

Optimality Theory and Human Sentence Processing: The Case of Coordination

John C. J. Hoeks (j.c.j.hoeks@rug.nl)

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

Petra Hendriks (p.hendriks@rug.nl)

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

Abstract

In line with recent studies we propose a model of human sentence processing that is based on Optimality Theory (OT). Rather than explaining parsing preferences through extralinguistically motivated parsing strategies or frequencies in the hearer's linguistic environment, our model explains these preferences as the intermediate results of the incremental application of our OT grammar. In contrast to most other current OT-approaches to language processing, we use constraints from OT *semantics* rather than from OT *syntax* to explain on-line comprehension. We illustrate the workings of our model by investigating the comprehension of coordination, a phenomenon which is ill-understood from a competence perspective and sparsely investigated from a processing perspective. The psycholinguistic evidence that is currently available strongly suggests that the on-line comprehension of coordinate structures is influenced by constraints from many different information sources: pragmatics, discourse semantics, lexical semantics, and syntax. The competence / performance model we propose is able to formalize this *cross-modular* constraint interaction, and to yield concrete predictions with respect to both intermediate parsing preferences and ultimate interpretations.

Introduction

A fundamental issue in linguistic research is the relation between linguistic theory (competence) and human sentence processing (performance). A quite influential line of research aims at trying to explain human sentence processing through parsing strategies that are not themselves part of the competence, such as, for example, Frazier's (1987a) Minimal Attachment principle (i.e., in case of ambiguity, choose the syntactic structure that is least complex) or her right association principle Late Closure (i.e., in case of ambiguity, attach the new node to the most recently processed constituent). According to this type of approach, structural preferences in processing are the result of general cognitive limitations, such as limitations on working memory or bounds on complexity. Another approach is to view processing preferences as tuning to frequencies in the hearer's linguistic environment (MacDonald, Pearlmutter, & Seidenberg, 1994; Mitchell, Cuetos, Corley, & Brysbaert, 1995). An important assumption underlying both of these approaches is that linguistic competence provides a hearer with a set of possible structures for a given sentence, from which the human sentence processor then selects the preferred structure. However, an alternative approach is to try and

derive processing characteristics from linguistic theory itself (Artstein, 2000; de Hoop & Lamers, to appear; Fanselow, Schlesewsky, Cavar, & Kliegl, 1999; Gibson & Broihier, 1998; Phillips, 1995; Stevenson & Smolensky, 2005; Weinberg, 2001). In its strongest form this approach considers all theoretical constraints to be processing restrictions and vice versa, and it is postulated that this model can be a model of competence (i.e., traditionally called 'grammar') and a model of performance (i.e., traditionally called 'parser') at the same time. It is the aim of this paper to show that such a model is viable, and even necessary, for the understanding of a linguistic phenomenon that otherwise seems to resist an adequate theoretical analysis, namely *coordination*.

Linguistic Theory and Coordination

Linguistic theory is concerned with the syntactic and semantic structure of all sentences and phrases that constitute a given language. However, it is generally assumed that coordinate structures (e.g., "X and Y") are somehow special, and distinct from all other structures, especially in the syntactic domain. This exceptional position for coordination is implicit in works such as Ross (1967) and others, but has recently surfaced in more explicit terms. For example, the syntactic analyses of coordination such as developed by Munn (1987) and by Johannessen (1993) are incompatible with the X-bar schema, which is assumed to underlie all syntactic structure (Borsley, 1994). For one thing, the fact that the categorial makeup of coordinate structures reflects that of the conjuncts cannot be derived from X-bar schema and associated standard mechanisms. That is, if two conjuncts are noun phrases, as in the coordination *cats and dogs*, the entire coordinate structure syntactically behaves as a noun phrase as well. Secondly, it does not follow from standard syntactic mechanisms that the conjuncts must generally be of the same syntactic type or category (i.e., 'coordination of likes'). Thirdly, conjunctions differ from other heads in that they may also take non-maximal projections as their conjunct complements, for example prepositional heads, as in *on or under the table*. So it is impossible to derive the main properties of coordination from standard syntactic assumptions. Consequently, coordination is treated as a special type of structure, generated by special rules that enforce the structural properties typical for coordination. Some syntacticians have suggested that coordination can only be represented by means of non-standard three-dimensional structures (e.g., de

Vries, 2003). Others have even gone so far as to argue that, because of its deviant properties, coordination must be viewed as an extragrammatical parsing procedure rather than a linguistic structure (cf. Kempen, 2004, who suggests to analyze coordination as a grammaticalized form of self-repair). But if coordination must indeed be treated distinctly from regular syntactic structures, it is not clear how the syntactic rules that determine regular structure, and the syntactic rules or procedures that determine coordinate structure can coexist and interact. Moreover, it is not obvious why there would be such a distinction between coordination and other linguistic forms. We will argue that no special status for coordination is necessary in Optimality Theory (OT).

