Integrating Discourse in a Computational Model of the Production and Comprehension of Referring Expressions

Jacolien van Rij (J.C.van.Rij@rug.nl)
Center for Language and Cognition, University of Groningen
P.O.Box 716, 9700AS Groningen, The Netherlands

Petra Hendriks (P.Hendriks@rug.nl)
Center for Language and Cognition, University of Groningen
P.O.Box 716, 9700AS Groningen, The Netherlands

Hedderik van Rijn (D.H.van.Rijn@rug.nl)
Department of Psychology, University of Groningen
Grote Kruisstraat 2/1, 9712TS Groningen, The Netherlands

Abstract

In this paper we propose a cognitive model that simulates the acquisition of object pronouns. The model will be implemented to generate precise and testable predictions about the comprehension and production of object pronouns. During sentence processing, the model incrementally builds a structural representation of the sentence, and the model incrementally constructs a discourse representation as well. Furthermore, for adult-like performance on the comprehension of pronouns the model needs to take into account the perspective of the speaker in addition to its own perspective as a hearer. Using this model, we can investigate the interaction between grammar, cognitive constraints, and discourse in the production and comprehension of referring expressions.

Keywords: referring expressions; language acquisition; cognitive modeling.

Introduction

The production and comprehension of referring expressions is not only guided by the interaction between grammar and cognitive constraints, but also by discourse. In previous work, we developed a cognitive model that simulates the acquisition of object pronouns (Van Rij, Hendriks, Spenader, & Van Rijn, 2009). The model predicts that children who show difficulties in pronoun comprehension but not in reflexive comprehension will perform better on the comprehension of object pronouns if they are given more time for interpretation. This hypothesis was supported by the results of a psycholinguistic experiment. In this paper we propose a re-implementation of that model that also simulates discourse effects on the comprehension of object pronouns. Using this model, we can investigate the interaction between grammar, cognitive constraints, and discourse in the production and comprehension of referring expressions.

Object pronouns

Several studies have shown that children are able to comprehend reflexive sentences such as (1a) correctly from the age of 3;0 on, but show difficulties in the interpretation of pronoun sentences like (1b) up to the age of 6;6. They incorrectly allow the pronoun to corefer with the local subject about half the time (a.o., Chien & Wexler, 1990; Jakubowicz, 1984; Koster, 1993; Spenader, Smits, & Hendriks, 2009). At the same time, these children show adult-like performance in their *production* of reflexives and pronouns (e.g., De Villiers, Cahillane, & Altreuter, 2006; Spenader et al., 2009). So children produce pronouns correctly before they have acquired adult-like comprehension of pronouns.

(1) a. The penguin$_i$ is hitting himself$_{i/*j}$ with a pan.
b. The penguin$_i$ is hitting him$_{*i/j}$ with a pan.

For adults, the reflexive *himself* in (1a) has to co-refer with *the penguin* because a reflexive must be bound in its local domain (Principle A of Binding Theory, Chomsky, 1981). In contrast, a pronoun must be locally free (Principle B of Binding Theory). Therefore the pronoun *him* in (1b) cannot refer to *the penguin*, and can only refer to another antecedent present in the context. An important question is why children still make comprehension errors with Principle B as late as age 6, while they perform adult-like on its sister principle from a young age on.

Bidirectional optimization

Hendriks and Spenader (2005/2006) propose in the context of Optimality Theory (OT, A. Prince & Smolensky, 2004) that children's difficulties with pronoun comprehension are caused by a direction-sensitive grammar (for alternative explanations, see a.o., Conroy, Takahashi, Lidz, & Phillips, 2007; Reinhart, 2006; Thornton & Wexler, 1999.) In OT, the grammar consists of a set of ranked and violable constraints. For every input, which can be either a form in comprehension or a meaning in production, a set of potential outputs, or candidates, is generated. The optimal candidate is the candidate that commits the least severe constraint violations. One violation of a higher ranked constraint is more important than many violations of lower ranked constraints. The same grammar can be used for production and comprehension. Importantly, the constraints are direction-sensitive, which means that they can have different effects in production and comprehension.

