Gradience in Linguistic Data

Antonella Sorace

Theoretical and Applied Linguistics
School of Philosphy, Psychology and Language Sciences
University of Edinburgh
40 George Square, Edinburgh EH8 9LL, UK
phone: +44-131-650-3493, fax: +44-131-650-3962

Frank Keller

Institute for Communicating and Collaborative Systems
School of Informatics
University of Edinburgh
2 Buccleuch Place, Edinburgh EH8 9LW, UK
phone: +44-131-650-4407, fax: +44-131-650-6626

Abstract

This paper provides a survey of the theoretical and experimental findings on degrees of grammaticality, with a special focus on gradience in syntax. We first discuss the theoretical relevance of gradient data, and argue that such data should be elicited experimentally in order to be reliable. We then review a set of experimental findings on gradience, which lead to the hypothesis that linguistic constraints come in two types: hard constraints whose violations trigger strong unacceptability, and soft constraints that lead to only mild unacceptability. There is experimental evidence that both types of constraints are subject to constraint ranking and show cumulativity and ganging up effects. However, soft and hard constraints differ with respect to context effects, crosslinguistic variation, and developmental optionality. This set of properties forms the basis for an evaluation of existing proposals for models of gradience in grammar, with a particular focus on models that extend Optimality Theory.

Key words: linguistic data, gradience, grammaticality, magnitude estimation, Optimality Theory

Email addresses: `antonella@ling.ed.ac.uk` (Antonella Sorace), `keller@inf.ed.ac.uk` (Frank Keller).

1   Introduction

Acceptability judgments are the basic data that linguists rely on to formulate their theories. In certain cases, these data fail to provide a clear-cut division between fully acceptable sentences and fully unacceptable sentences. Rather, relevant linguistic examples are *gradient*, i.e., they come in varying degrees of acceptability.

The aim of this paper is to present a unified view of gradient data in syntax, based on recent experimental and theoretical results. In the current section, we give an overview of the general theoretical and empirical issues related to gradient linguistic data. Section 2 reviews the experimental findings relating to gradience and constraint ranking, cumulativity, context effects, and crosslinguistic effects. Based on these findings, a fundamental distinction between hard and soft constraints is proposed. Section 3 presents developmental evidence (from language acquisition and attrition) that supports the hard/soft distinction. Section 4 discusses the implications of the experimental and developmental findings for a number of models of gradience that have been proposed in the literature.

1.1   Theoretical Relevance of Gradient Data

Our guiding assumption is that gradient data can contribute to linguistic theory as they provide more information than the binary judgments on which linguists traditionally rely. Often these judgments are not in fact binary, but constitute an idealization, namely an artificial classification into acceptable and unacceptable examples. In what follows, we argue that it is preferable to give up this idealization and develop a theory that permits us to analyze realistic, gradient data.

The potential benefits of a theory of gradient grammaticality include an expansion of the empirical base of linguistics and an increase of the predictive power of linguistic theory. As Hayes (1997, 15) puts it: "Linguistics at present is *not hard enough*; we are not presenting our theories with sufficient demands to distinguish which ones are true. The task of analyzing data with gradient well-formedness puts a theory to a stiffer test." Note that accounting for gradience was part of the research program of early generative grammar. Chomsky, for instance, insists that "an adequate linguistic theory will have to recognize degrees of grammaticalness" (Chomsky, 1975, 131), based on the observation that "there is little doubt that speakers can fairly consistently order new utterances, never previously heard, with respect to their degree of 'belongingness' to the language" (Chomsky, 1975, 132).

However, the issue of gradience has become increasingly difficult to accommodate in contemporary formal models of grammar. While the Principle and Parameters model left some limited scope for gradience and optionality (Fukui, 1993), more recent models (e.g., Minimalism, Optimality Theory) assume a set of candidates

competing for well-formedness and an evaluation metric, based on economy principles, which selects only one candidate as the winner (Müller, 1999b). Gradience, which presupposes the recognition that suboptimal candidates have grammatical status, is therefore incompatible with these models.

On the other hand, a number of experimental studies demonstrate that by taking gradient judgment data into account, one can both discover new linguistic facts that have eluded the conventional, informal approach to data collection, and resolve data disputes that exist for certain linguistic phenomena in the literature. The underlying hypothesis is that such disputes arise because conventional linguistic analysis fails to do justice to the gradient nature of these phenomena, both in its data collection methodology and in its analytic approach.

Note that there is an important methodological caveat here. Arguably, the aim of formulating precise, testable theories of linguistic competence is at the heart of the generative enterprise. We have to make sure that this aim carries over to an extended theoretical framework that is capable of dealing with gradience. In other words, we have to make sure that a *formal* theory of gradience is possible, countering "[c]ritics of generative grammar [who] might take the existence of gradient well-formedness judgments as an indication that the entire enterprise is misconceived [. . .]. In this eliminativist view, gradient well-formedness judgments constitute evidence that generative linguistics must be replaced by something very different, something much 'fuzzier'" (Hayes, 2000, 88). We follow Hayes (1997, 15) in adopting the guiding assumption that "we don't have to trash existing theories of what constraints are like just to get gradient well-formedness". The challenge is to develop a grammatical framework that is permissive enough to account for gradient data without idealizing it, but restrictive enough to allow us to formulate precise, testable linguistic analyses. Section 4 discusses this issue in some detail, focusing in particular on the question of how existing linguistic frameworks such as Optimality Theory (Prince and Smolensky, 1993) can be extended to deal with gradient data.

1.2   Eliciting Gradient Judgments

Acceptability judgments by native speakers are generally accepted as the main type of evidence for linguistic theory. The use of judgment data is typically justified by a set of key arguments that Schütze (1996, 2) summarizes as follows:

- Acceptability judgments allow us to examine sentences that rarely occur in spontaneous speech or corpora.
- Judgments constitute a way of obtaining negative evidence, which is rare in normal language use.
- In observing naturally occurring speech data, it is difficult to distinguish errors

(slips of the tongue, unfinished utterances, etc.) from grammatical production.

- The use of acceptability judgments allows us to minimize the influence of communicative and representational functions of language. Judgment data allow us to study the structural properties of language in isolation.

This set of advantages explains the popularity of acceptability judgments as primary data for linguistic theory. However, as Schütze (1996) argues, judgment data are often used by linguists in a dangerously uncritical fashion. "In the vast majority of cases in linguistics, there is not the slightest attempt to impose any of the standard experimental control techniques, such as random sampling of subjects and stimulus materials or counterbalancing for order effects" (Schütze, 1996, 4). Linguists typically rely on a naive, intuitive way of collecting judgments, ignoring psycholinguistic findings that show that acceptability judgments are subject to a considerable number of biases, for which a naive methodology fails to control.

Schütze (1996) also points out that current linguistic research makes crucial use of subtle (and thus potentially controversial) judgments; it does not confine itself to cases of clear acceptability or unacceptability (which arguably can be established without using an experimental methodology). To substantiate this claim, he discusses the use of subtle judgments in the widely cited studies by Aoun et al. (1987), Belletti and Rizzi (1988), and Lasnik and Saito (1984).

Belletti and Rizzi's 1988 study is particularly interesting as it makes extensive use of gradient acceptability judgments, de facto employing a seven point scale for acceptability. However, the authors fail to provide an explicit account of degrees of grammaticality:

> But there is no general theory of which principles *should* cause worse violations. The theory makes no prediction about the relative badness of, say, θ-Criterion versus Case Filter violations, let alone about how bad each one is in some absolute sense. The notion of relative and absolute badness of particular violations is ad hoc, and is used in just those cases where it is convenient. (Schütze, 1996, 43)

This problem is not limited to Belletti and Rizzi's 1988 paper. Even a well-known syntax textbook such as Haegeman's 1994 suffers from similar difficulties. Haegeman (1994) makes extensive use of intermediate acceptability ratings, which in the absence of clear criteria on how to record and interpret intermediate judgments can lead to serious inconsistencies, as Bard et al. (1996) discuss in some detail.

These examples indicate that the use of gradient acceptability judgments is common in the linguistic literature. However, the reliance on subtle data is not matched by the necessary concern for experimental methodology. This problem has recently been addressed by a number of authors by applying the experimental technique of *magnitude estimation* to linguistic data.

Gradient acceptability judgments should be measured experimentally, since the informal elicitation technique traditionally used in linguistics is unlikely to be reliable for such data, as argued in the previous section. A suitable experimental paradigm is magnitude estimation (ME), a technique standardly applied in psychophysics to measure judgments of sensory stimuli (Stevens, 1975). The magnitude estimation procedure requires subjects to estimate the perceived magnitude of physical stimuli by assigning values on an interval scale (e.g., numbers or line lengths) proportional to stimulus magnitude. Highly reliable judgments can be achieved in this way for a whole range of sensory modalities, such as brightness, loudness, or tactile stimulation (see Stevens 1975 for an overview).

The ME paradigm has been extended successfully to the psychosocial domain (see Lodge 1981 for a survey) and recently Bard et al. (1996), Cowart (1997), and Sorace (1992) showed that linguistic judgments can be elicited in the same way as judgments for sensory or social stimuli. Unlike the five- or seven-point scale conventionally employed in the study of psychological intuition, ME allows us to treat linguistic acceptability as a continuum and directly measures acceptability differences between stimuli. ME's use of an interval scale means that parametric statistical tests can be applied for data analysis.

ME has been shown to provide 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 has been applied successfully to phenomena such as auxiliary selection (Bard et al., 1996; Sorace, 1992, 1993a,b; Sorace and Vonk, forthcoming; Keller and Sorace, 2003), binding (Cowart, 1997; Keller and Asudeh, 2001), resumptive pronouns (Alexopoulou and Keller, 2002; McDaniel and Cowart, 1999), *that*-trace effects (Cowart, 1997), extraction (Cowart, 1997), gapping (Keller, 2001), and word order (Keller and Alexopoulou, 2001; Keller, 2000a).

