

Constraints in Language Processing: Do Grammars Count?

Marieke van der Feen

Department of Artificial
Intelligence,
University of Groningen,
Grote Kruisstraat 2/1,
9712 TS Groningen,
The Netherlands

mvdfeen@ai.rug.nl

Petra Hendriks

Center for Language and
Cognition Groningen,
University of Groningen,
P.O. Box 716,
9700 AS Groningen,
The Netherlands

p.hendriks@rug.nl

John Hoeks

Center for Language and
Cognition Groningen,
University of Groningen
P.O. Box 716,
9700 AS Groningen,
The Netherlands

j.c.j.hoeks@rug.nl

Abstract

One of the central assumptions of Optimality Theory is the hypothesis of strict domination among constraints. A few studies have suggested that this hypothesis is too strong and should be abandoned in favor of a weaker cumulativity hypothesis. If this suggestion is correct, we should be able to find evidence for cumulativity in the comprehension of Gapping sentences, which lack explicit syntactic clues in the form of the presence of a finite verb. On the basis of a comparison between several computational models of constraint evaluation, we conclude that the comprehension of Gapping sentences does not yield compelling evidence against the strict domination hypothesis.

1 Introduction

A linguistic framework which has gained a considerable amount of attention in recent years is Optimality Theory (Prince and Smolensky, 1993/2004). Optimality Theory (henceforth OT) is not only used for analyzing and explaining linguistic phenomena in the domain of phonology, but also in the domains of morphology, syntax, semantics and pragmatics. In contrast to more traditional linguistic frameworks, OT assumes grammatical constraints to be violable. Because constraints are formulated in such a way that they are maximally general (and perhaps even universal across languages), these constraints may conflict. To resolve conflicts among constraints, constraints are assumed to differ in

strength. It is better to violate a weaker constraint than it is to violate a stronger constraint. The grammatical structure is the one that violates the least highly ranked (i.e., strong) constraints.

A fundamental property of OT is the principle of strict domination. This means that each constraint has complete priority over all constraints ranked lower in the constraint hierarchy. A number of recent studies, however, have called into question this fundamental property of OT. Keller (2001) argues that constraint violations must be cumulative to account for the pattern of relative acceptability with respect to the phenomenon of Gapping. Jäger and Rosenbach (to appear) draw a similar conclusion on the basis of the observed variation with respect to the English genitive (*the king's palace* versus *the palace of the king*).

In this study, we focus on the linguistic phenomenon of Gapping. The central question is whether the comprehension of Gapping sentences provides evidence in favor of cumulativity of constraint violations. In section 2, we introduce the phenomenon and discuss the possibility of an OT model of Gapping. In section 3, we consider different kinds of cumulativity. Section 4 discusses the way we modeled four different evaluation algorithms based on these kinds of cumulativity. A comparison between our computational models of constraint evaluation in section 5 suggests that the comprehension of Gapping does not provide compelling evidence for abandoning the strict domination hypothesis.

2 Gapping

Gapping is a grammatical operation that deletes certain subconstituents in the second conjunct of a coordinate structure, as in (1):

- (1) Some ate beans, and others rice.

The deleted material always includes the finite verb, but may also include further constituents such as the direct object. As a result, it may not always be possible to uniquely identify which elements were left out. As an example, consider the following sentence:

- (2) John greeted Paul yesterday and George today.

This sentence is ambiguous between reading (3), where first John greeted Paul, and then John greeted George, and reading (4), where first John greeted Paul, and then George greeted Paul.

- (3) John greeted Paul yesterday and ~~John greeted~~ George today.
 (4) John greeted Paul yesterday and George ~~greeted Paul~~ today.

The reading in (3) is traditionally analyzed as resulting from the operation of conjunction reduction, whereas the reading in (4) is analyzed as resulting from Gapping of the finite verb and the direct object.

2.1 Functional constraints on Gapping

Based on previous work on Gapping, Kuno (1976) notes that several non-syntactic factors affect the acceptability and interpretation of Gapping. One of these factors is the distance between the remnants in the second conjunct and their counterparts in the first conjunct:

- (5) *The Minimal Distance Principle:*
 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.

