J. Child Lang. **37** (2010), 731–766. © Cambridge University Press 2010
doi:10.1017/S0305000909990560

Cognitive architectures and language acquisition: A case study in pronoun comprehension*

JACOLIEN VAN RIJ, HEDDERIK VAN RIJN AND PETRA HENDRIKS

University of Groningen

(*Received* 23 *December* 2008 – *Revised* 31 *August* 2009 – *Accepted* 20 *December* 2009 – *First published online* 30 *March* 2010)

ABSTRACT

In this paper we discuss a computational cognitive model of children's poor performance on pronoun interpretation (the so-called Delay of Principle B Effect, or DPBE). This cognitive model is based on a theoretical account that attributes the DPBE to children's inability as hearers to also take into account the speaker's perspective. The cognitive model predicts that child hearers are unable to do so because their speed of linguistic processing is too limited to perform this second step in interpretation. We tested this hypothesis empirically in a psycholinguistic study, in which we slowed down the speech rate to give children more time for interpretation, and in a computational simulation study. The results of the two studies confirm the predictions of our model. Moreover, these studies show that embedding a theory of linguistic competence in a cognitive architecture allows for the generation of detailed and testable predictions with respect to linguistic performance.

INTRODUCTION

An influential but also controversial distinction in linguistic research is the distinction between linguistic competence and linguistic performance (Chomsky, 1965). Linguistic competence pertains to the idealized linguistic

[*] This investigation was supported in part by a grant from the Netherlands Organization for Scientific Research, NWO, awarded to Petra Hendriks (grant no. 277-70-005). The authors thank the children, teachers and parents of the Nassauschool in Groningen for participating in the experiment, Sanne M. Berends and Sanne Kuijper for assisting with the experiment, Petra van Berkum and Robert Prins for drawing the pictures, Margreet Drok-van der Meulen for recording the sentences, and Jennifer Spenader, the Groningen Cognitive Modeling Group and the University of Groningen Language Acquisition Lab for helpful discussions regarding this study. Jacolien van Rij, Center for Language and Cognition Groningen (CLCG), University of Groningen. PO Box 716, 9700AS Groningen, The Netherlands. tel: +31 (0)50 363 5873; fax: +31 (0)50 363 6855; e-mail: J.C.van.Rij@rug.nl

knowledge a language user has of his or her language, which is often contrasted with linguistic performance, the actual use of this knowledge in concrete situations. This distinction between competence and performance provided a rationale for studying linguistic phenomena separately from cognitive factors. However, this distinction also created the methodological problem that it became impossible to empirically test theories of linguistic competence solely by studying linguistic performance. As a result, linguistic analyses appealing to aspects of linguistic performance such as insufficient working memory capacity, processing limitations or pragmatic skills are difficult to evaluate. Nevertheless, such analyses have been proposed in many areas of language acquisition to explain differences in linguistic performance between children and adults.

The aim of this paper is to show that embedding a theory of linguistic competence in a cognitive architecture may allow for the generation of detailed and testable predictions with respect to linguistic performance. A cognitive architecture is a general framework that incorporates built-in and well-tested parameters and constraints on cognitive processes. Within a cognitive architecture, computational models can be built that simulate the cognitive processes involved in performing a task such as interpreting a sentence. The predictions generated by these computational models can be tested on the basis of empirical data, for example the performance results obtained from a psycholinguistic experiment. As a case study, we present an account of the Delay of Principle B Effect in language acquisition (e.g. Chien & Wexler, 1990; Jakubowicz, 1984; Koster, 1993). The Delay of Principle B Effect (DPBE) concerns the observation that children's comprehension of pronouns is delayed in comparison with their comprehension of reflexives. Initially, children show incorrect performance on pronoun comprehension as well as on reflexive comprehension. However, when they have mastered reflexive comprehension, they still show incorrect performance on pronoun comprehension. This phenomenon in language acquisition is referred to as the DPBE. It can take several years before children show correct performance on both pronoun comprehension and reflexive comprehension.