The ME procedure for linguistic acceptability is analogous to the standard procedure used to elicit judgments for physical stimuli. It 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, to 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.

The obvious difference between ME of physical stimuli (brightness, loudness, etc.) and ME of linguistic stimuli is that no objective measurements of acceptability are available to compare linguistic judgments against. However, the validity of linguistic ME can be established by showing that ME data are consistent when

elicited cross-modally: two groups of subjects judge the same stimuli in two different modalities and the correlation of the resulting data sets is determined. Using this approach, Bard et al. (1996) were able to obtain a high correlation between ME data elicited using numeric values and line lengths as response modalities. Similar results are reported by Cowart (1997).

Since magnitude estimation of linguistic acceptability yields data on an interval scale, it is possible to subtract the estimate given to an unacceptable sentence from the estimate given to its corresponding acceptable counterpart: the relative magnitude of the number obtained is a direct indication of the speaker's ability to discriminate between acceptable and unacceptable sentences, and therefore a correlate of the strength of preference for the acceptable sentence.

1.4   Other Experimental Paradigms

Magnitude estimation is an experimental paradigm that is very close to the methodology of intuitive judgments traditionally applied in theoretical linguistics, but at the same time imposes the rigorous controls that are standard practice for psycholinguistic experimentation.

A number of experimental paradigms are in use in psycholinguistics that can serve as alternatives to magnitude estimation. In many cases, acceptability judgments are elicited using a conventional binary scale (acceptable/unacceptable) or an ordinal scale, typically with five or seven points (e.g., Gordon and Hendrick 1997; Scheepers 1997). The data elicited this way have been shown to correlate well with magnitude estimation data (Keller, 2000b; Keller and Asudeh, 2001). [1] The disadvantage of using an ordinal scale is that it only allows for non-parametric statistics to be carried out on the results. Another widely used technique are speeded acceptability judgments, where the subject is required to classify sentences as acceptable or unacceptable as quickly as possible, and reaction times are measured (e.g., McElree and Griffith 1995).

Other experimental techniques do not require subjects to make explicit acceptability judgments, thus avoiding the problems associated with metalinguistic experimental procedures (Schütze, 1996). Pechmann et al. (1994) compared acceptability judgment data with data elicited using a number of online comprehension and production tasks, including sentence matching, delayed articulation, rapid serial visual presentation, and sentence generation. They report that the judgment data are broadly compatible with the data obtained using the online tasks (at least for word order variation, the domain they investigate).

---

[1]   See also Sorace (1992), who compared magnitude estimation judgments with data obtained from an acceptability ranking task.

Sentence matching is an experimental paradigm that is particularly relevant for the distinction between soft and hard constraints put forward in this paper. In sentence matching, subjects are presented with a sentence (fully grammatical or with a constraint violation), followed by a short pause. Then a second sentence is displayed below the first one and the subject is required to decide as quickly as possible if the two sentences are identical or not. The matching times for identical sentences are then analyzed. Freedman and Forster (1985) and Forster and Stevenson (1987) found that certain types of constraint violations (e.g., subjacency violations) do not lead to a decrease in matching times, while other violations (e.g., agreement violations) lead to increased matching times, compared to fully grammatical sentences. [2] This results is compatible with the distinction between soft and hard constraints proposed in the present paper. The prediction is that soft constraint violations should not lead to increased matching times, but hard constraint violations should.

2   Properties of Gradient Data

This section reviews the experimental literature on gradience in linguistic data. We assume a fundamental dichotomy between hard constraints (that trigger categorical linguistic judgments) and soft constraints (that trigger gradient judgments). Based on this dichotomy, we discuss how constraint ranking, cumulative effects, context, and crosslinguistic effects influence soft and hard constraints.

2.1   Soft and Hard Constraints

Many authors in the syntactic literature have drawn an intuitive distinction between two types of syntactic constraints: [3] constraints that lead to serious unacceptability when violated and constraints that merely trigger intermediate unacceptability when violated.

A widely cited example are the constraints on extraction: Belletti and Rizzi (1988), for instance, work on the assumption that a violation of the Empty Category Principle leads to serious unacceptability, while a violation of Subjacency only results in mild unacceptability. This distinction can then be used to diagnose which constraint is violated in a given structure: serious unacceptability indicates a violations of the Empty Category Principle, while mild unacceptability indicates a Subjacency violation.

---

[2]   This finding is not uncontroversial; an alternative explanation in terms of error correction strategies has been offered by Crain and Fodor (1987).

[3]   We use the term constraint in a fairly theory-neutral sense, referring to a principle or rule of grammar that can be either satisfied or violated in a given linguistic structure.

A similar distinction is proposed by Sorace (2000b), who deals with auxiliary selection in Western European languages. The starting point of her study is the following set of facts which characterize split intransitivity in a number of Western European languages: (a) across languages, some verbs tend to show consistent unaccusative/unergative behavior, whereas others do not; (b) within languages, some verbs are invariably unaccusative/unergative regardless of context, whereas others are exhibit variation. Sorace (2000b) provides supporting evidence for these generalizations, mostly based on experiments testing native speakers' intuitions about auxiliary selection (perhaps the best known diagnostic of unaccusativity) in various languages that have a choice of perfective auxiliaries (such as Dutch, German, French, Italian, and Paduan). In all these languages, unaccusative verbs tend to select the counterpart of English auxiliary *be* and unergative verbs tend to select the counterpart of auxiliary *have*; however, native intuitions on auxiliaries are categorical and consistent for certain types of verb, but much less determinate for other types. For example, native speakers have a very strong preference for auxiliary *be* with change of location verbs, but express a weaker preference for the same auxiliary (or have no preference at all) with stative verbs.

Sorace's 2000b account of these systematic differences within the syntactic classes of unaccusative and unergative verbs is that there exist gradient dimensions ("hierarchies") which distinguish "core" unaccusative and unergative verbs from progressively more "peripheral" verbs. These hierarchies, which are based on (potentially universal) aspectual parameters, identify the notion of "telic dynamic change" at the core of unaccusativity and that of "agentive non-motional activity" at the core of unergativity. The outcome is an Auxiliary Selection Hierarchy whose extremes consist of maximally distinct core verbs—verbs of change of location (e.g., *arrive*) and verbs of agentive non-motional activity (e.g., *work*)—which show categorical auxiliary choice. In contrast, peripheral verb types between the extremes are susceptible to variation and exhibit indeterminate auxiliary selection: either both auxiliaries are equally acceptable, or one auxiliary is only marginally more acceptable than the other.

Drawing on the same basic intuition that underlies Sorace's 2000b core/periphery distinction, Keller (2000b) proposes the term *hard constraints* for constraints that cause strong unacceptability when violated. *Soft constraints* are constraints whose violation causes only mild unacceptability. In this terminology the Empty Category Principle constitutes a hard constraint on extraction, while Subjacency is a soft constraint. In the same vein, core verbs violate hard constraints when they occur with the wrong auxiliary, while peripheral verbs only violate soft constraints. In Section 2.5, we will provide a more detailed explanation of this proposal.

When constraint violations are investigated experimentally (e.g., using the magnitude estimation paradigm introduced above), then violations of hard constraints are expected to be significantly less acceptable than violations of soft constraints. In addition to this basic pattern, a number of other properties of hard and soft constraints

have been investigated experimentally: constraint ranking, cumulativity, context effects, and crosslinguistic effects. In what follows, we will provide an overview of these experimental findings.

2.2 Constraint Ranking

The term constraint ranking refers to the fact that some constraints are more important than others for the grammaticality of a linguistic structure. Experimental studies generally employ an operational definition of constraint ranking based on the relative unacceptability caused by a constraint violation; the higher the degree of unacceptability caused by the constraint violation, the more important the constraint. In other words, a constraint C_1 is ranked higher than a constraint C_2 if a violation of C_1 leads to a higher degree of unacceptability than a violation of C_2. [4]

We will illustrate the concept of constraint ranking using data from Keller and Alexopoulou (2001), a magnitude estimation study of gradient word order preferences in Greek. It is sufficient in the present context to consider a subset of the constraints investigated by Keller and Alexopoulou (2001):

(1) **Constraints on word order**
    a. ACCENTALIGN (ACCAGN): in broad focus, accent has to fall on the rightmost constituent.
    b. DOUBLEALIGN (DOUAGN): preverbal objects have to be clitic doubled.
    c. DOUBLEGROUND (DOUG): clitic doubled objects have to be interpreted as ground.

The sentences in (2) provide examples for violations of the constraints in (1). The assumption is that these sentence are presented in a null context, which behaves in the same way as an all focus context. An all focus context is an instance of broad focus.

Sentence (2a) illustrates the word order svO (subject–verb–object, accent on the object, no clitic doubling). This order does not violate any constraints: ACCAGN is satisfied as the accent falls on the rightmost constituent, DOUAGN is satisfied as there is no preverbal object, and DOUG is satisfied as there is no clitic doubling.