According to this principle, interpretation (3) should be preferred for sentence (2) because it is more preferable to couple *George* in the second conjunct to the direct object *Paul* in the first conjunct, than to the more distant subject *John*. This preference is confirmed by experimental evidence (Carlson, 2001). A further principle about Gapping is that the deleted material has to represent contextually given information, whereas the remnants in the second conjunct have to consti-

tute new information. This is captured in the following principle:

- (6) *The Functional Sentence Perspective (FSP) Principle of Gapping:*
 a. Constituents deleted by Gapping must be contextually known. On the other hand, the two constituents left behind by Gapping necessarily represent new information and, therefore, must be paired with constituents in the first conjunct that represent new information.
 b. It is generally the case that the closer a given constituent is to sentence-final position, the newer the information it represents in the sentence.
 c. Constituents that are clearly marked for nonanaphoricity necessarily represent new information in violation of (b). Similarly, constituents that appear closest to sentence-final position necessarily represent old information (in violation of (b)) if coreferential constituents appear in the corresponding position in the preceding discourse.

This principle explains the observation that in a suitable context, interpretation (4) can become the preferred interpretation for (2) (but see Hoeks et al. (2006) for experimental evidence that in addition to context also prosody has to be in accordance with a Gapping reading to make this reading the preferred reading):

- (7) When did John and George greet Paul?
 John greeted Paul yesterday and George ~~greeted Paul~~ today.

In this example, John, Paul and George are all contextually introduced. But only John and George are subjects in the context sentence and hence can be interpreted as contrastive topics in the target sentence. Contrast has a similar effect as newness. Because of this effect of context, the Gapping reading can become the preferred reading for (2). Two further principles proposed by Kuno are (8) and (9).

- (8) *The Tendency for Subject-Predicate Interpretation:*
 When Gapping leaves an NP and a VP behind, the two constituents are readily interpreted as constituting a sentential pattern, with the NP representing the subject of the VP.

(9) *The Requirement for Simplex-Sentential Relationship:*

The two constituents left over by Gapping are most readily interpretable as entering into a simplex-sentential relationship. The intelligibility of gapped sentences declines drastically if there is no such relationship between the two constituents.

The principle in (8) is meant to account for a difference in preference with object control verbs versus subject control verbs. The principle in (9) reflects the observation that Gapping cannot leave behind remnants that are part of a subordinate clause. Kuno notes that this latter constraint seems to be the strongest of the four principles, being nearly inviolable, but does not make the interaction between his principles explicit.

2.2 An OT model of Gapping

As Kuno already observes, the FSP Principle seems to be able to override the Minimal Distance Principle. This observation is regarded by Keller (2001) as evidence that Gapping is subject to constraint competition in an optimality theoretic sense. Based on Kuno's principles, Keller develops an OT model of Gapping, which is able to account for the pattern of relative acceptability of Gapping sentences. According to this model, the degree of acceptability of a candidate structure depends on the number and type of re-rankings required to make the structure optimal (Keller, 1998).

Keller's OT model differs from standard OT in a number of ways. Firstly, a distinction is made between soft and hard constraints. Hard constraints cause strong acceptability when violated, while violation of soft constraints causes only mild unacceptability. According to Keller, the Requirement for Simplex-Sentential Relationship is such a hard constraint. The distinction between soft and hard constraints is needed in Keller's model to avoid the problem of overgeneration of acceptability differences.

Secondly, Keller's model assumes that constraint violations are cumulative. According to his model, the degree of unacceptability increases with the number of constraints violated. In standard OT, on the other hand, no number of violations of weaker constraints can override one violation of a stronger constraint, in accordance with the principle of strict domination.

The aim of Keller's OT model is to account for the pattern of relative acceptability of Gapping

sentences. The aim of the present study, on the other hand, is to account for the comprehension of Gapping sentences. Nevertheless, we follow Keller in adopting Kuno's functional principles (reformulated as OT constraints) for our OT model because Kuno's principles are principles of comprehension.

Our model differs from Keller's model in several essential aspects, though. We assume that all constraints are violable, in accordance with the basic assumptions of OT. Because certain strong constraints are not violated by the data under discussion, they simply appear to be inviolable. Keller's second assumption, the assumption that constraint violations are cumulative, is the topic of investigation of this study.