Hendriks and Spenader (2005/2006) argue that children's performance on pronoun production is guided by a general preference for forms with less referential content. This preference is reflected in a constraint hierarchy Referential Economy, according to which reflexives are more economical than pronouns, and pronouns a more economical than full NPs (cf. Burzio, 1998; Gundel, Hedberg, & Zacharski, 1993). Pronominal forms are used only in cases when the constraint Principle A (the OT version of Principle A of Binding Theory, Chomsky, 1981) prohibits the use of a reflexive, for example when expressing disjoint reference. Because Referential Economy expresses a preference among forms, this constraint hierarchy only plays a role in production, and not in comprehension. In contrast, Principle A plays a role in both directions of optimization. Because Hendriks and Spenader do not assume a constraint Principle B, pronominal forms are in principle ambiguous. Therefore, children are predicted to show chance performance on the interpretation of sentences with object pronouns.

Hendriks and Spenader (2005/2006) argue that adults' performance does not reflect this direction-sensitivity of the grammar, because adults also take into account the perspective of the conversational partner. This pragmatic process is illustrated in Figure 1.

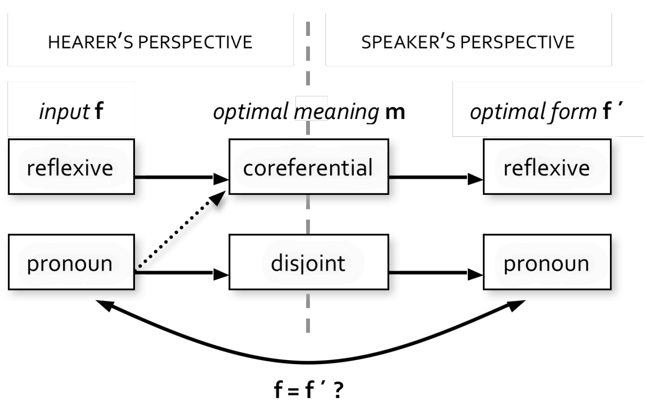

Figure 1. Taking into account the speaker's perspective in comprehension.

From the hearer's perspective, pronouns are ambiguous between a coreferential and a disjoint interpretation. However, if an adult hearer takes into account the speaker's perspective, the constraints of the grammar will allow the hearer to conclude that a coreferential interpretation is best expressed using a reflexive form, because of Principle A. As a result, a coreferential interpretation for pronouns (represented by the dotted line in Figure 1) is blocked. This disambiguation of the meaning of pronouns is formalized as *bidirectional optimization* (Blutner, 2000). Note that for correct *production* of pronouns the second step is not required. Because, in this particular case, the correct form is already fully determined by *unidirectional optimization*, the speaker need not take into account the hearer's perspective. Thus production and comprehension are closely related: The same grammar is used for both directions of optimization, and adult hearers adopt the perspective of a speaker and vice versa.

The OT explanation discussed above provides an explanation why children's comprehension of pronouns is acquired later than the comprehension of reflexives. Furthermore, it explains why children's production of pronouns may already be adult-like, while their comprehension of pronouns is still poor. However, since OT is a theory of linguistic competence and not a theory of linguistic performance, it does not provide an explanation for the change in optimization mechanism between children and adults. To be able to generate precise and testable predictions with respect to the development of the ability to apply bidirectional optimization, we implemented the OT explanation of the acquisition of object pronouns in a cognitive model (Hendriks, Van Rijn, & Valkenier, 2007; Van Rij, Hendriks, Spenader, & Van Rijn, to appear).

Cognitive model of the acquisition of object pronouns

Cognitive models are computational simulations of the cognitive processes involved in performing a task, for example comprehending an anaphoric sentence. The model is implemented within the cognitive architecture ACT-R (Anderson et al., 2004), a modeling environment that constrains simulation models on the basis of built-in and well-tested mechanisms and parameters to ensure psychological plausibility.

The ACT-R model simulates bidirectional optimization as two serial processes of unidirectional optimization: 1) selecting the optimal meaning for the (ambiguous) pronoun, and 2) checking whether a speaker would have expressed this meaning with the same form. Since bidirectional optimization is simulated as two serial processes of unidirectional optimization, initially bidirectional optimization takes more time to complete than unidirectional optimization. Given that time for interpretation is limited in sentence processing, the model predicts that children initially cannot use bidirectional optimization, as illustrated in Figure 2a. They lack the processing efficiency to complete two processes of unidirectional optimization within the limited amount of time. When the process is repeatedly performed, a learning mechanism in the ACT-R architecture called *production compilation* gradually increases the efficiency of the process (Taatgen & Anderson, 2002). Eventually, children's processing efficiency is sufficient for completing bidirectional optimization within the limited amount of time and adult-like performance is reached. This stage in development is depicted in Figure 2b. In other words, the computational model assumes that young children have the ability to optimize bidirectionally, but lack the processing efficiency to do so. So, if children who show difficulties in pronoun comprehension (but not in reflexive comprehension) are given more time for interpretation, the model predicts that their performance on pronoun comprehension will improve. Figure 2c illustrates this prediction.