---

[4] This definition of constraint ranking differs from the one standardly employed in Optimality Theory. In OT, constraint ranking is merely a tool for formalizing constraint competition; no direct correspondence between constraint ranks and degrees of acceptability is assumed. In OT, constraint rankings are used to determine the optimal candidate in a set of candidate structures. This optimal candidate is predicted to be grammatical; no predictions are made about suboptimal constraints and their degree of ungrammaticality. See Section 4.3 for a more detailed discussion.
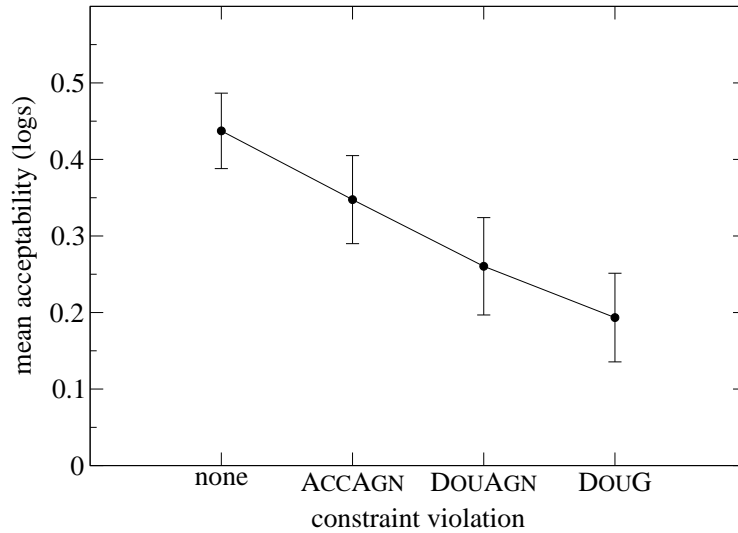
Fig. 1. Acceptability scores for word orders in Greek: svO (no violation), Svo (ACCAGN violation), ovS (DOUAGN violation), and sclvO (DOUG violation). This illustrates the ranking of soft constraints. (See Keller and Alexopoulou 2001, 344–355, for a statistical analysis of the data.)

(2) a. I Maria   apelise to YANI.
          the Maria-NOM fired-3SG the Yanis-ACC
          'Maria fi red Yanis.'
     b. ? I MARIA   apelise to Yani.
          the Maria-NOM fired-3SG the Yanis-ACC
     c. ?? To Yani   apelise i MARIA.
          the Yanis-ACC fired-3SG the Maria-NOM
     d. * I Maria   ton   apelise to YANI.
          the Maria-NOM him-CL fired-3SG the Yanis-ACC

Example (2b) illustrates the order Svo (subject–verb–object, accent on the subject, no clitic doubling). This order violates ACCAGN as the accent falls on *Maria*, i.e., it fails to fall on the rightmost constituent. The order ovS in (2c) (object–verb–subject, accent on the subject, no clitic doubling) violates DOUAGN, as there is a preverbal object, but it is not clitic doubled. Finally, (2d) is an example for the order sclvO (subject–verb–object, accent on the object, clitic doubling). This order violates DOUG as the doubled object is in focus (recall that we are assuming an all focus context).

Figure 1 shows the mean acceptability judgments reported by Keller and Alexopoulou (2001) for sentences such as the ones in (2). These data show that the constraints ACCAGN, DOUAGN, DOUBLEGROUND differ in constraint ranking. A violation of ACCAGN leads to a lower degree of unacceptability than a violation of DOUAGN, which in turn leads to a lower degree of unacceptability than a violation DOUG. This demonstrates that these three constraints are ranked in the sense of the operational defi nition of constraint ranking that we introduced above.

The ranking of constraints has also been investigated for word order constraints in German (Keller, 2000a) and for constraints on gapping (Keller, 2001) and binding (Keller and Asudeh, 2001; Cowart, 1997) in English. It is important to note that constraint ranking can occur not only for soft constraints such as the constraints on word order discussed in the present section. Also hard constraints can be shown to be ranked; an example for this will be given in the next section.

A general note on the data displayed in Figure 1, and in the other figures in this paper. All figures show log-transformed acceptability judgments, averaged over a population of native speakers. The error bars represent standard errors. Due to space limitations, we will omit a detailed description of the experiments on which the figures are based. However, each figure captions gives a reference to the source of the original data, which contains a full discussion of the materials and the procedure used, and a statistical analysis of the results.

2.3 Cumulativity

Another important question that has been addressed experimentally concerns the interaction of constraints. By constraint interaction we mean the behavior of linguistic structures that incur multiple constraint violations. Again, an operational definition can be used here: the interaction of two constraints can be determined by investigating the degree of unacceptability of a structure that violates both constraints, and comparing it to the degrees of unacceptability of structures that violate only one of the two constraints. An accurate picture of constraint interaction can be built up by investigating structures that violate constraints of different ranks and types.

There is substantial evidence indicating that constraints interact in a *cumulative* fashion: the more constraint violations a structure incurs, the less acceptable it becomes. This property is shared by hard and soft constraints (as will be discussed below). As an example consider Keller's 2000b magnitude estimation study of extraction from picture-NPs in English. This experiment investigated the interaction of three soft constraints and three hard constraints: [5]

(3) **Soft constraints on extraction**
 a. DEFINITENESS (DEF): a picture NP has to be marked $[-\text{DEFINITE}]$.
 b. VERBCLASS (VERB): a verb subcategorizing for a picture NP has to be marked $[-\text{EXISTENCE}]$.
 c. REFERENTIALITY (REF): an NP extracted from a picture NP has to be marked $[+\text{REFERENTIAL}]$.

---

[5] The precise formulation of the constraints is irrelevant here (definitions for features such as $[-\text{EXISTENCE}]$, etc.), as we are only interested in the interaction of the constraints.

(4) **Hard constraints on extraction**
    a. INVERSION (INV): subject and auxiliary have to be inverted.
    b. AGREEMENT (AGR): subject and verb have to agree in number.
    c. RESUMPTIVE (RES): resumptive pronouns are disallowed.

The examples in (5) illustrate the soft constraints in (3). Sentence (5a) does not incur any constraint violations. It is more acceptable than (5b), which incurs a DEF violation, i.e., extraction takes place out of a definite picture NP. (5c) violates VERB, the constraint that requires that the matrix verb does not presuppose the existence of its object. VERB is satisfied by *paint*, but violated by *tear up*.[6]

(5d) shows how the extraction of a non-referential *wh*-phrase violates the constraint REF, leading to a decrease in acceptability compared to (5a).

(5) a.   Which friend has Thomas painted a picture of?
    b. ?Which friend has Thomas painted the picture of?
    c. ?Which friend has Thomas torn up a picture of?
    d. ?How many friends has Thomas painted a picture of?

The examples in (6) illustrate violations of the hard constraints in (4). Sentence (6a) incurs a violation of the constraint INV, as the subject and the auxiliary are not inverted. (6b) violates AGR, as there is no number agreement between the subject and the verb. Finally, example (6c) includes the resumptive *her* instead a gap, which violates RES, the constraint against resumptives.

(6) a. *Which friend Thomas has painted a picture of?
    b. *Which friend have Thomas painted a picture of?
    c. *Which friend has Thomas painted a picture of her?

Figure 2 graphs the mean acceptability judgments reported by Keller (2000b) for sentences such as the ones in (2). Each sentences incurs between zero and three violations of the soft and hard constraints listed in (3) and (4). Note that there is a clear distinction between soft and hard constraints: hard constraint violations trigger a significantly higher degree of unacceptability than soft constraint violations. Figure 2 also illustrates the cumulativity effect for constraint violations: the degree of unacceptability of a structures increases with the number of violations it incurs, both for hard and soft constraints.

Figure 3 (also taken from Keller's 2000b experiment on extraction from picture NPs) graphs the acceptability scores for single violations of the hard constraints INV, AGR, and RES. Statistical tests show that a single violation of RES is significantly more unacceptable than single violations of INV and AGR. This illustrates

---

[6] The verb *paint* is used here in the sense of 'produce an image with a brush", which presupposes that the image being painted does not exist yet. Other senses of *paint* do not carry this presupposition.
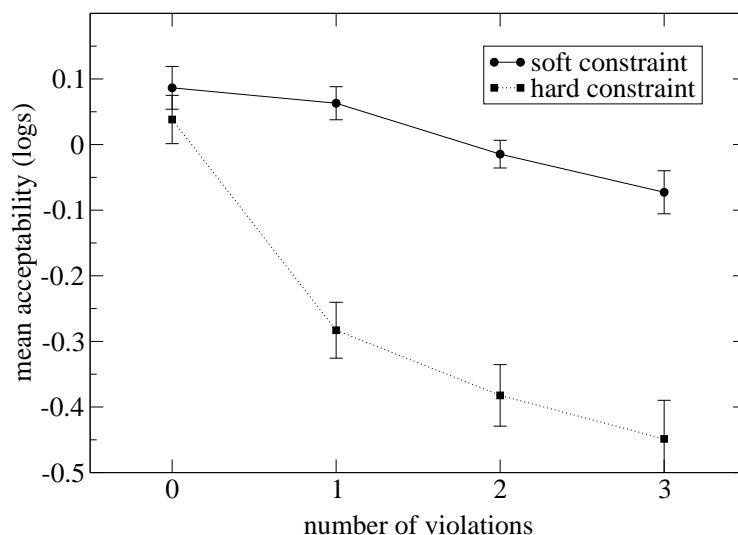
Fig. 2. Acceptability scores for extraction from picture-NPs in English involving multiple constraint soft and hard constraint violations. This illustrates the cumulativity of soft and hard constraint violations. (See Keller 2000b, 90–94, for a statistical analysis of the data.)

that constraint ranking also occurs for hard constraints.

Figure 3 also illustrates the *ganging up* of constraints. The constraint RES is ranked higher than both INV and AGR. However, the combined violation of INV and AGR is as unacceptable as a single violation of RES. In other words, the two lower ranked constraints INV and AGR gang up against the higher ranked constraint RES, resulting in the same degree of unacceptability.

Such ganging up effects are predicted to be impossible under the assumption of *strict domination* of constraints (as in Optimality Theory). Strict domination means that multiple violations of lower ranked violations are not able to compensate for a single violation of a higher ranked constraint (for a more detailed discussion of the issue of strict domination see Guy 1997; Keller and Asudeh 2002).