The DPBE has received a variety of explanations, many of which appeal to performance factors to account for children's errors in comprehending pronouns. One such explanation is formulated within the linguistic framework of Optimality Theory (Hendriks & Spenader, 2005/2006). We show how a cognitive model can be built within the cognitive architecture ACT-R (Anderson, Bothell, Byrne, Douglass, Lebiere & Qin, 2004) that implements an optimality theoretic explanation of the DPBE. The resulting cognitive model predicts that children will make fewer errors in their interpretation of pronouns but not in their interpretation of reflexives if they are given more time for comprehension, for example by slowing down

the speech rate. We tested this prediction empirically in a psycholinguistic study as well as in a computational simulation study.

The organization of this paper is as follows. First, we discuss the DPBE and several of the proposed explanations to account for this delay in language acquisition, including a detailed account of the optimality theoretic explanation of the DPBE. Then we present a cognitive model that is based on the optimality theoretic explanation of the DPBE. The hypotheses derived from this cognitive model are first tested in a psycholinguistic experiment involving 75 Dutch children aged 4;1 to 6;3. Then a simulation study is discussed in which the effects of speech rate on the comprehension of sentences with pronouns and reflexives are modeled. In this second study, the performance of a group of children is simulated and compared to the results of the psycholinguistic experiment. The paper concludes with a discussion of the considerations and limitations in using cognitive models to study theories of language acquisition.

DELAY OF PRINCIPLE B EFFECT (DPBE)

A well-established finding in language acquisition research is the observation that, in languages such as English, French and Dutch, children's comprehension of pronouns is delayed in comparison with their comprehension of reflexives (e.g. Chien & Wexler, 1990; Jakubowicz, 1984; Koster, 1993; Philip & Coopmans, 1996; Spenader, Smits & Hendriks, 2009). This phenomenon is called the Delay of Principle B Effect (DPBE). Principle B is one of the two principles of Binding Theory that relate to the adult use and interpretation of reflexives and pronouns (Chomsky, 1981):

(1) a. Principle A: a reflexive must be bound in its local domain.
    b. Principle B: a pronoun must be free in its local domain.

The local domain is defined as the minimal clause containing both the lexical anaphor and a subject. An anaphor is bound when it is co-indexed with and c-commanded by an antecedent.[1] Sentences (2a) and (2b) illustrate the application of Principles A and B:

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

The reflexive *himself* in (2a) can only co-refer with the local subject *the penguin*, in accordance with Principle A, and may not co-refer with another referent. In contrast, Principle B prevents the pronoun *him* in (2b) from co-referring with *the penguin*. Therefore, *him* must co-refer with another

[1] The definition of C-COMMAND used here is: node A c-commands node B if the first branching node of the syntax tree that dominates A, also dominates B.

referent present in the linguistic or extralinguistic context. From the age of 3;0 on, children are able to interpret sentences with reflexives, like (2a), correctly, thus displaying knowledge of Principle A. However, up to the age of 6;6, children show difficulties in the interpretation of pronouns in sentences like (2b) (e.g. Chien & Wexler, 1990). They seem to choose freely between a co-referential interpretation, in which the pronoun co-refers with the local subject, and a disjoint interpretation, in which the pronoun co-refers with an antecedent outside its local domain. Thus, in comprehension, children act as if they only have access to Principle A. Their acquisition of Principle B seems to be delayed.

EXPLANATIONS OF THE DPBE

To explain the Delay of Principle B Effect, several theories have been proposed within a nativist framework (a notable exception is the usage-based account of Matthews, Lieven, Theakston & Tomasello, 2009). In this section, we limit ourselves to two well-accepted theories: the pragmatic account of Thornton & Wexler (1999) and the processing account of Reinhart (2006). Both Thornton and Wexler's and Reinhart's account proceed from a nativist view on language. Hence, they assume that children have knowledge of both Principle A and B, and should in principle be able to apply this knowledge. However, the accounts differ in their explanation of why Principle B is delayed.