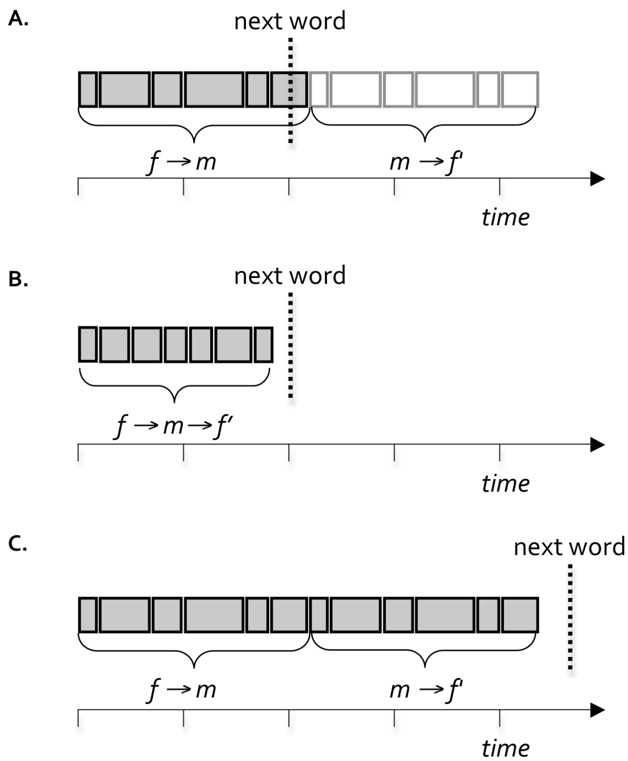

Figure 2. Schematic illustration of the predictions of the cognitive model: a) children's processing; b) adult-like processing; c) children's processing when they are given more time for interpretation. The blocks represent the cognitive operations involved in pronoun interpretation (Van Rij, Hendriks, Spenader, & Van Rijn, to appear to appear).

We performed a psycholinguistic experiment to compare children's pronoun comprehension at a normal speech rate with their comprehension at a slow speech rate (Van Rij et al., 2009). By slowing down the speech rate, children have more time for interpretation. As a result of slowed-down speech, the performance of children who showed difficulties in pronoun comprehension was found to improve on pronoun sentences, but not on reflexive sentences. These results support the hypothesis that children's difficulties with pronoun comprehension are caused by children's insufficient processing efficiency. As a result of their lack of processing efficiency, they cannot take into account the speaker's perspective as a hearer (Van Rij et al., 2009).

Although this model was designed with pronoun comprehension in mind, it also predicts the correct pattern of performance for pronoun production. The optimization process in production is similar to that in comprehension, but proceeds in the opposite direction. That is, production is modeled as optimization from an input meaning to the optimal form for expressing that meaning. Interestingly, the model predicts that the use of bidirectional optimization in production increases with time, although bidirectional optimization is not strictly necessary for the correct production of object pronouns since the output of bidirectional optimization is not different from the output of unidirectional optimization

Modeling discourse processing

Our previous work has shown that the acquisition of object pronouns can be explained by the interaction between an optimality theoretic grammar and general constraints on cognition. It is generally assumed that binding phenomena are sentence-level phenomena, therefore discourse processing was not implemented in the current model. However, Spenader, Smits and Hendriks (2009) showed that children's difficulties with pronoun comprehension disappeared when the introduction sentence established the intended referent of the pronoun as the topic of the discourse. Children showed almost correct performance on sentences like (1b) when preceded by (2b), but showed just above chance performance on these sentences with the classic introduction (2a).

(2) a. Here you see a penguin and a sheep.

b. Here you see a sheep.

Their results indicate that discourse can affect the interpretation of pronouns in binding contexts. Therefore, any cognitively plausible theory of sentence processing also needs to explain how discourse structure can affect the resolution of object pronouns. To incorporate the ability to process discourse information in our model, the sentence-processing component should be able to build a representation of the discourse structure during sentence processing.

Proposal for re-implementation

We propose a cognitive model of sentence processing and generation in ACT-R that not only processes structural information (cf. Lewis & Vasishth, 2005), but also semantic and discourse information. The model contains three main components, namely 1) a sentence-processing component, 2) a discourse component, and 3) the bidirectional optimization component of the model described in the previous sections. This section describes the sentence-processing component and the discourse component of the model.