Optimality Theory

OT is in essence a hybrid cognitive architecture, combining rule-governed symbolic processing with parallel subsymbolic, or connectionist, processing. It was originally introduced as a model of linguistic competence, and as such has been very successful in the linguistic domains of phonology (e.g., Prince & Smolensky, 2004), morphology (e.g., McCarthy & Prince, 1993), syntax (e.g., Barbosa et al., 1998), and semantics/pragmatics (e.g., Hendriks & de Hoop, 2001). In OT, inputs are mapped onto outputs by first generating all possible candidates for each input (by the generator function GEN), and then selecting the optimal candidate from among them (through the evaluator function EVAL). GEN has the property of 'freedom of analysis' (Kager, 1999): the number of outputs is *infinite*, and the outputs can have *any* conceivable form (cf. Beaver, 2004). These two features are crucial, because if the number of candidates were finite, or would have a predefined structure, the grammar would have to place restrictions on the candidates *before* the constraints are evaluated. And since the grammar is actually formed by the set of constraints, the grammar would have to apply twice, which is an unwelcome situation. Therefore it is stipulated that there are infinitely many outputs of every conceivable structure. Importantly, the number of output candidates is only *virtually* infinite: at the level of the neural computation that underlies OT models, these candidates are not actually physically represented; they can be said to be present as *possible* outcomes (cf. Smolensky & Legendre, 2005). Evaluation consists of the simultaneous application of a hierarchically ordered set of constraints to the candidate outputs. The constraints differ in strength, and crucially, the strongest constraint has absolute dominance over all the weaker, i.e. lower ranked, constraints. In contrast to many other approaches to linguistics, especially those that are concerned with the nature of the grammar rather than the parser, these constraints are soft, which means that they can be violated by well-formed candidates, but only if violation of a weaker constraint is necessary to satisfy a stronger constraint. Importantly, in the present model, constraints are included that arise from many different sources of information.

Constraints on Coordination

On the basis of the currently available empirical evidence, we will argue that the processing of coordinate structures is

dependent on 1) pragmatic constraints, 2) discourse-semantic constraints, 3) syntactic constraints, and 4) lexical-semantic constraints. Adopting the framework of OT allows us to formalize this *cross-modular* constraint interaction. Most psycholinguistic studies on the comprehension of coordinated structures have focused on the NP- versus S-coordination ambiguity (e.g., Frazier, 1987b; Hoeks, Vonk, & Schriefers, 2002; Hoeks, Hendriks, Vonk, Brown, & Hagoort, in press). For instance, Frazier (1987b) showed that readers prefer NP-coordination over S-coordination in sentences such as (1a) and (1b); slashes indicate how sentences were divided into segments.

- 1a. Joe kissed Ann and / Mary / today. (NP-coordination)
 1b. Joe kissed Ann and / Mary / laughed. (S-coordination)

It was assumed that readers prefer to take the ambiguous NP *Mary* as part of the direct object of *kissed* as in (1a). Consequently, they will run into trouble when reading the final segment of (1b), where the finite verb *laughed* indicates the ambiguous NP is actually the subject of a conjoined sentence. And indeed, Frazier found significantly longer reading times for *laughed* in (1b) than for *today* in (1a). This finding was replicated by Hoeks et al. (in press) who corrected for a number of confounds in Frazier's earlier study. S-coordinated sentences such as (2a) and (2b) were used, the first of which was temporarily ambiguous, whereas the latter served as a control sentence, made unambiguous by inserting a comma after the first object NP (Hoeks et al., 2002). Underlined is the critical verb opened which forces an S-coordination reading, where *the photographer* in (2a) is the subject of a conjoined sentence, which goes against the preferred reading in which *the photographer* is part of the direct object of *embraced* (i.e., conjoined with *the designer*).