Thornton & Wexler (1999) propose that the DPBE is caused by a deficiency in pragmatic knowledge. The starting point for their theory is the observation that in certain special contexts a pronoun may receive a co-referential interpretation, for example when the event being described is unexpected or uncharacteristic. To indicate that such an exceptional co-referential interpretation is intended, speakers stress the pronoun (*Mama Bear is washing HER*; see Thornton & Wexler, 1999: 94), in addition to providing special pragmatic context. Thornton and Wexler argue that children do not yet have sufficient world knowledge and pragmatic knowledge to determine whether the event described by the sentence reflects a typical or atypical situation, that is, to evaluate whether the context licenses a co-referential interpretation. Furthermore, Thornton and Wexler argue that children do not recognize stress on a pronoun as an indication that the speaker intended to express an atypical interpretation. As a result, children accept a co-referential interpretation of a pronoun sentence such as *Mama Bear is washing her*. For adult language users, only a disjoint interpretation is possible for this sentence, because adults do not allow a co-referential interpretation in the absence of stress. So children over-accept co-referential interpretations of pronouns because they are unable to distinguish the contexts that license co-referential interpretations

from the contexts that do not license such interpretations. Children will have to acquire the world knowledge and pragmatic knowledge necessary to disallow a co-referential interpretation of a pronoun in non-exceptional contexts.

Although their account focuses on the comprehension of pronouns, Thornton and Wexler point out that this lack of pragmatic knowledge has ramifications for children's production as well (1999: 95). However, under their account it remains a mystery why children who show difficulties on pronoun comprehension at the same time show adult-like performance on pronoun production (see De Villiers, Cahillane & Altreuter, 2006; Spenader *et al.*, 2009).

In contrast to Thornton and Wexler, Reinhart (2006) argues that children possess all knowledge required for the interpretation of pronouns. The crucial difference between children and adults is that children fail to complete the operation of REFERENCE-SET COMPUTATION. Reference-set computation is an operation that is performed by the parser to choose between multiple interpretations generated by the grammar. The operation is required for determining whether a co-referential interpretation is permitted for a pronoun. For a sentence such as (3), for example, the grammar generates two different derivations: one giving rise to a bound variable interpretation (3a), and one giving rise to a co-referential (3b) or disjoint (3c) interpretation. A co-referential interpretation arises if the two variables x and y both happen to be resolved to the same referent, in this case Lili, whereas a disjoint interpretation arises if x and y are resolved to different referents.

(3) Only Lili thinks she's got the flu. (adapted from Reinhart, 2006: 167)
   a. *Bound variable interpretation*: Only Lili (λx (x thinks x has got the flu))
   b. *Co-referential interpretation*: Only Lili (λx (x thinks y has got the flu) & $y = $ Lili)
   c. *Disjoint interpretation*: Only Lili (λx (x thinks y has got the flu) & $y \neq$ Lili)

The grammar allows the bound variable interpretation (3a) for sentence (3), because the pronoun *she* is not bound within its local domain (cf. Principle B). The grammar also allows the pronoun to be interpreted as a free variable, giving rise to the disjoint interpretation (3c). Whether co-referential interpretation (3b) is allowed, however, must be determined through reference-set computation. Reference-set computation involves the comparison of pairs of derivations and their corresponding interpretations. A co-referential interpretation is allowed for (3) only if this interpretation is different from the bound variable interpretation. If these interpretations are indistinguishable, a co-referential interpretation is not allowed because

it is inefficient to revert back to an interpretation that is ruled out by the grammar through the discourse option of co-reference. With respect to sentence (3), the co-referential interpretation is allowed, because (3a) and (3b) have slightly different meanings. Interpretation (3a) entails that other people do not think that they have got the flu, whereas interpretation (3b) entails that other people do not think that Lili has got the flu. The situation is slightly different for the sentence in (4).

(4) Mama Bear is washing her.
   a. *Bound variable interpretation*: Mama Bear (λx (x is washing x))
   b. *Co-referential interpretation*: Mama Bear (λx (x is washing y) & y = Mama Bear)
   c. *Disjoint interpretation*: Mama Bear (λx (x is washing y) & $y \neq$ Mama Bear)

For this sentence, the grammar (Principle B) disallows the bound variable interpretation (4a), because the pronoun *her* would be bound within its local domain. Although (4a) is disallowed by the grammar, reference-set computation nevertheless requires that a bound variable derivation is constructed and its interpretation is compared with the co-referential interpretation (4b). Because the two interpretations are indistinguishable, the co-referential interpretation is not allowed for sentence (4). Consequently, only the disjoint interpretation (4c) is possible for this sentence.