Sentence processing.

The sentence-processing component of the cognitive model is based on the sentence processing model of Lewis and Vasishth (2005) in ACT-R. Our model implements OT sentence processing (e.g. De Hoop & Lamers, 2006; Fanselow, Schlesewsky, Cavar, & Kliegl, 1999; Stevenson

& Smolensky, 2006) to be able to account for discourse effects.

Lewis and Vasishth (2005) modeled sentence processing as a series of very efficient memory retrievals. In ACT-R, information is represented by one or more chunks, i.e., pieces of knowledge. These memory chunks have a certain activation that fluctuates and is dependent on the usage history of that information and the recency of the last retrieval. The activation increases when the chunk is being retrieved. In general the activation of chunks decays with time, unless they are re-activated again. The activation determines the time it takes to retrieve the information: the higher the activation, the less time retrieval will take. Chunks that are retrieved or manipulated spread activation to other chunks. So the activation of a chunk can increase as a result of the association with a chunk that is being processed. Lewis and Vasishth (2005) argue that the fluctuating activation of memory elements and associative interference result in differences in memory retrieval times. Longer memory retrieval times will lead to longer reading times, which are generally considered indications of processing difficulties.

Sentence processing is simulated as an incremental process. Every time a word is encountered, syntactical and lexical information related to that word, including argument structure, becomes available. Furthermore, a tree structure that represents the structural analysis of the previous words is retrieved and the new word is attached to that tree structure. The syntactical and lexical information of the encountered word contains a kind of syntactic expectation that guides the retrieval of the tree structure. Our model differs from the model of Lewis and Vasishth (2005) in the mechanism of the model that determines a single interpretation at the time. Whereas Lewis and Vasishth implemented the Left Corner parsing algorithm (Aho & Ullman, 1972) to resolve ambiguities, we implement OT sentence processing to do that, so that our model is able to explain the asymmetry in the acquisition of pronouns as well.

In our model the optimality theoretic grammar is used to determine the optimal interpretation at that point in the sentence on the basis of an optimization process. The input of this process consists of one or more words from the start of the sentence up to the last pronounced word. The output is the optimal meaning at that point in the sentence. To determine the optimal interpretation for a given input, two candidate interpretations are evaluated at a time. These candidates are tree structures to which the word can be attached. In the cognitive model, only one constraint can be applied at a time to evaluate the two candidates. Because chunks are ordered with respect to their activation value, the system will retrieve the most relevant candidates first. The candidate that violates the constraint is replaced by another candidate, and another process of comparison takes place. If there is no other candidate, the remaining candidate will be selected as the optimal meaning. If the two candidates show the same pattern of constraint violations (both violating or satisfying this constraint), the next constraint will be retrieved. If none of the constraints distinguishes the two candidate meanings, then one of the candidates is randomly selected as the optimal meaning. This process in principle always finds the optimal candidates by first evaluating the candidates against the higher ranked constraints and, if necessary, further evaluating the candidates against lower ranked constraints. After one of the candidate structures is selected, a new structure is created that attaches the current word to the candidate interpretation.

Note that this process of sentence-processing is subject to the production compilation mechanism, the learning mechanism of ACT-R that gradually increases the efficiency of the bidirectional optimization process. Therefore, sentence processing will eventually become largely proceduralized, so that the model is able to analyze a sentence within the limited amount of time that is available during online sentence processing.

Discourse processing.

It is generally assumed that the form used to produce a referring expression and the interpretation of a referring expression are dependent on the saliency or accessibility of a referent (e.g., Ariel, 1988; Arnold, 1998; Givón, 1983; Gundel et al., 1993). For example, Gundel et al. (1993) have proposed the Givenness Hierarchy, an implicational scale that relates the givenness of the referent with the referring form that is used. A speaker will use a form that is as short and unspecific as possible, but at the same time informative enough for the listener to understand the intended meaning (cf. Grice's Maxim of Quantity, Grice, 1975). So less specific forms, such as pronouns, are only used for sufficiently accessible referents. However, different theories have been proposed about the factors that determine saliency (e.g., Ariel, 1988; Arnold, 1998; Givón, 1983; Gundel et al., 1993; E. Prince, 1981).