- 2a. The model embraced the designer and the photographer opened a bottle of expensive champagne
 2b. The model embraced the designer, and the photographer opened a bottle of expensive champagne.

With these materials it is possible to compare sentences that are identical in sentence meaning, and to compare regions that are identical in length, frequency, and syntactic category. Thus, these studies show that readers prefer to interpret an ambiguous NP as an argument of the first main verb (NP-coordination), instead as the subject of a new clause (S-coordination). As a result, temporarily ambiguous S-coordinations give rise to processing difficulty. What has been ignored, however, is that there is an earlier point at which the sentence is ambiguous: at the connective itself. Before the ambiguous NP is read, the sentence can continue as an NP- or an S-coordination, but also as a *VP-coordination*, as shown in sentence (3).

3. The model embraced the designer and laughed.

And indeed, recent evidence from sentence completion studies has shown that language users strongly prefer to continue a fragment such as (4) as a VP-coordination.

4. The model embraced the designer and ...

In about 86 % of all cases coordinated VPs were produced, as opposed to 9 % NP-coordinations and 5 % S-coordinations (Hoeks et al., 2002, Exp. 1). This outcome suggests that language comprehenders expect the connective to be followed by a VP, not by an NP. Only when the NP is actually presented, and VP-coordination is no longer possible, NP-coordination becomes the preferred structure. This finding provides us with important clues as to which constraints may be necessary to describe the processing of coordinate structures.

Pragmatic Constraints

Why should there be a VP-coordination preference at the connective? According to Hoeks et al. (2002) this preference is the consequence of language users, especially readers, constructing their own default 'topic-structure' in the absence of prosodic or other topic-marking cues. Topic-structure can be loosely defined as describing the relation between the *topic* of a sentence: the element referring to an entity about which information is given, and the *information* that is expressed by the sentence. In VP-coordinations there is only one topic, which is presumed to be the default and most frequently occurring situation, whereas for instance S-coordinations contain an additional topic, e.g., *the photographer*. Having more than one topic, Hoeks et al. argue, is unexpected and leads to processing difficulty because readers will have to accommodate an entity that has not been introduced as a second topic in their model of the discourse. This preference for a single topic is formally described in the UNIQUE TOPIC constraint, adopted from Beaver's (2004) work on anaphor resolution:

UNIQUE TOPIC: With respect to any sentence, there is exactly one discourse entity which is the topic of that sentence.

Importantly, the UNIQUE TOPIC constraint (UT) does not differentiate between VP-coordination and NP-coordination, as both constructions have but one topic, namely the subject of the sentence (see leftmost constraint "UT" in tableau 1). However, language users clearly prefer VP-coordination over NP-coordination at the connective. Thus, another constraint is in order to explain this preference.

Discourse-Semantic Constraints

We would like to propose that the constraint prohibiting NP-coordination is the discourse-semantic constraint FORWARD DISCOURSE:

FORWARD DISCOURSE: Introduce new events into the discourse.

If additional information has to be incorporated into the hearer's model of the discourse, this constraint prefers the introduction of a new event to elaboration of a previously introduced event. Its effect is the promotion of discourse structures consisting of a sequence of minimally elaborated events. It is important to note that only further empirical

testing can confirm the validity of this specific constraint, or whether it should be replaced by or augmented with another type of constraint.

To summarize, with the two constraints we just defined, we can describe one step in the incremental comprehension of coordinated structures, and explain how the VP-coordination preference arises at the connective in structures such as (4). Tableau 1 displays how the optimal VP-coordination interpretation is chosen from among the alternatives. We will only show a subset of candidates, namely the ones that have some chance of being optimal, but recall that there are infinitely many candidates (that clearly cannot be shown here). In OT it is assumed that constraints are hierarchically ordered, that is, from strongest constraint to weakest constraint. However, in this case both orderings, namely UT >> FD and FD >> UT produce the same optimal candidate, VP-coordination. In such instances, where there is no direct conflict between constraints, more evidence is needed to specify the correct ranking. This is signified by a dashed instead of a solid boundary between the constraints (see tableau 1).