Reinhart argues that children may be unable to perform this operation of reference-set computation because of working memory limitations. If children fail to complete the operation of reference-set computation, they resort to a guessing strategy and arbitrarily choose between a co-referential and a disjoint interpretation. Other strategies are conceivable as well and are used with other marked forms requiring reference-set computation, such as contrastive stress. Only when children have developed sufficient working memory capacity will they be able to complete the operation of reference-set computation and disallow the co-referential interpretation for pronouns. Because the grammar generates two derivations for pronoun sentences but not for reflexive sentences, reference-set computation is not involved in the interpretation of reflexives. With respect to the production of pronouns, as speakers know which meaning they intend for the utterance, reference-set computation is not involved in production either. This would explain why children are able to produce pronouns correctly from a young age while still having difficulties with the comprehension of pronouns (De Villiers *et al.*, 2006; Spenader *et al.*, 2009).

In the next section, we contrast these theories with an alternative theory: the optimality theoretic account of Hendriks & Spenader (2005/2006),

which assumes that only Principle A is part of grammar and Principle B is a derived effect.

OPTIMALITY THEORY EXPLANATION OF THE DPBE

A third type of explanation of the DPBE is provided by Hendriks & Spenader (2005/2006). They argue that the DPBE is the result of a direction-sensitive grammar, that is, a grammar that has different effects in production and comprehension. Their account is formulated within the framework of Optimality Theory (OT), a linguistic framework that models the relationship between a surface form and its underlying structure by means of optimization from a particular input to the optimal output for that input (Prince & Smolensky, 2004). In the domain of semantics, OT describes the relation between an input form and the optimal meaning for that form (e.g. Hendriks & De Hoop, 2001). Applied to syntax, OT describes the relation between an input meaning and the optimal form for expressing that meaning. OT thus provides an account of linguistic competence with respect to language production (i.e. OT syntax) as well as language comprehension (i.e. OT semantics). In OT, the grammar consists of a set of violable constraints, rather than inviolable rules. For every input, which can be either a form or a meaning, a set of potential outputs, or candidates, is generated. These candidates are evaluated on the basis of the constraints of the grammar. In OT, constraints are as general as possible and hence may conflict. OT resolves conflicts among constraints by ranking the constraints in a language specific hierarchy on the basis of their strength. One violation of a stronger (i.e. higher ranked) constraint is more important than many violations of a weaker (i.e. lower ranked) constraint. The optimal candidate is the candidate that commits the least severe constraint violations. Only the optimal candidate is realized.

Direction-sensitive grammar

For their explanation of the DPBE, Hendriks & Spenader (2005/2006) exploit the fact that an OT grammar is inherently direction-sensitive: The form–meaning relations defined by the OT grammar are not necessarily the same from the speaker's perspective (involving optimization from meaning to form) as from the hearer's perspective (involving optimization from form to meaning) (Smolensky, 1996). This property of OT is a result of the output orientation of the markedness constraints in OT. OT assumes two kinds of constraints. FAITHFULNESS constraints evaluate the similarity between input and output. Because faithfulness constraints pertain to the mapping between input and output, these constraints are direction-insensitive and also apply in the reverse direction of optimization.

An example is the constraint PRINCIPLE A (5), which prohibits reflexives from being locally free. This constraint induces hearers to assign a locally bound interpretation to reflexives and at the same time prohibits speakers from expressing a disjoint interpretation by using a reflexive.

(5) PRINCIPLE A: A reflexive must be bound in its local domain.

MARKEDNESS constraints on forms, on the other hand, reflect a preference for unmarked forms, irrespective of their meaning. Because they pertain to the output only, markedness constraints on form only have an effect when a form must be selected from a set of candidate forms. That is, they only have an effect from the speaker's perspective. An example is the constraint AVOID PRONOUNS. For hearers, this constraint does not have any effect, because for hearers the form is already given as the input. The hearer's task is to select the optimal meaning for this form. Since the constraint AVOID PRONOUNS does not distinguish between potential meanings, it does not have any effect from the hearer's perspective. The constraint AVOID PRONOUNS is part of the constraint hierarchy REFERENTIAL ECONOMY (6). This constraint hierarchy consists of several markedness constraints, of which AVOID REFLEXIVES is the lowest ranked. The hierarchy reflects a preference for less referential content: Reflexives are preferred over pronouns, and pronouns over full NPs.

