to model the relative frequency of forms that are in free variation; serveral relevant models have been put forward in the OT literature (Anttila 1997; Boersma 1997; Hahn 1998; Hayes and MacEachern 1998). It seems highly desirable to provide a uniform account of gradience and frequency, two related forms of numeric information in the grammar.[6] We leave it to future research to determine how frequency information can be integrate into the re-ranking model proposed in the present paper.

6 Conclusions

This paper used experimental data on gradient grammaticality for a double purpose: firstly, to lend plausibility to the concepts assumed in OT, and secondly, to devise an explicit model of gradience based on these concepts. Our main claims can be summarized as follows:

- OT's assumptions about constraint ranking and constraint interaction can be validated by testing their compatibility with experimental data. Experimental data on gradient judgments are particularly suitable for such a validation, as they allow to investigate the relative grammaticality of suboptimal linguistic structures.

- Experimental findings show that two types of constraints can be distinguished: soft and hard ones. Violations of soft constraints are judged as significantly less serious than violations of hard constraints.

- There is evidence that constraint violations are cumulative, i.e., that the degree of ungrammaticality increases with the number of constraints violated. Also, the effect from constraint type (hard or soft) can be shown to outweigh the cumulativity effect.

- These findings provide limited support for OT's notions of constraint ranking and constraint interaction, including the concept of strict domination.
- OT can be extended to provide an explicit account of gradient judgment data (instead of just being compatible with it). However, a naive model of gradience that equates relative grammaticality with relative optimality is not compatible with the experimental findings.
- A plausible model of gradience can be devised based on the concept of selective constraint re-ranking. This approach, which is grounded in OT learnability theory, predicts the cumulativity of constraint violations, and allows to model the distinction between hard and soft constraints found experimentally.

Notes

[1]A simple alternative approach to constraint interaction would be the summation of violations: the structure that incurs the least number of violations is optimal. Such an approach would not necessarily require the ranking of constraints; it could just count the number of violations of each candidate, no matter how serious they are.

[2]Only two levels per factor were included: RF0 (*who*) and RF1 (*how many N*), MV0 (state verb) and MV1 (activity verb), DF0 (*a picture*) and DF1 (*the picture*).

[3]A similar result was obtained by Chapman (1974), who investigated two types of violations (selectional restrictions and subcategorization requirements). He found an effect of number of violations, but not of type of violation. Chapman's (1974) study differs from the ones reported here in that he investigated multiple violations of the same constraint.

[4]This is particularly implausible as the fact that soft constraints are ranked higher than hard constraints entails that this ranking is subject to crosslinguistic variation. Intuitively, however, we would expect hard constraint (such as agreement) to be hard in all languages. (The case is less clear for soft constraints, whose status might well vary across languages.)

[5]Strictness bands seem a rather ad hoc concept at first, but it might be possible to motivated them with results from learnability theory. The alternative OT learning algorithm developed by Boersma (1997, 1999), for instance, assumes a stochastic constraint evaluation mechanism that is conceptually similar to Hayes's (1999) strictness bands.

[6]Frequency and gradience are related, but separate grammatical concepts, as argued extensively by Keller (1996a).

References

Anttila, Arto. 1997. Deriving Variation from Grammar: A Study of Finnish Genitives. Variation, Change, and Phonological Theory, edited by Frans Hinskens, Roeland van Hout, and W. Leo Wetzels. Amsterdam: John Benjamins.

Bard, Ellen Gurman, Dan Robertson, and Antonella Sorace. 1996. Magnitude Estimation of Linguistic Acceptability. Language 72:32–68.

Boersma, Paul. 1997. How we Learn Variation, Optionality, and Probability. Proceedings of the Institute of Phonetic Sciences, vol. 21, 43–58. University of Amsterdam.

—. 1999. Learning a Grammar in Functional Phonology. Dekkers et al. 1999.

Chapman, Robin S. 1974. The Interpretation of Deviant Sentences in English: A Transformational Approach. The Hague: Mouton.

Cowart, Wayne. 1997. Experimental Syntax: Applying Objective Methods to Sentence Judgments. Thousand Oaks, CA: Sage Publications.

Dekkers, Joost, Frank van der Leeuw, and Jeroen van de Weijer, eds. 1999. Optimality Theory: Phonology, Syntax, and Acquisition. Oxford: Clarendon Press.

Ellison, T. Mark. 1994. Constraints, exceptions and representations. Proceedings of the 1st Meeting of the ACL Special Interest Group in Computational Phonology, 25–32. Las Cruces, NM.

Grimshaw, Jane. 1997. Projection, Heads, and Optimality. Linguistic Inquiry 28:373–422.

Hahn, Axel. 1998. Variation, Grammar, and the Power of the Optimal: German Obstruent Devoicing. Unpubl. ms., Heinz Nixdorf Institute, University of Paderborn.

Hayes, Bruce P. 1999. Gradient Well-Formedness in Optimality Theory. Dekkers et al. 1999.

Hayes, Bruce P., and Margaret MacEachern. 1998. Folk Verse Form in English. Language 74:473–507.

Karttunen, Lauri. 1998. The Proper Treatment of Optimality in Computational Phonology. Prodeedings of the International Workshop on Finite-state Methods in Natural Language Processing. Ankara.

Keller, Frank. 1996a. Extraction from Complex Noun Phrases: A Case Study in Graded Grammaticality. Master's thesis, University of Stuttgart.

—. 1996b. How Do Humans Deal with Ungrammatical Input? Experimental Evidence and Computational Modelling. Natural Language Processing and Speech Technology: Results of the 3rd KONVENS Conference, Bielefeld, October 1996, edited by Dafydd Gibbon, 27–34. Berlin: Mouton de Gruyter.

—. 1997a. Extraction, Gradedness, and Optimality. Proceedings of the 21st Annual Penn Linguistics Colloquium, edited by Alexis Dimitriadis, Laura Siegel, Clarissa Surek-Clark, and Alexander Williams, no. 4.2 in Penn Working Papers in Linguistics, 169–186. Department of Linguistics, University of Pennsylvania.

—. 1997b. Grammaticality Effects from Cumulative Constraint Violations. Paper presented at the 3rd Conference on Architectures and Mechanisms for Language Processing, Edinburgh.

Lodge, Milton. 1981. Magnitude Scaling: Quantitative Measurement of Opinions. Beverley Hills, CA: Sage Publications.

Müller, Gereon. 1998. German Word Order and Optimality Theory. Working Papers of the SFB 340, Report 126, University of Tübingen.

Neville, Helen, Janet L. Nicol, Andrew Barss, Kenneth I. Forster, and Merrill F. Garrett. 1991. Syntactically Based Sentence Processing Classes: Evidence from Event-Related Brain Potentials. Journal of Cognitive Neuroscience 3:151–165.

Prince, Allan, and Paul Smolensky. 1993. Optimality Theory: Constraint Interaction in Generative Grammar. Technical Report 2, Center for Cognitive Science, Rutgers University.

Schütze, Carson T. 1996. The Empirical Base of Linguistics: Grammaticality Judgments and Linguistic Methodology. Chicago: University of Chicago Press.

Stevens, Stanley S., ed. 1975. Psychophysics: Introduction to its Perceptual, Neural, and Social Prospects. New York: John Wiley.

Tesar, Bruce, and Paul Smolensky. 1998. Learnability in Optimality Theory. Linguistic Inquiry 29:229–268.