Input: <i>Fragment (4)</i>	UT	FD
S-coordination	*	ok
● VP-coordination	ok	ok
NP-coordination	ok	*

*Tableau 1: Optimization in fragment (4); ●=optimal candidate for interpretation of the input; * = constraint violation; ok=constraint satisfaction; dashed boundary=ranking between the constraints has not been determined; UT=Unique Topic; FD=Forward Discourse.*

Syntactic Constraints

Now let us consider the situation in which the conjunction *and* is followed by an NP, as in sentence fragment (5).

5. The model embraced the designer and the photographer ...

We have seen that at the time the conjunction *and* is read, VP-coordination is the optimal parse candidate. However, if a VP-coordination parse is adopted in (5), this entails that the VP must have non-canonical word order. In English and Dutch, VPs do not normally begin with an NP in a non-embedded clause. The canonical word order in main clauses in these languages is SVO. Because, in (5), the first constituent following the conjunction is an NP, if the coordinate structure is preferably interpreted as VP-coordination the VP must have the non-canonical word order OV rather than VO. This option is ruled out by the same constraints on syntactic structure that also rule out OV word order in the first conjunct in (5). In this specific case it could be the hearer-oriented variant of the constraint STAY (or: "Do not move"), which prohibits movement of lexical items (taken from work by Grimshaw, 1997; see also Ackema & Neeleman, 1998):

STAY: Respect canonical word order.

So, the VP-coordination parse violates the constraint STAY

if the second conjunct starts with an NP. If the constraints STAY and UNIQUE TOPIC both outrank FORWARD DISCOURSE, as in tableau 2, this will account for the empirical observations discussed above that support the NP-coordination preference for (5). This fact also settles the indeterminacy of the ordering of UNIQUE TOPIC and FORWARD DISCOURSE. If namely the ordering S >> FD >> UT is assumed, this will yield S-coordination as the optimal parse, which goes against our empirical observations.

Input: <i>Fragment (5)</i>	S	UT	FD
S-coordination	ok	*	ok
○ VP-coordination	*	ok	ok
● NP-coordination	ok	ok	*

Tableau 2: Optimization in fragment (5); * = constraint violation; ok=constraint satisfaction; ●=currently preferred interpretation; ○=previously preferred interpretation; S=Stay; UT=Unique Topic; FD=Forward Discourse.

Note that the optimal parse of (5), corresponding to the NP-coordination interpretation, violates the FORWARD DISCOURSE constraint, which made NP-coordination sub-optimal in partial sentence (4) (see Tableau 2). Nevertheless, NP-coordination is optimal in (5) because the competing analyses violate stronger constraints.

Finally, when in a sentence such as (5) the ambiguous NP is followed by a finite verb (as in (6)), all options but S-coordination are rejected by STAY and also other, not further specified syntactic constraints, as no NP-coordinated or VP-coordinated sentence can be construed from the current ordered set of words (for convenience we will use STAY as a label for all of those).

6. The model embraced the designer and the photographer laughed ...

Input: <i>Fragment (6)</i>	S	UT	FD
● S-coordination	ok	*	ok
VP-coordination	*	ok	ok
○ NP-coordination	*	ok	*

Tableau 3: Optimization in fragment (6); * = constraint violation; ok=constraint satisfaction; ●=currently preferred interpretation; ○=previously preferred interpretation; S=Stay; UT=Unique Topic; FD=Forward Discourse.

We have seen that during the processing of the S-coordinated sentence (6), there are two occasions where there is a shift from one interpretation to another: When the ambiguous NP is read, the preference for VP-coordination shifts to a preference for NP-coordination, as VP-coordination becomes structurally impossible, and NP-coordination does not violate the UNIQUE TOPIC constraint. On the arrival of the disambiguating verb however, the NP-coordination reading becomes impossible and the S-coordinated alternative that has long been suboptimal, is now the optimal interpretation of the sentence. Under the 'Linking Hypothesis' (i.e., linking linguistic competence

and performance) proposed by Stevenson and Smolensky (2005), each of these shifts from one interpretation to another will give rise to processing difficulty. In the beginning of this section, we summarized a number of studies showing that there is indeed processing difficulty at the disambiguating verb of the S-coordinated sentences that are in focus here. Unfortunately, there is no empirical work explicitly testing whether there is processing difficulty due to the VP-coordination preference at the ambiguous NP in structures such as (5). Preliminary results from a recent ERP-study done in our lab, however, suggest that there is processing difficulty at the *article* of the ambiguous NP following the connective. Thus, the OT model formulated in tableau 4 seems to adequately capture all relevant aspects of processing coordinated sentences such as (6).