(6) REFERENTIAL ECONOMY:
Avoid full NPs ≫ Avoid pronouns ≫ Avoid reflexives[2]

In this discussion we limit ourselves to the choice between pronouns and reflexives, and hence only consider the constraint AVOID PRONOUNS. This constraint is violated by any pronoun in the output, and is satisfied by any reflexive in the output. The presence of markedness constraints such as AVOID PRONOUNS can lead to an asymmetry between production and comprehension, as is shown below.

The evaluation of candidates on the basis of the constraints of the grammar can be illustrated with an OT tableau. Figure 1 displays the two comprehension tableaux representing the comprehension of a reflexive (1a) and the comprehension of a pronoun (1b), respectively. The input to a comprehension tableau is a form and the output is the optimal meaning for this form. The constraints are presented in columns in order of descending strength, from left to right. PRINCIPLE A must be ranked higher than AVOID PRONOUNS because otherwise pronouns would never be selected. The relevant candidate outputs (in this case, potential meanings for the input form) are listed in the first column. A violation of a constraint is

[2] A ≫ B means that constraint A is higher ranked, i.e. stronger, than constraint B.

https://doi.org/10.1017/S0305000909990560 Published online by Cambridge University Press

a. OT tableau of the comprehension of a reflexive.

	Input: reflexive	PRINCIPLE A	AVOID PRONOUNS
☞	**co-referential**		
	disjoint	*!	

b. OT tableau of the comprehension of a pronoun.

	Input: pronoun	PRINCIPLE A	AVOID PRONOUNS
☞	**co-referential**		
☞	**disjoint**		

NOTE: Constraint violations are marked with a '*', fatal violations with a '!' and the optimal output with a '☞'.

Fig. 1.

marked with a '*', and a fatal violation with a '!'. The optimal output is marked by '☞'.

When a hearer encounters a pronoun or a reflexive, he has to choose between a co-referential interpretation (first row) and a disjoint interpretation (second row). The co-referential interpretation is the optimal interpretation for a reflexive (Figure 1a), because the disjoint interpretation violates the strongest constraint PRINCIPLE A, whereas the co-referential interpretation satisfies this constraint. When comprehending a pronoun (Figure 1b), PRINCIPLE A is not relevant because it does not define the antecedent possibilities of pronouns. Because AVOID PRONOUNS does not apply in comprehension, both the co-referential interpretation and the disjoint interpretation are optimal candidates according to the grammar. As a result, pronouns are ambiguous. Hence, children might randomly choose one of the two candidate meanings when no contextual clues are available.

Figure 2 shows the tableaux for the production of a co-referential interpretation (2a) and a disjoint interpretation (2b), respectively.

When a speaker wishes to express a co-referential meaning (Figure 2a), the relevant competing candidate forms are a pronoun and a reflexive. PRINCIPLE A does not distinguish between these two candidates, because this constraint allows a co-referential interpretation to be expressed by a reflexive as well as a pronoun. However, AVOID PRONOUNS prefers reflexives over pronouns. Therefore, a reflexive is a better form for expressing a co-referential meaning than a pronoun. On the other hand, a pronoun is the optimal form for expressing a disjoint meaning (Figure 2b), because it satisfies PRINCIPLE A, whereas a reflexive does not.

https://doi.org/10.1017/S0305000909990560 Published online by Cambridge University Press

a. OT tableau of the production of a co-referential meaning

	Input: co-referential	PRINCIPLE A	AVOID PRONOUNS
☞	**reflexive**		
	pronoun		*!

b. OT tableau of the production of a disjoint meaning.

	Input: disjoint	PRINCIPLE A	AVOID PRONOUNS
	reflexive	*!	
☞	**pronoun**		*

NOTE: Constraint violations are marked with a '*', fatal violations with a '!' and the optimal output with a '☞'.

Fig. 2.