3 Cumulativity of constraint violations

In this section we discuss the different ways OT constraints can interact. In principle, OT constraints can interact in an unrestricted way, or in one of several more or less restricted ways.

3.1 Unrestricted constraint interaction

OT as a linguistic theory is derived from Harmonic Grammar (Legendre et al., 1990). In Harmonic Grammar (henceforth HG), each constraint is associated with a positive or negative numerical weight value. For each candidate, a so-called Harmony value is calculated by summing the numerically weighted constraints. From the set of candidates, the candidate with the highest Harmony value is selected as the optimal candidate. Consequently, the interaction among constraints in HG is cumulative. Each constraint violation lowers the Harmony value of the candidate. This type of constraint interaction is essentially unrestricted.

To account for natural language interpretation, however, unrestricted cumulativity is too liberal, as is shown by OT analyses of other phenomena. With respect to Gapping, if Kuno and Keller are right, no amount of violations on weaker constraints of an interpretation satisfying Simplex-Sentential Relationship can make an interpretation violating Simplex-Sentential Relationship the preferred one:

- (10) Who did John promise to examine who?
John promised Paul to examine George,
and Ringo Bob.

If Simplex-Sentential Relationship indeed is a strong constraint, (10) should only mean that

Ringo promised to examine Bob (satisfying Simplex-Sentential Relationship but violating the Minimal Distance Principle and the FSP), and never that John promised to examine Bob (violating Simplex-Sentential Relationship).

For the analysis of natural language, therefore, but also for the establishment of cross-linguistic generalizations (see Legendre et al., 2006), we seem to require a type of constraint interaction which is more restricted than simple numerical constraint weighting.

3.2 Restricted constraint interaction

In this section we discuss four ways to restrict constraint interaction: (1) strict domination, (2) local restricted cumulativity, (3) global restricted cumulativity, and (4) Keller’s counting cumulativity.

| | A | B | C | D |
|---------------|----|---|---|---|
| ☞ Candidate 1 | | * | * | * |
| Candidate 2 | *! | | | |

Tableau 1: Strict domination

Strict domination is illustrated in tableau 1. The constraints are ordered from left to right in the top row in order of descending strength. Under strict domination, no number of violations of the weaker constraints B, C and D is able to override a violation of the strongest constraint A.

| | A | B | C | D |
|---------------|---|---|----|----|
| Candidate 1 | | | *! | *! |
| ☞ Candidate 2 | | * | | |

Tableau 2: Local restricted cumulativity

Tableau 2 illustrates local restricted cumulativity. When the weaker constraints C and D are simultaneously violated, their joint effect can be stronger than their linear sum. As a result, together they are able to override the immediately dominating constraint B. This type of cumulativity is similar to the effects of local conjunction. The result is a conjoined constraint C&D, which is ranked immediately above constraint B in the hierarchy.

| | A | B | C | D |
|---------------|---|---|----|----|
| Candidate 1 | | | *! | *! |
| ☞ Candidate 2 | * | | | |

Tableau 3: Global restricted cumulativity

An illustration of global restricted cumulativity is given in tableau 3. In this case, the weaker constraints C and D together are able to override a stronger, but not necessarily immediately dominating, constraint A. Again, this type of cumulativity is similar to the effects of local conjunction. The result is a conjoined constraint C&D, which is ranked anywhere above C and D in the hierarchy.

| | A | B | C | D |
|---------------|---|---|----|---|
| Candidate 1 | | * | *! | * |
| ☞ Candidate 2 | * | | | |

Tableau 4: Keller’s counting cumulativity

Keller’s counting cumulativity is illustrated in tableau 4. For Keller’s cumulativity, the hierarchical relation between the constraints is irrelevant. The candidate with the fewest constraint violations is always optimal. In Keller’s model, constraint violations are assumed to result in a gradient pattern. The more constraints are violated by a given Gapping construction, the less acceptable the construction is predicted to be. Of course, this type of cumulativity will greatly overgenerate in production as well as in comprehension if every constraint violation counts as an equally serious violation. For this reason, a system employing this type of cumulativity must make a distinction between soft and hard constraints. Hard constraints cause strong unacceptability. This extra assumption serves to restrict the overgenerating power of this type of cumulativity.