Lexical-Semantic Constraints

An important factor in the interpretation of linguistic utterances that we haven't dealt with yet is plausibility. Plausibility can be thought of as involving three interrelated categories of conceptual knowledge: 1) lexical-semantic knowledge (e.g., regarding the fillers of thematic roles), 2) knowledge about the discourse that is presently under consideration, and 3) general knowledge about the world. For our present purposes, we will only discuss the role the first kind of plausibility plays in our model. The chances for alternative interpretations to be optimal decline if the thematic fit of an argument is poor given the requirements of the thematic role assigner that is associated with it. This can be clearly seen in a sentence fragment such as (7).

7. Jasper sanded the board and the carpenter ...

Here, the poor thematic fit between *carpenter* and *sanded* argues against NP-coordination, which would normally have been the preferred structure (cf. Tableau 2). If on the basis of the thematic fit information S-coordination is assumed instead of NP-coordination, no processing difficulty is expected to ensue when a disambiguating verb comes in. And indeed, the results of the experiments by Hoeks et al. (2004), using sentences such as (8a), as compared to (8b), indicated that information regarding poor thematic fit was used very rapidly, and processing difficulty was largely eliminated (cf. Tableau 3).

8a. Jasper sanded the board and the carpenter repaired ...
8b. Jasper sanded the board, and the carpenter repaired ...

Thus a lexical-semantic factor such as thematic fit is of great influence on the processing of coordination. We will call the associated constraint THEMATIC FIT (adapted from Lamers and De Hoop, 2005).

THEMATIC FIT: A thematic element must meet the requirements of the thematic role that is assigned to it.

Tableau 4 shows how the constraints interact at the time the ambiguous NP *the carpenter* is read. Tableau 5 shows the constraint interaction at the time the disambiguating verb *repaired* is presented.

Input: <i>Fragment (7)</i>	TF	S	UT	FD
● S-coordination	ok	ok	*	ok
VP-coordination	ok	*	ok	ok
NP-coordination	*	ok	ok	*

Tableau 4: Optimization in fragment (7). Legend: see below.

Input: <i>Fragment (8a)</i>	TF	S	UT	FD
● S-coordination	ok	ok	*	ok
VP-coordination	ok	*	ok	ok
NP-coordination	*	*	ok	*

Tableau 5: Optimization in fragment (8a); * = constraint violation; ok=constraint satisfaction; ●=currently preferred interpretation; dashed boundary=ranking between the constraints has not been determined yet; TF=Thematic Fit; S=Stay; UT=Unique Topic; FD=Forward Discourse.

Note that on the basis of the currently available evidence we cannot decide on the ordering of the constraints THEMATIC FIT (TF) and STAY; these do not conflict in the structure at hand and appear in the tableaux in random order.

OT Versus Other Models

We have seen that in OT alternative interpretations for potentially ambiguous sentences compete and are evaluated on the basis of interacting constraints coming from multiple sources of information (e.g., syntactic, semantic, discourse-related). The interpretation that satisfies the total set of constraints best is the optimal interpretation, and will be selected by the hearer. In a sense, OT is very similar to ‘standard’ constraint-based models of human sentence processing as for instance proposed by MacDonald et al., (1994). In these models, it is assumed that lexical and syntactic processing is interleaved with interpretative processing and the construction of a discourse model, much like in the OT model that we propose here. Constraint-based models are usually contrasted with *syntax-first* models of human sentence processing (Frazier, 1987a). In syntax-first models, non-syntactic constraints only play a role after the initial phase of building a syntactic tree. In principle, syntax-first models can be viewed as a special case of constraint-based models, namely one in which syntax-based constraints (such as, e.g., Minimal Attachment and Late Closure) apply before other constraints. A syntax-first approach can easily be modeled in OT as well, by assuming that constraints pertaining to syntactic structure are stronger than other constraints. As a result, it will be more important to satisfy the syntactic constraints than it will be to satisfy the non-syntactic constraints.