It is important to note that cumulative effects and ganging up effects are not limited to extraction. Keller (2000b) demonstrates both effects also for word order constraints and for constraints on gapping. Cumulativity effects were also obtained for multiple violations of selectional restrictions and subcategorization requirements by Chapman (1974). His results also demonstrate that cumulativity effects do not only occur for multiple violations of different constraints, but also for multiple violations of the same constraint. Furthermore, cumulativity also seems to hold for phonological constraints: Guy and Boberg (1997) presents data that show that violations of the Obligatory Contour Principle are cumulative. [7]

---

[7] Guy and Boberg (1997) formulate the Obligatory Contour Principle as *[αF] [αF]: feature sharing with the preceding segment is disallowed.
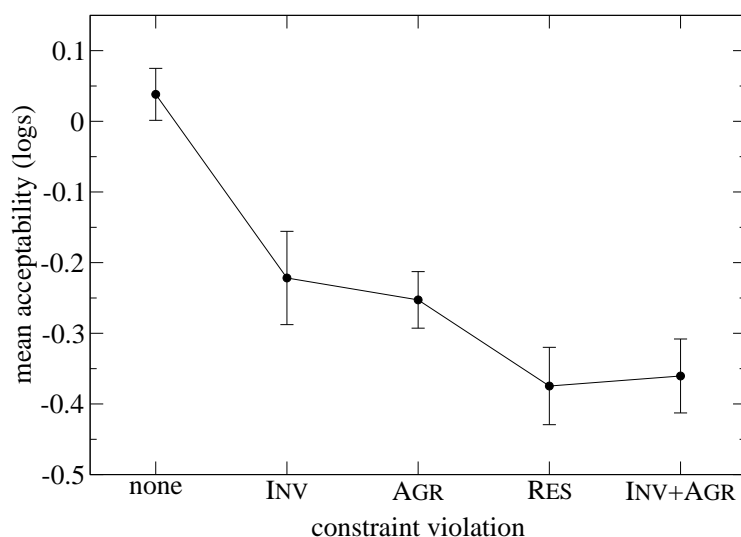
Fig. 3. Acceptability scores for extraction from picture-NPs in English involving single and double hard constraint violations. This illustrates the ranking of hard constraints and shows the existence of ganging up effects. (See Keller 2000b, 90–94, for a statistical analysis of the data.)

2.4 Context Effects

When linguists are confronted with gradient evidence, e.g., with a sentence S that is of reduced acceptability, but not fully unacceptable, they often resort to an argumentation like the following. They try to find a specific context C in which S is fully acceptable (or at least of increased acceptability). Having found such a context, they conclude that the structure instantiated by S is grammatical, a fact from which certain theoretical conclusions can be drawn. This strategy implies that "S is acceptable" actually means "there is a context in which S is acceptable". The problem with such an approach is that it fails to recognize the distinction between sentences that are acceptable *without* requiring a specific context, and ones that are *only* acceptable in a specific context, as pointed out by Chomsky (1964, 385).

However, this distinction traditionally plays an important role in the research on word order variation, going back to Lenerz (1977) and Höhle (1982) (among others). To capture the influence of context on word order, Höhle (1982, 102, 122) defines the notion of *markedness* as follows: a sentence S_1 is less marked than a sentence S_2 if S_1 can occur in more context types than S_2. Based on this definition Müller (1999a) proposes that markedness corresponds to speakers' intuitions about gradient acceptability, i.e., that "relative degrees of markedness can be empirically determined [...] either by directly invoking speakers' judgments, or by adhering to the number of context types in which the candidate [i.e., the sentence] is possible" (Müller, 1999a, 782f). This statement can be reinterpreted based on the soft/hard dichotomy proposed in the present paper: soft constraints are subject to context effects, while hard constraints are immune to context effects.

14

Keller's 2001 magnitude estimation study of gapping in English provides evidence for this hypothesis by investigating a set of constraints on gapping (originally proposed by Kuno (1976)). We will focus on the constraints MINDIS and SIMS, defined as follows:

(7) **The Minimal Distance Principle** [MINDIS] (Kuno, 1976, 306)
The two constituents left behind by Gapping can be most readily coupled with the constituents (of the same structures) in the first conjunct that were processed last of all.

(8) **The Requirement for Simplex-Sentential Relationship** [SIMS] (Kuno, 1976, 314)
The two constituents left over by Gapping are most readily interpretable as entering into a simplex-sentential relationship. The intelligibility of the gapped sentence declines drastically if there is no such relationship between the two constituents.

Kuno's 1976 examples in (9) illustrate the constraint MINDIS: In (9a), the remnant *Tom* has to be paired with *Mary*, yielding the interpretation in (9b). It is not possible to pair *Tom* with the more distant subject *John*, yielding the interpretation in (9c).

(9) a. John believes Mary to be guilty, and Tom to be innocent.
b. John believes Mary to be guilty, and John believes Tom to be innocent.
c. John believes Mary to be guilty, and Tom believes Mary to be innocent.

The constraint SIMS states that gapping cannot leave behind remnants that are part of a subordinate clause: (10a) cannot be understood as a gapped version of (10b).

(10) a. John persuaded Dr. Thomas to examine Jane and Bill Martha.
b. John persuaded Dr. Thomas to examine Jane and Bill persuaded Dr. Thomas to examine Martha.

A felicitous context for gapped sentences such as the ones in (9) and (10) is a context in which the constituents deleted by gapping are contextually known, while the constituents left behind by gapping represent new information (Kuno, 1976). An infelicitous context represents the inverse situation: the gapped constituents are new, while the remnant constituents are contextually known.

Figure 4 graphs the mean acceptability judgments reported by Keller (2001) for sentences violating the constraints MINDIS and SIMS in felicitous and non-felicitous contexts. The constraint MINDIS shows context effects: in the non-felicitous context, sentences violating MINDIS are significantly less acceptable than sentences not violating MINDIS. In the felicitous context, however, there is no significant difference between these two conditions (though a small acceptability difference remains). The constraint SIMS, on the other hand, fails to show context effects: a SIMS violation is equally unacceptable in a non-felicitous and in a
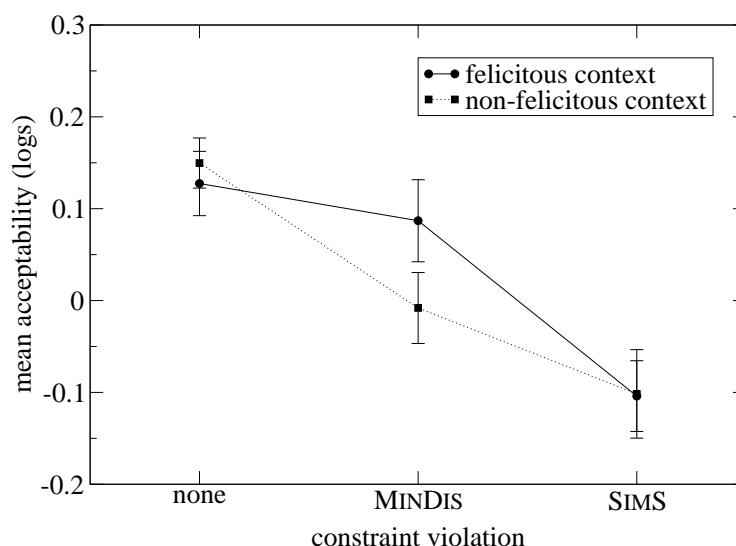
Fig. 4. Acceptability score for gapping constructions in English involving violations of the soft constraint MINDIS and the hard constraint SIMS in four different contexts. This illustrates that there are context effects in soft constraints, but not in hard constraints. (See Keller 2001, 236–241, for a statistical analysis of the data.)

felicitous context. [8]

These results support the hypothesis that hard constraint violations are equally un-acceptable in all contexts, while soft constraint violations are context dependent, i.e., the degree of unacceptability triggered by soft constraint violations can change from context to context. Further experimental support for this hypothesis is provided by Keller (2000a) and Keller and Alexopoulou (2001), who investigate context effects for word order constraints in Greek and German.

2.5 Crosslinguistic Effects

Sorace (2000b) provides a review of experimental data on gradient auxiliary selection in French, Italian, Dutch, and German (for the original experiments see Bard et al. 1996; Sorace 1992, 1993a,b; Sorace and Vonk forthcoming; Sorace and Cennamo 2000; Keller and Sorace 2003). She demonstrates that core verbs show categorical auxiliary selection behavior and are not subject to crosslinguistic variation.

---

[8] MINDIS is a soft constraint, and SIMS is a hard constraint, hence it is expected that SIMS violations lead to substantially higher unacceptability than MINDIS violations. This does not seem to be the case in the non-felicitous context in Figure 4, where the difference between the MINDIS and the SIMS violation is small. This leads Keller (2001) to argue that seriousness of violation is not the only criterion for assessing the type of a constraint. Rather, in order to determine if a constraint is soft or hard, one has to take into account other factors as well, such as context effects, crosslinguistic effects, and evidence from language development.

This means that core verbs select the same auxiliary across languages, and that the other auxiliary is severely unacceptable. Peripheral verbs, on the other hand, are subject to gradience and crosslinguistic variation: both auxiliaries are acceptable to a certain degree, and the choice of auxiliary can vary from language to language.

In Section 2.1, we argued that hard constraints are violated if a core verb occurs with the wrong auxiliary. Peripheral verbs, on the other hand, only violate soft constraints if they occur with the wrong auxiliary, triggering mild unacceptability. Together with Sorace's 2000b observations regarding the crosslinguistic validity of the distinction between core and peripheral verbs, this leads to the hypothesis that the type of a constraint is crosslinguistically stable: hard constraints are hard across languages, while soft constraints are soft across languages. This means that crosslinguistic variation can change the rank of a constraint, but not its type (hard or soft).