The four types of cumulativity discussed here differ in the amount of freedom they allow. Strict domination is the most restricted type of constraint interaction, local restricted cumulativity the one but most restricted type, global restricted cumulativity the two but most restricted type, and Keller’s cumulativity the least restricted type. As a result, strict domination yields the strongest hypothesis, and Keller’s cumulativity the weakest hypothesis. The question we set out to answer in the next section is how strongly constraint interaction must be restricted to account for the comprehension of Gapping sentences.

4 Testing the evaluation algorithms

To test the predictions of the four types of cumulativity discussed in the previous section, a computer model was developed in Prolog. The input

to the model is a Gapping sentence in Dutch. The first conjunct is manually parsed. Information about the givenness of its constituents, the selectional restrictions of the main verb of the first conjunct, and featural information for all NPs is added. The output of the model is formed by the possible couplings of constituents in the second conjunct with constituents in the first conjunct. In addition, for each possible coupling the constraint profile is given. For each possible coupling, the model also gives a reconstruction of the second conjunct by placing the constituents from the second conjunct in the position of the constituents they are coupled with in the first conjunct.

4.1 Constraint ranking

The constraints implemented in the model were Kuno's principles, reformulated as OT constraints, augmented with constraints on parallelism (cf. Carlson, 2001), thematic selection (Hoeks and Hendriks, 2005) and word order (Lamers and de Hoop, 2004). The constraint ranking used is:

- (11) Categorical Parallelism >> Simplex-Sentential Relationship >> FSP >> Thematic Selection >> Subject Precedes Object >> Syntactic Parallelism >> Minimal Distance >> Subject-Predicate Interpretation >> Featural Parallelism

The constraint Categorical Parallelism is added to ensure that constituents are coupled with constituents of the same syntactic category only. It prevents, for example, that in (2) *today* is coupled with *Paul*. Thematic Selection expresses the selectional restrictions verbs may impose on their arguments. For example, the verb *bake* requires an inanimate object, the verb *introduce* requires an animate object, and the verb *take* can combine with either an animate or an inanimate object (see section 4.3). The constraint Thematic Selection is violated if the candidate interpretation does not satisfy these selectional restrictions, for example if the object of the verb *bake* is animate. According to the constraint Subject Precedes Object, the subject must linearly precede the object. Syntactic Parallelism requires the two conjuncts to have the same syntactic structure. The constraint Featural Parallelism, finally, promotes the coupling of constituents which share features such as animacy, definiteness, number and gender. The ranking of these constraints was determined on the basis of the literature (Carlson,

2001; Kuno, 1976) and via comparison of relevant sentences and their meanings.

4.2 Computational considerations

The different types of cumulativity were computationally modeled as different ways of evaluating the constraint profiles.

Strict domination can be modeled as numerical weighting with exponential weights.

Local restricted cumulativity can be modeled as numerical weighting as well, if the weights are chosen in such a way that the sum of two adjacent constraints is larger than the weight of the directly dominating constraint. This is the case if, for example, B is 0.50, C is 0.26, and D is 0.25 in tableau 2. In our model, local restricted cumulativity only applies to the constraints Thematic Selection, Subject Precedes Object and Syntactic Parallelism, and allows the constraints Subject Precedes Object and Syntactic Parallelism together to override the directly dominating constraint Thematic Selection.

Global restricted cumulativity, on the other hand, cannot be captured straightforwardly in a system with weight values. To implement this evaluation method, therefore, we made explicit use of constraint conjunction. The newly formed conjoined constraint C&D was located in the hierarchy somewhere above its constituting constraints C and D. Because violation of this conjoined constraint is dependent on the violation of each of the constituting constraints, the new constraint can only be evaluated in a second round of evaluation after all other constraints have been evaluated. This is an unfortunate complication of our implementation. Legendre et al. (2006: 352) show that this type of cumulativity can be implemented with weight values if constraint conjunction is assumed to involve a superlinear combination of weights (through summation as well as multiplication). In our model, only the constraints Minimal Distance and Subject-Predicate Interpretation were allowed to conjoin. The resulting conjoined constraint was located above Categorical Parallelism in the hierarchy.