Thus, at face value OT seems to be compatible with both the standard constraint-based models and syntax-first models of human sentence processing. But if OT is more or less similar to current models of human sentence processing, then what is the advantage of OT? To begin with, one important advantage of using OT in the domain of processing is that it also can serve as a complete model of competence, so no separate module describing the

knowledge of a language is required (as in other processing theories). In addition, OT provides a relatively restricted framework within which to interpret empirical data. Standard constraint-based models are probabilistic, which means that constraints are defined in terms of the amount of (probabilistic) support they provide for one syntactic alternative over the other. In the strong version of OT that we propose here, however, any parsing preference produced by the constraints is discrete, as constraints are violated or not; they cannot be violated to a certain degree. Thus, OT makes stronger predictions that can be easily tested.

A related difference between OT and standard constraint-based models is that constraint violations in our version of OT are not cumulative. This means that any higher-ranked constraint takes absolute priority over any lower-ranked constraint. A single violation of a higher-ranked constraint is always worse than any number of violations of any number of lower-ranked constraints. In this respect, OT is also more restricted than constraint-based models. As a result, the OT framework is able to prohibit certain output patterns, while a system of numerical constraints is not. The fact that certain output patterns never arise can therefore only receive a principled explanation in OT, whereas such an observation would be completely coincidental in ‘standard’ constraint-based models.

Note that, although OT constraints produce discrete preferences, this does not imply that OT models cannot give rise to probabilistic effects. With the partial ranking model of Anttila and Cho (1998), or the stochastic OT variant proposed by Boersma (1998), very specific probabilistic output patterns can be predicted.

Adopting the OT approach has also considerable advantages in the domain of competence. Whereas many competence theories apply the grammar at the level of the entire sentence, our proposed OT performance theory incrementally optimizes interpretation one word at a time. Such an approach is impossible within grammatical frameworks of inviolable constraints such as current generative syntax, because sentence fragments simply do not meet inviolable constraints and hence cannot be interpreted. By contrast, in OT any input receives a parse because the constraints are violable. Consequently, many theories of linguistic competence do not seem to be extendible to the domain of processing. Although attempts have been made to apply competence results from generative syntax to the domain of processing (e.g., Weinberg, 2001), these attempts do not involve the entire grammar but rather a very limited set of conditions that can be viewed as violable conditions as well, such as those regarding syntactic economy. And although lexicalist linguistic theories like Categorical Grammar also allow for the interpretation of sentence fragments, they require these sentence fragments to be syntactically well-formed. In OT, on the other hand, every input receives an interpretation, even if this input would be ungrammatical. This corresponds to the observation that we understand people even when their utterances are not completely well-formed. There is currently no other theory that has a principled account for this phenomenon.

Conclusion

In this paper we developed an OT model of coordinated structures, and we argued that adopting an OT framework has many advantages: 1) it solves some of the problems that, until now, beset syntactic theories; 2) it incorporates linguistic competence as no other current model of language processing; 3) as a processing model, it accounts for all processing phenomena associated with coordinated structures; 4) it allows for the interpretation of all structure, including ungrammatical utterances; and finally 5) it provides very clear and testable predictions. Importantly, the model as it is formulated here represents only the first step towards a complete theory of sentence processing. Many aspects of the model still have to be developed and tested. For one thing, though we aimed to only include constraints that have been validated in earlier research (e.g. UNIQUE TOPIC, adopted from Beaver's (2004) work on discourse anaphora, or THEMATIC FIT, adapted from studies by Lamers & De Hoop regarding the processing of Subject-Object ambiguities), it is essential to further investigate the FORWARD DISCOURSE constraint, as it has not yet been empirically supported. In addition, it is important to find out whether the present model makes the right predictions for other linguistic phenomena. Nevertheless, we are confident that the OT approach to coordination that was put forward here can be successfully transferred to other areas of language interpretation, and that constructing an OT grammar of language interpretation in general is indeed a feasible enterprise.

Acknowledgements

This research was partly funded by grant # 015.001.103 from the Netherlands Organization for Scientific Research, awarded to Petra Hendriks. We thank Helen de Hoop, Monique Lamers, and Laurie Stowe for helpful comments.