This idea is developed further by Legendre et al. (to appear) and Legendre and Sorace (2003) within an OT framework. Based on the notion of *harmonic alignment of prominence scales* (Prince and Smolensky, 1993), the Auxiliary Selection Hierarchy is reformulated in terms of a universal hierarchy of alignment constraints shown in (11), where "1" is subject, "2" is direct object, and the constraints are to be understood as "avoid" in the sense of Aissen (1999):

(11) *1/TELIC ≫ *1/MOTIONAL ≫ *1/DIRECTIONAL ≫ *1/CONTINUATIVE ≫ *1/STATIVE ≫ *1/CONTROL

The hierarchy represents semantic classes of verbs (roughly) corresponding to points on the Auxiliary Selection Hierarchy and embodies the fact that avoiding the projection of the argument of telic verb onto the subject position is more important than avoiding it with other verb classes. The hierarchy in (11) is consistent with the assumption that, in languages that have auxiliary selection, telicity is the main factor that distinguishes BE verbs from HAVE verbs, whereas control is a secondary factor that further differentiates among HAVE verbs (these further differentiations are not shown here). Violations of higher-ranked (hard) telicity constraints therefore lead to more serious ungrammaticality than violations of lower-ranked (softer) constraints.

Crosslinguistic variation in split intransitivity is given by an intervening constraint *2 ("don't project to direct object"), whose location may vary from language to language, and which splits the hierarchy into verbs that display unaccusative syntax (to the left of *2) and verbs that display unergative syntax (to its right). The difference between French and Italian, for example, would be represented in (2.5)

(i) a. **French:**
        1/TELIC ≫ *1/MOTIONAL ≫ *2 ≫ *1/DIRECTIONAL ≫
        *1/CONTINUATIVE ≫ *1/STATIVE ≫ *1/CONTROL
    b. **Italian:**
        1/TELIC ≫ *1/MOTIONAL ≫ *1/DIRECTIONAL ≫ *1/CONTINUATIVE
        ≫ *1/STATIVE ≫ *2 ≫ *1/CONTROL

The different position of the intervening constraint in the two languages accounts for the fact that the verbs selecting auxiliary *être* in French are a subset of the verbs selecting *essere* in Italian and are largely restricted to inherently telic verbs of change of location (see Legendre and Sorace 2003, for discussion).

3 Gradience and Language Development

It has been argued that the transitional grammars constructed by language learners, both in first and in second language acquisition, exhibit syntactic optionality, that is the existence of more than one realization of a given input. Typically, such optional variants differ in their degree of acceptability and the relative acceptability of optional structures changes over time in the course of development. Gradience in language development is therefore correlated with optionality.

In child grammars, target and non-target structures co-exist, while in adult second language acquisition, learners usually approximate target acceptability patterns. However, optionality tends to persist into advanced stages of development, and "residual" optionality may be regarded as one of the features of the near-native grammatical competence. Sorace (1993a,b) demonstrates this effect for auxiliary selection in French and Italian, while Robertson and Sorace (1999) show that there is residual optionality with respect to verb-second in German advanced speakers of English.

Optionality and gradience have also been recently investigated in a different domain of language development, namely language attrition. Attrition refers to the changes that occur in the native language competence of experienced second language users. According to Sorace (2000c), one manifestation of attrition is "emerging" optionality in areas of the grammar that are normally determinate or categorical in monolingual speakers. A relevant example is the use and interpretation of null and overt referential subjects in Italian. A referential subject is obligatorily null in Italian if it is coreferential with a topic antecedent (Grimshaw and Samek-Lodovici, 1998); otherwise the subject is overtly expressed. It is this discourse constraint on the distribution of subject pronouns that is subject to attrition in native Italian speakers who have been exposed to English for a long time. As a result, these speakers may overproduce overt subjects when coreferentiality with a topic obtains; however, they never produce null subjects in inappropriate contexts. In contrast, the syntac-

18

tic parameter responsible for licensing null subjects is not affected by attrition (see arguments in Sorace 2000c, 2003). In fact, native speakers of English under sustained exposure to Italian do not exhibit null subjects in English, although they do overproduce overt subject pronouns just like Italians under attrition. Advanced non-native grammars thus present a parallel pattern of optionality to native grammars in a situation of attrition.

The overall indication that can be drawn from these findings is that developmental optionality seems to be limited to certain types of constraints. Optionality in auxiliary selection, for example, seems to concern only peripheral verb classes, whereas the auxiliary selection behavior of core verb classes is rapidly acquired in non-native grammars. Verb-second phenomena are ultimately related to the instantiation of the illocutionary force of a sentence: they are a particular syntactic realization of speech act information, such as discourse anchoring and information structure. The residual optionality exhibited by German non-native speakers of English is related to their imperfect acquisition of the different (and much more restrictive) conditions required by English. Attrition affects the discourse conditions on the realization of subjects pronouns, but not the syntactic parameters that license null subjects. In sum, this lends plausibility to the hypothesis that soft constraints are subject to developmental optionality, whereas hard constraints are relatively immune to optionality effects.

4 Modeling Gradient Data

The current section summarizes the main properties of gradient judgments to be modeled and provides a critical assessment of the most important proposals for models of gradience that have been proposed in the literature. Two main types of models will be discussed: traditional models based on weighted rules and recent approaches based on Optimality Theory. We will focus on models of gradience in adult native speakers, which is what the bulk of the modeling literature deals with.

4.1 Properties to be Modeled

Based on experimental data covering a range of syntactic phenomena in several languages, we were able to identify a number of sources of gradience in grammar in Sections 2 and 3.

The two central findings that we reported are that constraints are ranked and that constraint violations are cumulative. Constraint ranking means that some constraint violations are significantly more unacceptable than others. Cumulativity means the multiple constraint violations are significantly more unacceptable than single vio-

lations. These properties seem to be fundamental to the behavior of gradient linguistic judgments and therefore should form the basis of a model of gradience in grammar. Cumulativity also accounts for the ganging up effect that was observed for constraints on word order, extraction, and gapping: multiple violations of lower ranked constraints can be as unacceptable as a single violation of a higher ranked constraint.

Another central experimental result is that constraints can be classified into two types, soft and hard. While both types of constraint share the properties of ranking and cumulativity, they differ in another set of properties (as summarized in Table 1). First, soft constraint violations are associated with mild unacceptability, while hard violations trigger serious unacceptability. Second, soft constraints are context-dependent, while hard constraints are immune to context effects. Third, only soft constraints are subject to developmental optionality in first and second language acquisition and in language attrition: hard constraints are developmentally stable, even in second language speakers and in first language speakers subjects to attrition.

A fourth property of gradient structures concerns crosslinguistic variation. Both hard and soft constraints are subject to crosslinguistic variation; however, we hypothesizes that crosslinguistic variation cannot affect the type of a constraint, i.e., we predict that there are no constraints that are soft in one language and hard in another. (Note that this presumes that there is an inventory of universal constraints for all languages, as assumed in Optimality Theory.)

Based on these properties, we can operationalize the notion of constraint type. If a constraint violation induces strong unacceptability and fails to show context effects and developmental optionality, then it can be classified as a hard constraint. If a constraint triggers only mild unacceptability and is subject to contextual variation and developmental optionality, then the constraint is soft. The classification can be verified by investigating the crosslinguistic behavior of the constraint; the type of a constraint (soft or hard) should remain the same across languages.

A model of gradience in grammar should account for the experimental properties of gradient linguistic structures that were summarized in this section. In the following section, we will use these properties as a yardstick against which to evaluate the proposal for such a model in the literature.

4.2 Weighted Rules

Some researchers in the generative tradition have adopted weighted rule models as a way of accounting for degrees of grammaticality. An example is Uszkoreit's 1987 account of word order preferences. In this framework, grammatical rules are annotated with numeric weights that reflect their importance in determining grammati-

Table 1
Properties of hard and soft constraints

	hard constraints	soft constraints
universal effects	ranking effects	ranking effects
	cumulativity effects	cumulativity effects
	ganging up effects	ganging up effects
type-specific effects	strong unacceptability	mild unacceptability
	no context effects	context effects
	no developmental optionality	developmental optionality
crosslinguistic effects	constraints are hard across languages	constraints are soft across languages

cality (for a similar proposal see Jacobs 1988). Uszkoreit (1987) assumes constraint competition, i.e., not all constraints are necessarily satisfiable in a given linguistic structure. In this model, grammaticality is a gradient notion; the degree of grammaticality of a linguistic structure is computed as the sum of the weights of the constraint violations the structure incurs.

However, Uszkoreit's 1987 approach remains on an intuitive level; it is not founded on experimental evidence (but a partial experimental confirmation was later provided by Pechmann et al. 1994). Also, the Uszkoreit model only deals with word order variation; it remains to be seen if the approach is general enough to handle gradience in other parts of the grammar as well. Another problem is that Uszkoreit (1987) fails to make explicit how the rule weights are estimated from judgment data; it seems that this is left to the intuition of the linguist.

Another weighted grammar model is the Variable Rule model proposed by Labov (1969) and Cedergren and Sankoff (1974) to account for sociolinguistic variation (mainly in phonology). The Variable Rule model differs from the approach proposed by Uszkoreit (1987) in that it is probabilistic, and comes with an algorithm for parameter estimation (the Variable Rule model is essentially a log-linear model of the frequency distribution of grammatical forms in a corpus). The model is in principle compatible with the data discussed in Sections 2 and 3; the ranking of constraints can be naturally expressed through rule weights. In this setting, hard constraints receive high weights, while soft ones receive low weights. Furthermore, the Variable Rule model assumes that rule violations are cumulative, which is in line with the experimental evidence discussed in Section 2.3.