In our re-implementation of the model, discourse representation is built up incrementally during sentence processing. The discourse representation is modeled as a collection of activated concepts introduced by the linguistic context. During processing the semantic meaning of the incoming words is activated. The meaning of a word or a group of words is represented by a chunk. As a result, the activation of a discourse element is determined by the recency of the last retrieval, the history of retrievals and the association with other discourse elements. The activated semantic chunks in the declarative memory together form the discourse representation. In this account, the discourse representation is not limited to the referents in the previous utterance (in contrast to a.o., Grosz, Weinstein, & Joshi, 1995). However, the activation of a discourse element decays over time. Therefore, a referent that is introduced many sentences before disappears from the discourse representation, unless it is re-activated by association or unless it is referred to again.

So saliency of discourse referents is modeled as a gradient property that is determined by the linguistic context and cognitive constraints. To be able to simulate the effect of discourse on the processing of pronoun sentences, it is necessary to specify how the grammar interacts with the discourse representation. In our model the notion of discourse topic is necessary for the grammar to integrate discourse information in the analysis of the sentence. For the production and comprehension of pronouns, it is important to establish the topic of the current sentence because pronouns preferably refer to the topic (Grosz et al., 1995; Spenader et al., 2009). In our cognitive model, the optimality theoretic grammar determines the topic from the discourse structure (cf. Beaver, 2004). Only two candidates can be evaluated at the same time, and the constraints to evaluate these candidates are applied in sequence. The model will first consider the most active, i.e., the most salient, referents in the discourse as possible topics, because the most relevant or most recent referent is probably the most activated discourse element.

In summary, the saliency of discourse referents, reflected by their activation in the discourse, is modeled as a gradient property (cf. Arnold, 1998) following from the linguistic context and cognitive constraints, such as the fluctuation of activation. The activated semantic meanings together form the discourse representation. In contrast to saliency, topicality is not modeled as a gradient property: the grammar establishes exactly one discourse referent as the topic of the current utterance (cf. Beaver, 2004; Grosz et al., 1995; but see Arnold, 1998). The grammar establishes a unique topic of an utterance to be able to use this information in sentence processing.

Our re-implemented model thus processes structural information as well as discourse information by combining the incremental process of assigning a structure to the sentence with incrementally construing a representation of the saliency of discourse referents. Comprehension of referring expressions is modeled as a bidirectional process, whereby hearers not only consider their own perspective but also take into account the perspective of the speaker. Similarly, in production, speakers take into account the perspective of their hearers. These bidirectional processes are believed to be essential for adult processing but may be too taxing for children, resulting in delays in acquisition.

Future directions: Modeling the production of referring subjects

Based on the re-implementation of the cognitive model, simulations are able to provide more specific predictions with regard to discourse effects on the comprehension of object pronouns. These predictions will be investigated using psycholinguistic experiments. In addition to the processing of object pronouns, this model will also be tested on the processing of subject pronouns. If children are unable to take into account the perspective of their conversational partner, an acquisition delay is predicted in children's *production* of *subject* pronouns too (Hendriks, Englert, Wubs, & Hoeks, 2008). If children prefer forms with less referential content over forms with more referential content (i.e., Referential Economy), they will prefer pronouns over full NPs, even for reference to a discourse element that is not the discourse topic. However, a hearer will usually interpret a pronoun as referring to the topic. Therefore, the use of a pronoun to refer to a non-topic must be blocked by bidirectional optimization in the adult language. Since children are unable to optimize bidirectionally, they fail to block this use of pronouns and overuse pronouns to also refer to non-topics (Hendriks et al., 2008). Consequently, children's production of subject pronouns is delayed. An important question is whether and how this delay in the production of referring expressions is related to the finding that children show difficulties in the comprehension of object pronouns as a result of their processing efficiency. Running simulations with our cognitive model may help to clarify this issue.

References

Aho, A. V., & Ullman, J. D. (1972). *The theory of parsing, translation, and compiling*. Englewood Cliffs, NJ: Prentice Hall.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*(4), 1036-1060.

Ariel, M. (1988). Referring and accessibility. *Journal of Linguistics, 24*, 65-87.

Arnold, J. E. (1998). *Reference Form and Discourse Patterns.* Unpublished Ph.D. thesis, Stanford University.

Beaver, D. (2004). The optimization of discourse anaphora. *Linguistics and Philosophy, 27*(1), 3-56.

Blutner, R. (2000). Some aspects of optimality in natural language interpretation. *Journal of Semantics, 17*(3), 189-216.

Burzio, L. (1998). Anaphora and soft constraints. In P. Barbosa, D. Fox, P. Hagstrom, M. McGinnis & D. Pesetsky (Eds.), *Is the Best Good Enough? Optimality and Competition in Syntax* (pp. 93-114). Cambridge, MA: MIT Press.