In summary, an optimality theoretic grammar is direction-sensitive, because the optimal form–meaning pairs in production are not necessarily the same as the optimal form–meaning pairs in comprehension. Specifically, a pronoun can have a co-referential and a disjoint interpretation according to the grammar, whereas the best form for expressing a co-referential interpretation is a reflexive and the best form for expressing a disjoint interpretation is a pronoun, according to the same grammar. This fits the pattern typically displayed by four- to seven-year-old English and Dutch children, leading to an asymmetry between their production and their comprehension (De Villiers *et al.*, 2006; Spenader *et al.*, 2009).

Explanation for adults' comprehension of pronouns

In contrast to children, adult language users always interpret a pronoun as having a disjoint meaning. According to Hendriks & Spenader (2005/2006), the difference between children and adults is that adult hearers also take into account the perspective of the speaker, whereas children only consider their own perspective. The adult way of interpretation can be modeled as bidirectional optimization (Blutner, 2000). Figure 3 illustrates the serial implementation of bidirectional optimization proposed by Hendriks, Van Rijn & Valkenier (2007).

When an adult hearer encounters a pronoun or a reflexive, he has to determine the optimal meaning for this form. This requires the hearer to optimize in the hearer's direction of optimization: from form to meaning. In addition, however, the hearer must also check whether the selected meaning is indeed expressed by the encountered form. This requires that

https://doi.org/10.1017/S0305000909990560 Published online by Cambridge University Press

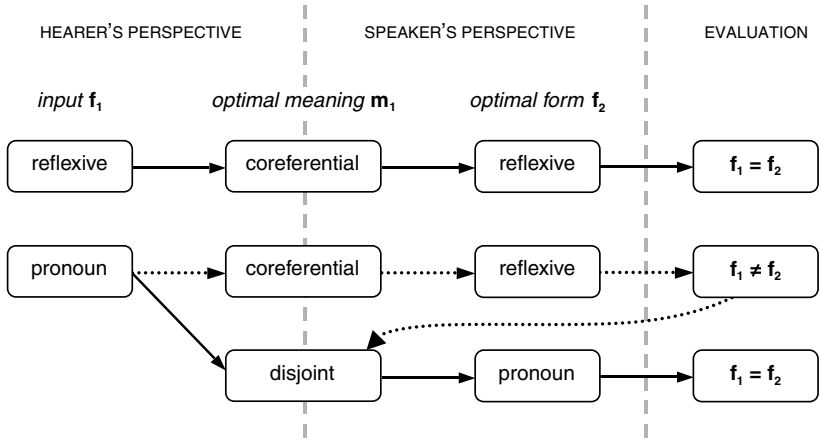

Fig. 3. Taking into account the speaker's perspective in comprehension. The co-referential interpretation for pronouns (represented by the dotted line) is blocked because a co-referential interpretation is best expressed by a reflexive.

the hearer also optimizes from meaning back to form, that is, that the hearer adopts the speaker's perspective. When comprehending reflexives, this process of bidirectional optimization leads to the same result as unidirectional optimization. In both cases, the optimal meaning for a reflexive is a co-referential interpretation. However, when comprehending pronouns, bidirectional optimization leads to a different result. Recall that, from a hearer's perspective, pronouns are ambiguous and can also receive a co-referential interpretation. From the speaker's perspective, however, a co-referential meaning is best expressed using a reflexive. If a hearer were to select the co-referential meaning for the ambiguous pronoun in the first step of optimization, he would find out in the second step of optimization that a co-referential meaning is best expressed with a reflexive. So the resulting form (a reflexive) is different from the encountered form (a pronoun). As a consequence, the co-referential interpretation is blocked for pronouns, and pronouns are only assigned a disjoint interpretation.

In summary, children's pattern of comprehension and production can be explained by unidirectional optimization, which is a formalization of the idea that children only consider their own perspective. The adult pattern can be explained by bidirectional optimization, which is a formalization of the idea that adults take into account the opposite perspective in addition to their own perspective. This OT explanation provides an adequate description of the Delay of Principle B Effect. It can account for the observation that the interpretation of pronouns is acquired later than the interpretation of reflexives. It also explains why children's production of

https://doi.org/10.1017/S0305000909990560 Published online by Cambridge University Press

pronouns may already be adult-like, while their comprehension of pronouns is still poor. Furthermore, the analysis of the DPBE can be generalized to other acquisition delays, either in comprehension or in production. This contrasts with the processing account of Reinhart (2006), which only predicts delays in comprehension.