The Variable Rule model is based on quantitative data (corpus data), and is equipped with an estimation scheme that determines the rule weights from corpus frequencies. It seems unlikely that this approach can be extended to deal with gradient phenomena like the ones investigated in this paper. Gradient structures are

typically extremely infrequent in a corpus (or fail to occur at all), and hence pose a serious sparse data problem for the estimation algorithm, making it difficult to obtain reliable estimates of the weights of gradient constraints. Note that this problem is an instance of the general unavailability of negative evidence from corpora—gradient structures constitute an instance of negative evidence (by virtue of being unacceptable, at least to a certain degree).

4.3   Optimality Theory

In line with all major linguistic frameworks, Standard Optimality Theory assumes a binary notion of grammaticality: the competition between candidate structures selects one candidate (or a set of candidates sharing the same constraint profile) as optimal and, hence, grammatical. All losing candidates, i.e., those structures that are *suboptimal*, are assumed to be ungrammatical; Standard OT makes no predictions about the relative ungrammaticality of suboptimal candidates. This binary view of grammaticality is inadequate for data that exhibit a continuum of degrees of acceptability, such as the data reported in the present paper. However, a number of proposals exist in the literature that propose to extend OT to deal with gradience.

4.3.1   Naive Extension of Standard Optimality Theory

A straightforward model of gradient grammaticality can be obtained by extending the Standard OT notion of grammaticality. We will refer to this approach as the *Naive Extension of Standard Optimality Theory*.

In Standard OT, grammaticality is defined as global optimality for the whole candidate set. This can be complemented by a definition of *suboptimality* as local optimality relative to a subset of the candidate set (Müller, 1999a). A structure S_1 is suboptimal with respect to a structure S_2 if there are subsets R_i and R_j of the candidate set such that S_1 is optimal for R_i and S_2 is optimal for R_j and R_j is a subset of R_i. To model gradience, we can then assume that S_1 is less grammatical than S_1 if S_1 is suboptimal with respect to S_2. Intuitively, this definition entails that the relative grammaticality of a structure corresponds to its optimality theoretic rank in the candidate set. This model predicts a grammaticality ordering for the structures in the candidate set, which can then be tested against the acceptability ordering found experimentally for the candidates.

However, the Naive Extension of Standard OT encounters a number of serious problems when faced with experimental evidence such as the one presented in Sections 2 and 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 Extension of Standard OT.

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 Extension of Standard OT predicts that there is a grammaticality difference whenever there is a difference in optimality. This carries the danger that the Naive Extension of Standard OT will overgenerate, in the sense of predicting more grammaticality differences than are justified by the data (for details see Müller 1999a).

The cumulativity of constraint violations poses a third problem for the Naive Extension of Standard OT as a 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 Extension: 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. In particular, strict domination is incompatible with the ganging up of constraint violations (see Section 2.3).

4.3.2 Grammaticality and Markedness

Müller (1999a) discusses the shortcomings of the Naive Extension of Standard OT and proposes an alternative, the markedness model. This approach assumes a distinction between grammaticality (manifested in binary judgments) and markedness (associated with gradient judgments). Grammaticality is handled in terms of Standard OT-style constraint competition. All candidates that are suboptimal in this competition are predicted to be categorically ungrammatical. For certain phenomena, the competition will produce not a single optimal candidate, but a set of optimal candidates. All of these candidates are predicted to be grammatical; however, they take part in a further optimality theoretic competition based on a separate set of constraints, so-called markedness constraints. The optimal candidate in this competition is *unmarked*; the suboptimal candidates are more or less marked (dispreferred) depending on their relative suboptimality.[9]

Müller's 1999a model provides a direct account for the distinction between hard and soft constraints in terms of markedness. It also has the advantage of avoiding the prediction of bogus grammaticality differences (a problem for the Naive Extension): gradience is only induced by a subset of the constraints (the markedness constraints), which take part in a separate constraint competition.

---

[9] Note that Müller (1999a) relies on the definition of markedness proposed by Höhle (1982, 102, 122): a given sentence S_1 is less marked than a sentence S_2 if S_1 can occur in more context types than S_2. See Section 2.4 for a brief discussion.

On the other hand, Müller's 1999a approach inherits from the Naive Extension of Standard OT the problem that only structures in the same candidate set can be compared as to their relative grammaticality. Just as the Naive Extension, Müller's 1999a model is unable to account for the cumulativity of and the ganging up of constraint violations because the competition of markedness constraints uses the Standard OT constraint evaluation scheme based on strict domination.

4.3.3 Constraint Re-ranking

Keller (1998) suggests an alternative model of gradience that draws on concepts from OT learnability theory (Tesar and Smolensky, 1998). The core idea of this approach is to compute the relative grammaticality of a suboptimal structure by determining which constraint re-rankings are required to make the suboptimal structure optimal. This information can then be used to compare structures with respect to their degree of grammaticality: the assumption is that the degree of grammaticality of a candidate structure S depends on the number and type of re-rankings required to make S optimal. Such a re-ranking model offers the necessary flexibility to accommodate the experimental findings on constraint ranking and constraint interaction discussed in Section 2:

- The re-ranking model allows us 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 allows us to account for the difference between hard constraints (serious re-rankings) and soft constraints (less serious re-rankings).
- 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.

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 can be dealt with by a single re-ranking, and hence we fail to predict a

cumulativity effect here. This prediction is not in accordance with the experimental facts: Keller (2000b) found that the cumulativity of constraint violations extends to multiple violations of the same constraints. A similar result was obtained by Chapman (1974), who investigated multiple violations of selectional restrictions and subcategorization requirements.

Another problem concerns the case of unmarked competitors. The algorithm proposed by Keller (1998) demotes the constraints violated by a structure S_1 below the ones violated by a given competitor S_2, so that S_1 becomes optimal. The degree of grammaticality of S_1 depends on the type and number of re-rankings required in this demotion process. Constraint demotion is impossible, however, if the competitor S_2 is completely unmarked, i.e., if it incurs no constraint violations at all, which entails that S_2 is optimal under any constraint ranking. The notion of relative grammaticality is not well-defined in this case, as no constraint demotion can take place.

A third problem with the re-ranking model concerns the absence of a learning algorithm for determining constraint ranks from a set of judgment data: Keller (1998) does not provide a systematic method for computing an adequate constraint hierarchy based on a set of gradient linguistic judgments.

4.3.4 Probabilistic Optimality Theory

Boersma and Hayes (2001) propose a probabilistic variant of Optimality Theory (POT) that is designed to account for corpus frequencies and degrees of grammaticality. The POT framework has been applied in phonology (Boersma and Levelt, 2000; Boersma, 1997, 1998; Boersma and Hayes, 2001; Boersma, 2000; Hayes, 2000; Hayes and MacEachern, 1998), morphology (Boersma and Hayes, 2001; Hayes, 1997), and syntax (Asudeh, 2001; Bresnan et al., 2001; Dingare, 2001; Koontz-Garboden, 2001).

The POT model can be viewed as a numeric version of the re-ranking model discussed in the last section: it stipulates a continuous scale of *constraint strictness*. Constraints are annotated with numerical strictness values; if a constraint C_1 has a higher strictness value that a constraint C_2, then C_1 outranks C_2. Boersma and Hayes (2001) assume *probabilistic constraint evaluation*, which means that at evaluation time, a small amount of random noise is added to the strictness value of a constraint. As a consequence, *re-rankings* of constraints are possible if the amount of noise added to the strictness values exceeds the distance between the constraints on the strictness scale.

For instance, assume that two constraints C_1 and C_2 are ranked $C_1 \gg C_2$, selecting the structure S_1 as optimal for a given input. Under Boersma and Hayes's 2001 approach, a re-ranking of C_1 and C_2 can occur at evaluation time, resulting in the opposite ranking $C_2 \gg C_1$. This re-ranking might result in an alternative optimal

candidate S_2. The probability of the re-ranking that makes S_2 optimal depends on the distance between C_1 and C_2 on the strictness scale (and on the amount of noise added to the strictness values). The re-ranking probability is assumed to predict the degree of grammaticality of S_2. The more probable the re-ranking $C_2 \gg C_1$, the higher the degree of grammaticality of S_2; if the rankings $C_1 \gg C_2$ and $C_2 \gg C_1$ are equally probable, then S_1 and S_2 are equally grammatical.

The POT framework comes with its own learning theory in the form of the Gradual Learning Algorithm (Boersma, 1998, 2000; Boersma and Hayes, 2001). This algorithm is a generalization of Tesar and Smolensky's 1998 Constraint Demotion Algorithm: it performs constraint promotion as well as demotion. The existence of a learning algorithms overcomes a limitation of Keller's 1998 re-ranking model, for which no learning algorithm is available. It means that a theory of language development is available for the POT, which might make it possible to account for the facts from language acquisition and attrition discussed in Section 3.

Just as the re-ranking model, POT has the advantage of allowing us to compare the relative grammaticality of arbitrary structures. It also allows provides a natural account for the dichotomy between hard and soft constraints: hard constraints have a very low (near-zero) re-ranking probability, while soft constraints have a higher re-ranking probability. This means that the POT model makes testable predictions for the behavior of hard and soft constraints with respect to context effects: hard constraints are expected to have the same re-ranking probability in all context, while the re-ranking probabilities of soft constraints are expected to be subject to contextual variation. POT is also able to naturally account for crosslinguistic effects (see Section 2.5), as it inherits the Standard OT account of crosslinguistic variation as constraint re-ranking.