Chien, Y. C., & Wexler, K. (1990). Children's Knowledge of Locality Conditions in Binding as Evidence for the Modularity of Syntax and Pragmatics. *Language Acquisition, 1*(3), 225-295.

Chomsky, N. (1981). *Lectures on government and binding: the Pisa lectures*. Dordrecht, The Netherlands: Foris Publications.

Conroy, A., Takahashi, E., Lidz, J., & Phillips, C. (2007). *Equal treatment for all antecedents: How children succeed with Principle B.* Unpublished manuscript, University of Maryland, College Park.

De Hoop, H., & Lamers, M. (2006). Incremental distinguishability of subject and object. In L. Kulikov, A. Malchukov & P. De Swart (Eds.), *Case, valency and transitivity* (pp. 269-287). Amsterdam: John Benjamins.

De Villiers, J., Cahillane, J., & Altreuter, E. (2006). What can production reveal about Principle B. *Proceedings of the Inaugural Conference on Generative Approaches to*

Language Acquisition–North America, University of Connecticut, 4.

Fanselow, G., Schlesewsky, M., Cavar, D., & Kliegl, R. (1999). *Optimal parsing: Syntactic parsing preferences and optimality theory.* Unpublished manuscript, University of Potsdam.

Givón, T. (1983). Topic and continuity in discourse: An introduction. In T. Givón (Ed.), *Topic continuity in discourse: A quantitative cross-language study* (pp. 1-42). Amsterdam: John Benjamins Publishing.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and Semantics* (Vol. 3: Speech acts, pp. 41-58). New York: Academic Press.

Grosz, B. J., Weinstein, S., & Joshi, A. K. (1995). Centering: a framework for modeling the local coherence of discourse. *Computational Linguistics, 21*(2), 203-225.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language(Baltimore), 69*(2), 274-307.

Hendriks, P., Englert, C., Wubs, E., & Hoeks, J. (2008). Age differences in adults' use of referring expressions. *Journal of Logic, Language and Information, 17*(4), 443-466.

Hendriks, P., & Spenader, J. (2005/2006). When production precedes comprehension: An optimization approach to the acquisition of pronouns. *Language Acquisition, 13*(4), 319-348.

Jakubowicz, C. (1984). On markedness and binding principles. *Proceedings of the North Eastern Linguistics Society, 14*, 154-182.

Koster, C. (1993). *Errors in anaphora acquisition.* Doctoral dissertation, Utrecht University, Utrecht, The Netherlands.

Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science, 29*(3), 375-419.

Prince, A., & Smolensky, P. (2004). *Optimality Theory: Constraint Interaction in Generative Grammar* Malden, MA: Blackwell.

Prince, E. (1981). Toward a taxonomy of given-new information. In P. Cole (Ed.), *Radical Pragmatics* (pp. 223-255). New York: Academic Press.

Reinhart, T. (2006). *Interface Strategies: Optimal and Costly Computations*: The MIT Press.

Spenader, J., Smits, E. J., & Hendriks, P. (2009). Coherent discourse solves the Pronoun Interpretation Problem. *Journal of Child Language, 36*(1), 23-52.

Stevenson, S., & Smolensky, P. (2006). Optimality in sentence processing. In P. Smolensky & G. Legendre (Eds.), *The Harmonic Mind: From Neural Computation to Optimality-Theoretic Grammar* (Vol. 2, Linguistic and philosophical implications). Cambridge, MA: MIT Press.

Taatgen, N. A., & Anderson, J. R. (2002). Why do children learn to say "Broke"? A model of learning the past tense without feedback. *Cognition, 86*(2), 123-155.

Thornton, R., & Wexler, K. (1999). *Principle B, VP Ellipsis, and Interpretation in Child Grammar*: MIT Press.

Van Rij, J., Hendriks, P., Spenader, J., & Van Rijn, H. (2009). From group results to individual patterns in pronoun comprehension. In J. Chandlee, M. Franchini, S. Lord & M. Rheiner (Eds.), *Proceedings of the 33rd annual Boston University Conference on Language Development* (Vol. 2, pp. 563-574). Somerville, MA: Cascadilla Press.

Van Rij, J., Hendriks, P., Spenader, J., & Van Rijn, H. (to appear). Modeling the selective effects of slowed-down speech in pronoun comprehension. In *Proceedings of the 3th Conference on Generative Approaches to Language Acquisition –North America, University of Connecticut*. Somerville, MA: Cascadilla Press.