For the fourth method of evaluation, Keller's counting cumulativity, simply counting the number of constraint violations suffices. By applying one of these four evaluation algorithms, the computational model yields an optimal interpretation for each combination of input and evaluation algorithm.

4.3 Input sentences

To test the four evaluation algorithms, we fed the model three types of input: (i) 10 Gapping sentences taken from a corpus, (ii) test sentences taken from all five conditions of Carlson’s (2001) study on Gapping, and (iii) 15 hand-crafted sentences.

The Eindhoven corpus (uit den Boogaart, 1975) is an annotated corpus of Dutch written text of about 750 000 words. We scanned the corpus for suitable Gapping sentences, which had to occur unembedded, contain an overt conjunction, and should not involve other deletion operations as well. Unfortunately, we only found 10 such Gapping sentences in the corpus, presumably because Gapping is quite rare. For all ten sentences, all evaluation methods produced the same outputs. Nine out of the ten optimal interpretations did not violate any of the constraints. One sentence involved a constraint violation by all models, namely a violation of the constraint Featural Parallelism:

- (12) Groep 1 trok de arm na vijftien minuten uit de testkamer, en groep 4 na een uur.
Group 1 pulled the arm after fifteen minutes from the test room and group 4 after an hour.

The most plausible interpretation of this sentence is the interpretation that group 4 pulled the arm from the test room after an hour. The interpretation selected by all evaluation methods, however, was that group 1 pulled group 4 from the test room after an hour, thus satisfying Minimal Distance but violating Featural Parallelism. It may be that the strong parallelism between *group 1* and *group 4* sets up a contrast which evokes the constraint FSP even in the absence of an explicit linguistic context. If this is true, Minimal Distance must be violated in order to satisfy FSP.

We also fed the models test sentences taken from Carlson’s (2001) written questionnaire. Carlson studied the interaction between Thematic Selection, Featural Parallelism and Minimal Distance by varying verb type (see the discussion of Thematic Selection in section 4.1) and properties of the noun phrases. She distinguished five conditions: the Bake A condition (*Alice bakes cakes for tourists and Caroline for her family*), the Bake B condition (*Alice bakes cakes for tourists and brownies for her family*), the Take A condition (*Josh visited the office during the vacation and Sarah during the week*), the

Take B condition (*Josh visited Marjorie during the vacation and Sarah during the week*) and the Introduce condition (*Dan amazed the judges with his talent and James with his musicality*).

The four evaluation algorithms behaved exactly the same on all five conditions of Carlson because none of Carlson’s sentences involves a simultaneous violation of Subject Precedes Object and Syntactic Parallelism (which would give rise to local restricted cumulativity in our model) or a simultaneous violation of Minimal Distance and Subject-Predicate Interpretation (which would give rise to global restricted cumulativity in our model). As a result, all models yielded a 100% Gapping response for Carlson’s Bake A condition (compared to Carlson’s subjects 81%) because for all models a violation of Thematic Selection is more serious than a violation of Minimal Distance. Furthermore, all models yielded a 100% non-Gapping response for her Bake B condition (compared to Carlson’s subjects 97%) because a Gapping response violates Thematic Selection, Minimal Distance and Featural Parallelism whereas a non-Gapping response satisfies all three constraints. Finally, all models yielded a 100% non-Gapping response for Carlson’s Take A condition (compared to Carlson’s subjects 60%), her Take B condition (compared to Carlson’s subjects 96%) and her Introduce condition (compared to Carlson’s subjects 79%) because for all models a violation of Minimal Distance is more serious than a violation of Featural Parallelism, given the constraint ranking in (11).

So all models correctly predicted the interpretational preferences found in Carlson’s experiment. However, subjects’ percentages of non-Gapping responses on the Take A, Take B and Introduce condition varied considerably. This variation seems to be due to differences between the features of the NPs involved. In particular, in the Take A condition the feature animacy played a role, which seems to have a stronger effect than the other grammatical features that were manipulated. However, our constraint Featural Parallelism does not distinguish between animacy and other grammatical features. Moreover, our OT model is unable to capture the gradience that seems to result from the interaction between features.