However, like the re-ranking model, the POT approach is not able to deal with unmarked competitors—an unmarked competitor does not incur any constraint violations, i.e., it is always optimal, no matter which re-rankings are assumed. Furthermore, the POT model is unable to account for cumulativity effects involving multiple violations of the same constraints: a multiple violation will only trigger a single re-ranking, no matter how often the constraint is violated. A related problem concerns ganging up effects, which cannot be accounted for in POT, as this framework assumes OT's strict domination of constraints. These problems with POT are discussed in more detail by Keller and Asudeh (2002).

4.3.5 *Linear Optimality Theory*

Keller (2000b) proposes an alternative model of gradience based on Optimality Theory, called Linear Optimality Theory (LOT). Although this model borrows central concepts (such as constraint ranking and competition) from Optimality Theory, it differs in two crucial respects from existing OT-based accounts. Firstly, it as-

sumes that constraint ranks are represented as numeric weights (it shares this feature with POT). Secondly, it assumes that the grammaticality of a given structure is proportional to the sum of the weights of the constraints it violates. This means that OT's notion of strict domination is replaced with an linear constraint combination scheme.

Linear Optimality Theory comes with a *learning algorithm*. This algorithm takes as its input a grammar (i.e., a set of linguistic constraints) and a training set, based on which it estimates the weights of the constraints in the grammar. [10] The training set is a collection of candidate structures, where the violation profile and the grammaticality score for each structure is specified.[11] It is important to note that LOT is not intended as a model of human language acquisition: it cannot be assumed that the learner has access to training data that are annotated with grammaticality scores. The purpose of the LOT learning algorithm is to perform parameter fitting for LOT grammars, i.e., to determine an optimal set of constraint weights for a given data set. The algorithm is not designed to be cognitively plausible.

LOT is able to account for the properties of gradient structures discussed in Section 2 and 3. Constraint ranking is modeled by the fact that LOT annotates constraints with numeric weights representing the contribution of a constraint to the unacceptability of a structure. Cumulativity is modeled by the assumption that the degree of ungrammaticality of a structure is computed as the sum of the weights of the constraint the structure violates. Once ranking and cumulativity are assumed as part of the LOT model, the other properties of gradient linguistic judgments follow without further stipulations. The ganging up effect is an obvious case; as constraint violations are additive, they can gang up. As an example assume that the weight of the constraint C_1 is twice that of the constraint C_2. Then a structure that violates C_1 once will be as ungrammatical as one that violates C_2 twice, i.e., C_2 gangs up against C_1.

The distinction between soft and hard constraints does not have to stipulated in LOT; rather it emerges a consequence of LOT's assumptions about ranking and cumulativity and its learning algorithm. The LOT learning algorithm assigns high weights to constraints that cause serious ungrammaticality when violated, and low weights to constraints that cause only mild ungrammaticality. This means that all hard constraints will receive a similar constraint weight, and the same will be true for all soft constraints.

---

[10] Note that the constraints themselves are not learned in LOT, only the constraint weights are. This makes it possible to assume that constraints are universal, i.e., they are the same across languages. Only the constraint weights (the constraint ranking in standard OT) vary from language to language.

[11] It can be shown that the problem of determining LOT constraint weights reduces to solving a system of linear equations, a familiar mathematical problem (Keller, 2000b). Efficient and well-understood algorithms such as Least Square Estimation can therefore be applied as learning algorithms for LOT.

Another difference between soft and hard constraints is that soft, but not hard constraints can be subject to context effects. In LOT this makes the prediction that the weight of a soft constraint will vary with context, while the weight of a hard constraint will be stable across contexts. This prediction can be tested by applying the LOT learning algorithm to two training sets, containing grammaticality scores for the same structures, but in two different contexts. The learning algorithm should compute the same weights for both training sets for hard constraints, while the weights for soft constraints can vary between the two sets. Keller (2000b) presents modeling results that show that this prediction is born out for data on gapping and word order.

The same line of reasoning can be applied to developmental optionality, which only occurs with soft constraints: in an LOT setting, this predicts that the weights of soft constraints will change during language development, while the weights of hard constraints will remain approximately constant. This prediction can be tested by applying the LOT learning algorithm to training sets that contain grammaticality scores for different stages of language development. Only the weights of hard constraints are expected to be constant across these training sets.

In Section 2.5 we argued that crosslinguistic variation does not affect the type of a constraint. In an LOT model, this means that while the weight of a constraint will vary from language to language, the constraint type will stay the same. In other words, a soft constraint should receive a low weight and trigger context effects and developmental optionality across languages, while a hard constraint is crosslinguistically associated with a high constraint weight and immune to contextual variation and developmental optionality.

LOT's predictions about language development and crosslinguistic variation still remain to be tested.

5 Conclusions

In this paper, we gave on overview of the literature on degrees of grammaticality in syntax. Based on a series of recent experimental findings, we argued that gradient data make it possible to distinguish two kinds of linguistic constraints, viz., soft and hard ones. Both types of constraints are subject to constraint ranking and show cumulativity and ganging up effects. However, soft and hard constraints differ with respect to context effects, crosslinguistic variation, and developmental optionality. This means that the notion of constraint type can be operationalized using gradient data: if a constraint violation induces strong unacceptability and fails to show context effects and developmental optionality, then it can be classified as a hard constraint. If a constraint triggers only mild unacceptability and is subject to contextual variation and developmental optionality, then the constraint is soft. The

classification can be verified by investigating the crosslinguistic behavior of the constraint; the type of a constraint (soft or hard) should remain the same across languages.

Throughout the paper, we took a purely formal view of the soft/hard dichotomy. We have not attempted to explain *why* constraints come in two types. This is a key area for further research on gradience; in the following, we will provide a number of speculations on future directions.

The hypothesis underlying the work reported in this paper is that hard constraints are purely structural (i.e., syntactic) in nature, while soft constraints are at the interface between syntax and other domains (i.e., semantics or pragmatics). This seems to account for the facts we illustrated in Section 2. For instance, the soft constraints REF, DEF, and VERB in the extraction data are clearly semantic or pragmatic in nature. The hard constraints AGR, INV, and RES on the other hand, are clearly structural. Similar observation can be made with respect to the data on word order, gapping, and binding presented by Keller and Asudeh (2001); Keller (2000a); Keller and Alexopoulou (2001); Keller (2001).

The hypothesis that soft constraints are at the syntax-pragmatics interface, while hard ones belong to the core of computational syntax, is also consistent with the findings on gradience in language development. Sorace (1999, 2000a) claims that interpretable features (in Chomsky's 1995 terms) are subject to language attrition and to residual optionality in L2 acquisition, whereas non-interpretable features are immune to attrition and don't exhibit residual optionality. Support for this distinction has recently begun to emerge from other areas of language development, such as bilingual first language acquisition (Müller and Hulk, 2001) and specific language impairment (Jakubowicz, 2000; Hamann et al., 1998). Further research is needed to explore—both theoretically and experimentally—these patterns of convergences, and more generally the modes of interaction of different constraints types in natural languages.

References

Aissen, J., 1999. Markedness and subject choice in Optimality Theory. Natural Language and Linguistic Theory 17, 673–711.

Alexopoulou, T., Keller, F., 2002. Resumption and locality: A crosslinguistic experimental study. In: Papers from the 38th Meeting of the Chicago Linguistic Society. Chicago.

Aoun, J., Hornstein, N., Lightfoot, D., Weinberg, A., 1987. Two types of locality. Linguistic Inquiry 18 (4), 537–577.

Asudeh, A., 2001. Linking, optionality, and ambiguity in Marathi. In: Sells, P. (Ed.), Formal and Empirical Issues in Optimality-theoretic Syntax. CSLI Publications, Stanford, CA, pp. 257–312.

Bard, E. G., Robertson, D., Sorace, A., 1996. Magnitude estimation of linguistic acceptability. Language 72 (1), 32–68.

Belletti, A., Rizzi, L., 1988. Psych-verbs and θ-theory. Natural Language and Linguistic Theory 6 (3), 291–352.

Boersma, P., 1997. How we learn variation, optionality, and probability. In: Proceedings of the Institute of Phonetic Sciences. Vol. 21. University of Amsterdam, pp. 43–58.

Boersma, P., 1998. Functional Phonology: Formalizing the Interactions between Articulatory and Perceptual Drives. Holland Academic Graphics, The Hague.

Boersma, P., 2000. Learning a grammar in functional phonology. In: Dekkers, J., van der Leeuw, F., van de Weijer, J. (Eds.), Optimality Theory: Phonology, Syntax, and Acquisition. Oxford Univeristy Press, Oxford, pp. 465–523.

Boersma, P., Hayes, B., 2001. Empirical tests of the gradual learning algorithm. Linguistic Inquiry 32 (1), 45–86.

Boersma, P., Levelt, C., 2000. Gradual constraint-ranking learning algorithm predicts acquisition order. In: Clark, E. V. (Ed.), Proceedings of the 30th Child Language Research Forum. CSLI Publications, Stanford, CA, pp. 299–237.

Bresnan, J., Dingare, S., Manning, C., 2001. Soft constraints mirror hard constraints: Voice and person in English and Lummi. In: Butt, M., Holloway King, T. (Eds.), Proceedings of the 6th Conference on Lexical-Functional Grammar. CSLI Publications Online, Stanford, CA, pp. 13–32.

Cedergren, H. J., Sankoff, D., 1974. Variable rules: Performance as a statistical reflection of competence. Language 50 (2), 333–355.

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

Chomsky, N., 1964. Degrees of grammaticalness. In: Fodor, J. A., Katz, J. J. (Eds.), The Structure of Language: Readings in the Philosophy of Language. Prentice-Hall, Englewood Cliffs, NJ, pp. 384–389.

Chomsky, N., 1975. The Logical Structure of Linguistic Theory. Plenum Press, New York.