TESTING LINGUISTIC THEORIES

In the previous sections, we discussed three different explanations for the DPBE: Thornton & Wexler's (1999) pragmatic account, Reinhart's (2006) processing account and Hendriks & Spenader's (2005/2006) OT account. These explanations illustrate the lack of consensus with respect to the cause of the DPBE. An important reason for this lack of consensus is that it is difficult to contrast the theories on the basis of linguistic data alone. The theories mentioned above attribute the DPBE to non-linguistic factors such as a lack of pragmatic knowledge, limited working memory capacity or the inability to take into account another person's perspective. However, without further specification of these non-linguistic factors and how they influence linguistic performance, it is difficult to evaluate these theories. To arrive at a full understanding of linguistic competence, it is therefore essential that theories of linguistic competence are tested in combination with viable theories of pragmatic reasoning, memory, parsing and other cognitive processes. Only then will it be possible to generate precise predictions for linguistic performance that can be empirically evaluated on the basis of experimental data.

A possible way to combine a theory of linguistic competence with theories of cognition and cognitive processes is by embedding the linguistic theory in a cognitive architecture. Cognitive architectures combine several theories of different cognitive subsystems into a single theory of the human cognitive system. A number of architectures have been proposed (e.g. EPIC: Meyer & Kieras, 1997; Soar: Newell, 1990; ACT-R: Anderson *et al.*, 2004) that offer a computational environment in which models can be constructed of the phenomena under study. By constructing a model in the context of an architecture, the model automatically respects the assumptions of the architecture.

Computational simulations are a powerful tool for testing theories since they allow for assessing the completeness of the theoretical account. Also, they make explicit which cognitive processes are required for explaining the phenomenon that is studied. The output of a simulation typically consists of the observed behavior and of estimates of the time it takes to perform the task. Therefore, precise predictions can be generated of human behavior (for a review of language-related computational models, see Dijkstra & De Smedt, 1996).

As we saw in the previous section, OT provides a way to account for children's and adults' linguistic competence with respect to pronouns. However, since OT is a theory of linguistic competence, it does not provide an explanation for the change in optimization mechanism between children and adults. Also, OT does not make any predictions about the time it takes to develop the ability to apply bidirectional optimization, or about the factors that are relevant in developing this ability.

The following section presents a computational cognitive model of the acquisition of pronoun comprehension that is based on the theoretical OT model of Hendriks & Spenader (2005/2006) and is implemented in the general cognitive architecture ACT-R (Anderson *et al.*, 2004). In the model, ACT-R interacts directly with the OT grammar to produce linguistic performance (cf. Hendriks *et al.*, 2007; see Misker & Anderson (2003) for an alternative approach to integrating OT and ACT-R). This is possible because of two important properties of OT: its robustness and its cross-modularity. Because OT is robust and does not pose any restrictions on the input, OT is able to assign an optimal output even to incomplete, dispreferred or ill-formed inputs. Hence, it is able to explain incremental parsing and certain parsing preferences without having to assume a separate parser (e.g. Stevenson & Smolensky, 2006). Furthermore, because OT can be applied to any linguistic domain, OT constraints can be ordered in one large constraint hierarchy. As a consequence, an OT grammar is inherently cross-modular and does not require any interfaces to mediate between different linguistic modules. These two properties allow us to implement the OT grammar (i.e. the constraints and their ranking) directly into the cognitive architecture ACT-R. As we will show, the resulting computational cognitive model generates testable predictions with respect to children's and adults' performance on pronoun comprehension.

COGNITIVE MODELS OF LANGUAGE ACQUISITION

The computational cognitive model we constructed is built in the cognitive architecture ACT-R (Adaptive Control of Thought-Rational; Anderson *et al.*, 2004). ACT-R is both a theory of cognition and a modeling environment. As a theory of cognition, its aim is to explain human cognition and to account for a broad range of data from psychological and neuro-cognitive experiments.[3] It has a modular structure: each of ACT-R's modules is based on smaller theories of cognition. For example, ACT-R contains a theory about retrieving declarative knowledge that is based on Anderson and Schooler's rational analysis of memory (Anderson & Schooler, 1991) and a theory about the processing of auditory stimuli that is

[3] For an overview of existing ACT-R models, see http://actr.psy.cmu.edu/publications.

loosely based on EPIC (Meyer & Kieras, 1997). ACT-R is also a modeling environment that can be used to implement a computational simulation of a specific task. The architecture constrains these simulation models to ensure psychological plausibility. The constraints imposed on the models are based on experimental data and define how information is processed, stored and retrieved within modules, and how information is communicated between modules (Anderson, 2007). Although many decisions have to be made when a linguistic analysis is translated into a computational simulation, mainly related to the non-linguistic aspects of the task, the constraints of the cognitive architecture guide these decisions.