4.4 Generating different predictions

Because the four evaluation algorithms behaved identically on all sentences taken from the corpus as well as on all sentences types from Carlson’s

study, we had to construct sentences on the basis of expected constraint violations in order to generate different predictions for the four evaluation algorithms. The following sentence is predicted to distinguish between strict domination and local restricted cumulativity:

- (13) John picked a rose, and a tulip Paul.

If hearers interpret this sentence as meaning that a tulip picked Paul, they will have violated the stronger constraint Thematic Selection in order to satisfy the two weaker constraints Subject Precedes Object and Syntactic Parallelism. This then would constitute evidence for local restricted cumulativity. If, on the other hand, hearers interpret this sentence as meaning that Paul picked a tulip, then this is evidence for strict domination. Sentence (14) distinguishes between strict domination and global restricted cumulativity:

- (14) John asked him to get Paul, and George to bring Ringo.

Because *him* is a pronoun, it counts as given for evaluating the constraint FSP. If hearers interpret this sentence as meaning that John asked George to bring Ringo, they will have violated the stronger constraint FSP in order to satisfy the weaker constraints Minimal Distance and Subject-Predicate Interpretation. Because FSP does not immediately dominate the weaker constraints, this would be evidence for global restricted cumulativity. To distinguish between strict domination and Keller's counting cumulativity, consider the following sentence:

- (15) The children promised John to stop, and the neighbors to continue.

If hearers interpret this sentence as meaning that the neighbors promised John to continue, they violate the single stronger constraint Minimal Distance in favor of satisfaction of the two weaker constraints Subject-Predicate Interpretation and Featural Parallelism. Because these constraints would all be considered soft constraints according to Keller's distinction between hard and soft constraints, Keller's counting cumulativity predicts that this interpretation is preferred. The strict domination hypothesis, on the other hand, predicts that the interpretation is preferred according to which the children promised the neighbors to continue, since it is more important

to satisfy the stronger constraint Minimal Distance than any number of weaker constraints.

5 Results and discussion

For all Gapping sentences occurring in the Eindhoven corpus and all Gapping sentences taken from the written part of Carlson's psycholinguistic study, the four evaluation algorithms yielded identical results. These sentences therefore do not shed any light on the central question of this study, namely whether the strict domination hypothesis should be abandoned in favor of a weaker cumulativity hypothesis.

To determine which evaluation algorithm models the way comprehenders process language best, we must look at the interpretations of sentences such as (13), (14) and (15). We presented 10 participants with a written questionnaire, which included 15 sentences distinguishing between the four evaluation algorithms. The reader is referred to van der Feen (2005) for the complete list of sentences. The results show that there does not seem to be a clear preference in interpretation for sentences such as (13), leaving the distinction between strict domination and local restricted cumulativity undecided. For sentences such as (14), on the other hand, there seems to be a clear preference for the reading supported by global restricted cumulativity. Sentences such as (15), finally, show no effects at all of Keller's counting cumulativity. For only one sentence only one subject preferred the interpretation according to which the neighbors promised John to continue, which favors the strict domination hypothesis and goes against Keller's cumulativity algorithm. This suggests that constraints on comprehension may be different from the principles governing acceptability judgments. Boersma (2004) argues that the paralinguistic task of providing acceptability judgments involves comprehension, but under a reverse mapping between meaning and form. An alternative view is that acceptability judgments involve a mapping from the given form to its optimal meaning ('what do I think the sentence means?'), followed by a mapping from that meaning to the optimal form for that meaning ('how would I express that meaning?'), thus involving principles of comprehension as well as production.

To conclude, there seems to be a slight indication of global restricted cumulativity in the comprehension of Gapping, but further study with a larger pool of subjects is required to confirm these initial findings.

However, a few remarks are in place here. First, note that for hearers to prefer the interpretation that Paul picked a tulip for (13), the hearer has to find some motivation in the linguistic context of the utterance for why the speaker chose to put the object first. In the absence of such a context supporting a non-canonical word order, the reading that Paul picked a tulip might be dis-preferred anyway.

Also in sentence (14), context seems to play a crucial role. Although in general pronouns may be used to refer to given material, in certain contexts pronouns can be emphatically stressed. If the pronoun in (14) is stressed, it is much easier to couple *George* to *him* to obtain the reading that John asked George to bring Ringo. This effect of context and prosody may have been the main reason for the observed preferences.