References

- Ackema, P., & Neeleman, A. (1998). Optimal questions. *Natural Language and Linguistic Theory*, 16, 443-490.
- Anttila, A., & Cho, Y.-M. Y. (1998). Variation and change in Optimality Theory. *Lingua*, 104, 31-56.
- Artstein, R. (2000). *Case constraints and empty categories in optimality theory parsing*. Ms.: Rutgers.
- Barbosa, P., Fox, D., Hagstrom, P., McGinnis, M., Pesetsky, D. (1998) (Eds.). *Is the Best Good Enough? Optimality and Competition in Syntax*. Cambridge: MIT.
- Beaver, D. (2004). The optimization of discourse anaphora. *Linguistics & Philosophy*, 27(1), 3-56.
- Boersma, P. (1998). *Functional phonology*. The Hague: Holland Academic Graphics.
- Borsley, R.D. (1994). In defense of coordinate structures. *Linguistic Analysis*, 24:3-4, 218-246.
- De Hoop, H. & Lamers, M. (to appear). Incremental distinguishability of subject and object. In: *Studies in Language Companion Series*. Benjamins, Amsterdam.
- De Vries, M. (2003). Three-dimensional grammar. In L. Cornips & P. Fikkert (Eds.), *Linguistics in the Netherlands 2003*, 2001-212.
- Fanselow, G., Schlesewsky, M., Cavar, D., & Kliegl, R. (1999). *Optimal parsing: Syntactic parsing preferences and optimality theory*. Ms.: U. Potsdam, ROA 367.
- Frazier, L. (1987a). Sentence processing: A tutorial review. In M. Coltheart (Ed.), *Attention and performance XII: The psychology of reading*. Hillsdale, NJ: Erlbaum.
- Frazier, L. (1987b). Syntactic processing: Evidence from Dutch. *Natural Language and Linguistic Theory*, 5, 519-559.
- Gibson, E., & Broihier, K. (1998). Optimality theory and human sentence processing. In: P. Barbosa et al. (Eds.), *Is the best good enough? Optimality and competition in syntax*. Cambridge: MIT.
- Grimshaw, J. (1997). Projections, heads and optimality. *Linguistic Inquiry*, 28, 373-422.
- Hendriks, P., & de Hoop, H. (2001). Optimality Theoretic Semantics. *Linguistics and Philosophy*, 24, 1-32.
- Hoeks, J.C.J., Hendriks, P., Vonk, W., Brown, C.M., & Hagoort, P. (in press). *Resolving the NP/S conjunction ambiguity*. Quarterly Journal of Experimental Psychology.
- Hoeks, J.C.J., Vonk, W., & Schriefers, H. (2002). Processing coordinated structures in context: the effect of topic-structure on ambiguity resolution. *Journal of Memory and Language*, 46, 99-119.
- Johannessen, J.B. (1993). *Coordination: A minimalist approach*. Ph.D. Dissertation, University of Oslo.
- Kager, R. (1999). *Optimality Theory*. Cambridge: CU Press.
- Kempen, G. (2004). *Symmetrical clausal coordination as grammaticalized self-repair*. Paper presented at Tabu-dag 2004, Groningen, June 4, 2004.
- Lamers, M., & de Hoop, H. (2005). Animacy information in human sentence processing: an incremental optimization approach. LNAI, 4338.
- MacDonald, M.C., Pearlmutter, N.J., & Seidenberg, M. (1994). The lexical nature of syntactic ambiguity resolution. *Psychological Review*, 101, 676-703.
- Mitchell, D.C., Cuetos, F., Corley, M.M.B., & Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained statistical records. *Journal of Psycholinguistic Research*, 24, 469-487.
- Munn, A. (1987). *Coordinate Structures, X-bar Theory and Parasitic Gaps*. Honours research paper, McGill University, Montreal.
- Phillips, C. (1995). Right Association in Parsing and Grammar. In: C. Schütze et al. (Eds.), *Papers on Language Processing and Acquisition*. Cambridge: MIT.
- Prince, A., & Smolensky, P. (2004). *Optimality Theory. Constraint Interaction in Generative Grammar*. Blackwell.
- Ross, J. R. (1967). *Constraints on variables in syntax*. MIT doctoral dissertation.
- Smolensky, P., & Legendre, G. (2005). *The harmonic mind: From neural computation to Optimality-Theoretic grammar* (2 volumes). Cambridge: MIT.
- Stevenson, S., & Smolensky, P. (2005). Optimality in sentence processing. In: P. Smolensky & G. Legendre (Eds.), *The Harmonic Mind*. Cambridge: MIT.
- Weinberg, A. (2001). A Minimalist Theory of Human Sentence Processing. In: Epstein & Hornstein (eds.), *Working Minimalism*. Cambridge: MIT.