Chomsky, N., 1995. The Minimalist Program. MIT Press, Cambridge, MA.

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

Crain, S., Fodor, J. D., 1987. Sentence matching and overgeneration. Cognition 26, 123–169.

Dingare, S., 2001. The effect of feature hierarchies on frequencies of passivization in English. Master's thesis, Stanford University.

Forster, K. I., Stevenson, B. J., 1987. Sentence matching and well-formedness. Cognition 26, 171–186.

Freedman, S. E., Forster, K. I., 1985. The psychological status of overgenerated sentences. Cognition 19, 101–131.

Fukui, N., 1993. Parameters and optionality. Linguistic Inquiry 24, 399–420.

Gordon, P. C., Hendrick, R., 1997. Intuitive knowledge of linguistic co-reference. Cognition 62, 325–370.

Grimshaw, J., Samek-Lodovici, V., 1998. Optimal subjects and subject universals.

In: Barbosa, P., Fox, D., Hagstrom, P., McGinnis, M., Pesetsky, D. (Eds.), Is the Best Good Enough? Optimality and Competition in Syntax. MIT Press and MIT Working Papers in Linguistics, Cambridge, MA, pp. 193–220.

Guy, G. R., 1997. Violable is variable: Optimality Theory and linguistic variation. Language Variation and Change 9, 333–347.

Guy, G. R., Boberg, C., 1997. Inherent variability and the obligatory contour principle. Language Variation and Change 9, 149–164.

Haegeman, L., 1994. Introduction to Government and Binding Theory, 2nd Edition. Basil Blackwell, Oxford.

Hamann, C., Penner, Z., Lindner, K., 1998. German impaired grammar: The clause structure revisited. Language Acquisition 7, 193–426.

Hayes, B. P., 1997. Gradient well-formedness in Optimality Theory, unpubl. handout, Department of Linguistics, University of California, Los Angeles.

Hayes, B. P., 2000. Gradient well-formedness in Optimality Theory. In: Dekkers, J., van der Leeuw, F., van de Weijer, J. (Eds.), Optimality Theory: Phonology, Syntax, and Acquisition. Oxford Univeristy Press, Oxford, pp. 88–120.

Hayes, B. P., MacEachern, M., 1998. Folk verse form in English. Language 74 (3), 473–507.

Höhle, T. N., 1982. Explikationen für "normale Betonung" und "normale Wortstellung". In: Abraham, W. (Ed.), Satzglieder im Deutschen: Vorschläge zur syntaktischen, semantischen und pragmatischen Fundierung. Gunter Narr, Tübingen, pp. 75–153.

Jacobs, J., 1988. Probleme der freien Wortstellung im Deutschen. In: Rosengren, I. (Ed.), Sprache und Pragmatik. Vol. 5 of Working Papers. Department of German, Lund University, pp. 8–37.

Jakubowicz, C., 2000. Functional categories in (ab)normal language acquisition, unpubl. ms., Université de Paris 5.

Keller, F., 1998. Gradient grammaticality as an effect of selective constraint reranking. In: Gruber, M. C., Higgins, D., Olson, K. S., Wysocki, T. (Eds.), Papers from the 34th Meeting of the Chicago Linguistic Society. Vol. 2: The Panels. Chicago, pp. 95–109.

Keller, F., 2000a. Evaluating competition-based models of word order. In: Gleitman, L. R., Joshi, A. K. (Eds.), Proceedings of the 22nd Annual Conference of the Cognitive Science Society. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 747–752.

Keller, F., 2000b. Gradience in grammar: Experimental and computational aspects of degrees of grammaticality. Ph.D. thesis, University of Edinburgh.

Keller, F., 2001. Experimental evidence for constraint competition in gapping constructions. In: Müller, G., Sternefeld, W. (Eds.), Competition in Syntax. Mouton de Gruyter, Berlin, pp. 211–248.

Keller, F., Alexopoulou, T., 2001. Phonology competes with syntax: Experimental evidence for the interaction of word order and accent placement in the realization of information structure. Cognition 79 (3), 301–372.

Keller, F., Asudeh, A., 2001. Constraints on linguistic coreference: Structural vs. pragmatic factors. In: Moore, J. D., Stenning, K. (Eds.), Proceedings of the 23rd

Annual Conference of the Cognitive Science Society. Lawrence Erlbaum Associates, Mahwah, NJ, pp. 483–488.

Keller, F., Asudeh, A., 2002. Probabilistic learning algorithms and Optimality Theory. Linguistic Inquiry 33 (2), 225–244.

Keller, F., Sorace, A., 2003. Gradient auxiliary selection and impersonal passivization in German: An experimental investigation. Journal of Linguistics 39 (1), 57–108.

Koontz-Garboden, A., 2001. A stochastic OT approach to word order variation in Korlai Portuguese. In: Andronis, M., Ball, C., Elston, H., Neuvel, S. (Eds.), Papers from the 37th Meeting of the Chicago Linguistic Society. Vol. 1: The Main Session. Chicago, pp. 347–362.

Kuno, S., 1976. Gapping: A functional analysis. Linguistic Inquiry 7, 300–318.

Labov, W., 1969. Contraction, deletion, and inherent variability of the English copula. Language 45 (4), 715–762.

Lasnik, H., Saito, M., 1984. On the nature of proper government. Linguistic Inquiry 15 (2), 235–289.

Legendre, G., Sorace, A., 2003. Auxiliaires et intransitivité en français et dans les langues romanes. In: Godard, D. (Ed.), Les langues romanes; problèmes de la phrase simple. Éditions du CNRS, Paris, pp. 185–234.

Legendre, G., Sorace, A., Smolensky, P., to appear. The Optimality Theory—Harmonic Grammar connection. In: Smolensky, P., Legendre, G. (Eds.), The Harmonic Mind: From Neural Computation to Optimality-theoretic Grammar. Blackwell, Oxford.

Lenerz, J., 1977. Zur Abfolge nominaler Satzglieder im Deutschen. Stauffenburg, Tübingen.

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

McDaniel, D., Cowart, W., 1999. Experimental evidence of a minimalist account of English resumptive pronouns. Cognition 70, B15–B24.

McElree, B., Griffith, T., 1995. Syntactic and thematic processing in sentence comprehension: Evidence for a temporal dissociation. Journal of Experimental Psychology: Learning, Memory and Cognition 21, 134–157.

Müller, G., 1999a. Optimality, markedness, and word order in German. Linguistics 37 (5), 777–818.

Müller, G., 1999b. Optionality in optimality-theoretic syntax. GLOT International 4 (5), 3–8.

Müller, N., Hulk, A., 2001. Crosslinguistic influence in bilingual language acquisition: Italian and French as recipient languages. Bilingualism: Language and Cognition 4 (1), 1–22.

Pechmann, T., Uszkoreit, H., Engelkamp, J., Zerbst, D., 1994. Word order in the German middle field: Linguistic theory and psycholinguistic evidence. CLAUS Report 43, Department of Computational Linguistics, Saarland University.

Prince, A., Smolensky, P., 1993. Optimality Theory: Constraint interaction in generative grammar. Technical Report 2, Center for Cognitive Science, Rutgers University.

Robertson, D., Sorace, A., 1999. Losing the V2 constraint. In: Klein, E. C., Marto-hardjono, G. (Eds.), The Development of Second Language Grammars: A Generative Approach. John Benjamins, Amsterdam, pp. 317–161.

Scheepers, C., 1997. Menschliche Satzverarbeitung: Syntaktische und thematische Aspekte der Wortstellung im Deutschen. Ph.D. thesis, University of Freiburg.

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

Sorace, A., 1992. Lexical conditions on syntactic knowledge: Auxiliary selection in native and non-native grammars of Italian. Ph.D. thesis, University of Edinburgh.

Sorace, A., 1993a. Incomplete vs. divergent representations of unaccusativity in non-native grammars of Italian. Second Language Research 9, 22–47.

Sorace, A., 1993b. Unaccusativity and auxiliary choice in non-native grammars of Italian and French: Asymmetries and predictable indeterminacy. Journal of French Language Studies 3, 71–93.

Sorace, A., 1999. Initial states, end-states and residual optionality in L2 acquisition. In: Greenhill, A., Littlefield, H., Taro, C. (Eds.), Proceedings of the 23rd Annual Boston University Conference on Language Development. Cascadilla Press, Somerville, MA, pp. 666–674.

Sorace, A., 2000a. Differential effects of attrition in the L1 syntax of near-native L2 speakers. In: Howell, S. C., Fish, S. A., Keith-Lucas, T. (Eds.), Proceedings of the 24th Annual Boston University Conference on Language Development. Cascadilla Press, Somerville, MA, pp. 719–725.

Sorace, A., 2000b. Gradients in auxiliary selection with intransitive verbs. Language 76 (4), 859–890.

Sorace, A., 2000c. Syntactic optionality in non-native grammars. Second Language Research 16 (2), 93–102.

Sorace, A., 2003. Near-nativeness. In: Doughty, C., Long, M. (Eds.), Handbook of Second Language Acquisition. Blackwell, Oxford, pp. 130–152.

Sorace, A., Cennamo, M., 2000. Aspectual constraints on auxiliary choice in Paduan, unpubl. ms., University of Edinburgh and University of Naples.

Sorace, A., Vonk, W., forthcoming. Gradient effects of unaccusativity in Dutch, Department of Theoretical and Applied Linguistics, University of Edinburgh and Max Planck Institute for Psycholinguistics, Nijmegen.

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

Tesar, B., Smolensky, P., 1998. Learnability in Optimality Theory. Linguistic Inquiry 29 (2), 229–268.

Uszkoreit, H., 1987. Word Order and Constituent Structure in German. CSLI Publications, Stanford, CA.