6 Conclusion

A central principle of Optimality Theory is the hypothesis of strict domination among constraints. In this paper we investigated whether this hypothesis should be abandoned in favor of the weaker hypothesis that constraint violations are cumulative. Studying the effects of four different evaluation algorithms (three of which display some kind of cumulativity) on the comprehension of Gapping sentences, we found a slight indication of cumulativity effects. However, these effects are likely to disappear if the context and prosodic structure of the utterance are taken into account.

Acknowledgments

The authors thank three anonymous reviewers for their useful comments and suggestions. This research was funded by grant # 015.001.103 from NWO, awarded to Petra Hendriks.

References

Paul Boersma. 2004. *A stochastic OT account of paralinguistic tasks such as grammaticality and prototypicality judgments*. Unpublished manuscript, University of Amsterdam. Rutgers Optimality Archive #648.

Katy Carlson. 2001. The Effects of Parallelism and Prosody in the Processing of Gapping Structures. *Language and Speech*, 44(1):1-26.

John Hoeks, Petra Hendriks, and Louisa Zijlstra. 2006. The Predominance of Nonstructural Factors in the Processing of Gapping Sentences. In: R. Sun and N. Miyake (eds.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society*.

John Hoeks, and Petra Hendriks. 2005. Optimality Theory and human sentence processing: The case of coordination. In: B.G. Bara, L. Barsalou, and M. Bucciarelli (eds.), *Proceedings of the 27th Annual Meeting of the Cognitive Science Society*, Erlbaum, Mahwah, NJ, pp. 959-964.

Gerhard Jäger, and Anette Rosenbach. To appear. The winner takes it all - almost. Cumulativity in grammatical variation. *Linguistics*.

Frank Keller. 1998. Gradient Grammaticality as an Effect of Selective Constraint Re-ranking. In: M.C. Gruber, D. Higgins, K.S. Olson, and T. Wysocki (eds.) *Papers from the 34th Meeting of the Chicago Linguistic Society*. Vol. 2: The Panels, Chicago, pp. 95-109.

Frank Keller. 2001. Experimental Evidence for Constraint Competition in Gapping Constructions. In: G. Müller and W. Sternefeld (eds.), *Competition in Syntax*, Mouton de Gruyter, Berlin, pp. 211-248.

Susumo Kuno. 1976. Gapping: A Functional Analysis. *Linguistic Inquiry*, 7:300-318.

Monique Lamers, and Helen de Hoop. 2004. The role of animacy information in human sentence processing captured in four conflicting constraints. In: H. Christiansen, P. Rossen Skadhauge, and J. Villadsen (eds.), *Constraint Solving and Language Processing*. Workshop proceedings, Roskilde Department of Computer Science, Roskilde University, pp. 102-113.

Géraldine Legendre, Yoshiro Miyata, and Paul Smolensky. 1990. Harmonic Grammar - A formal multi-level theory of linguistic well-formedness: An application. In: *Proceedings of the Twelfth Annual Conference of the Cognitive Science Society*, Erlbaum, Cambridge, MA, pp. 388-395.

Géraldine Legendre, Antonella Sorace, and Paul Smolensky. 2006. The Optimality Theory - Harmonic Grammar Connection. In: P. Smolensky and G. Legendre (eds.), *The Harmonic Mind*, Vol. 2, MIT Press, Cambridge, MA, pp. 339-402.

Alan Prince, and Paul Smolensky. 2004. *Optimality Theory: Constraint interaction in generative grammar*. Oxford, Blackwell. Previously distributed as Technical Report RuCCSTR-2, New Brunswick NJ, Rutgers Center for Cognitive Science, Rutgers University, 1993.

P.C. Uit den Boogaart. 1975. *Woordfrequenties in geschreven en gesproken Nederlands*. Werkgroep Frequentie-onderzoek van het Nederlands. Oosthoek, Scheltema & Holkema, Utrecht.

Marieke Van der Feen. 2005. *Do rules add up? A study of the application of Optimality Theory to the interpretation of gapping*. MSc Thesis Artificial Intelligence, University of Groningen.