General structure of ACT-R

ACT-R distinguishes several modules that are involved in different aspects of human cognitive functioning. The two main modules of ACT-R are declarative memory and the central production system. Declarative memory contains chunks of factual information. The central production system contains IF-THEN rules. The IF-clause of each production rule specifies a number of conditions that must be met for that production rule to be considered for execution. For example, a production rule that initiates a search in memory for alternative interpretations of a linguistic input is subject to the condition that a linguistic input is available in memory that has not already been fully processed, and that the memory system is currently not in use by another operation. The THEN-clause specifies which actions need to be performed if that production rule is selected for execution (for example, the instruction to initiate the retrieval of a memory element, or to initiate a key press). At each time step, the central production system matches the production rules to the current state of the system, and the most active matching rule is selected for execution. The activation value of production rules reflects the utility of that rule and is an expression of the expected benefits of executing that production rule discounted for the costs associated with that production rule. Elements in declarative memory are ranked on the basis of their activation value. The activation value of declarative memory elements (often called chunks) reflects the usefulness of that chunk in the current context. This activation value is based on a weighted average of the number of prior occurrences of that chunk in general, and the number of prior occurrences of that chunk in the current context.

An assumption of ACT-R that is important for the present study is the assumption that every operation, for example the retrieval of a fact from declarative memory or the execution of a production rule, takes a certain amount of time. The total execution time of a cognitive process is not simply the sum of the durations of all constituting operations, as the different

modules can operate in parallel. However, each module in itself can only perform a single action at a time. Thus, the duration of a process critically depends both on the timing of the serial processes within a module, and on how the different modules interact. To provide specific time estimations for a cognitive process, a computational simulation model can be constructed within the ACT-R system that provides precise predictions when it is run (Anderson *et al.*, 2004).

Modeling unidirectional optimization

In this section, we present an ACT-R model that implements Hendriks & Spenader's (2005/2006) theoretical account of the DPBE. Our computational DPBE/ACT-R model is a refined implementation of Hendriks *et al.* (2007) that enables us to derive more precise predictions related to the DPBE.[4] The main difference between our implementation and the original model (Hendriks *et al.*, 2007) is that our DPBE/ACT-R model is more generic than the original model. Our DPBE/ACT-R model can simulate not only the acquisition of pronoun COMPREHENSION, but the acquisition of pronoun PRODUCTION as well. Moreover, the current implementation allows for more principled timing of the processes involved in linguistic performance (see Lewis & Vasishth, 2005; Van Maanen, Van Rijn & Borst, 2009; Van Rijn & Anderson, 2003) for other approaches modeling temporal aspects of linguistic processing in ACT-R).

In our model, different candidate forms and candidate meanings are implemented as chunks in declarative memory. From a hearer's perspective, there are two possible candidate interpretations for a pronoun or a reflexive: a disjoint meaning and a co-referential meaning. From a speaker's perspective there are two possible candidate forms to express a disjoint meaning or a co-referential meaning: a pronoun or a reflexive. These four different candidates are represented as separate chunks. The optimality theoretic constraints are implemented in terms of the violations they incur. Each constraint is represented as a collection of chunks. The chunks specify for each possible input which candidate outputs violate this constraint. As there are four possible inputs in the current domain, each constraint is represented as four chunks. Production rules define strategies to retrieve forms, meanings and constraints from memory. Figure 4 illustrates the process of finding the optimal meaning. Although the process can be applied to comprehension as well as production, in this study we focus on comprehension.

[4] The code of the discussed ACT-R models can be retrieved from www.let.rug.nl/jacolienvanrij/modelcode.html

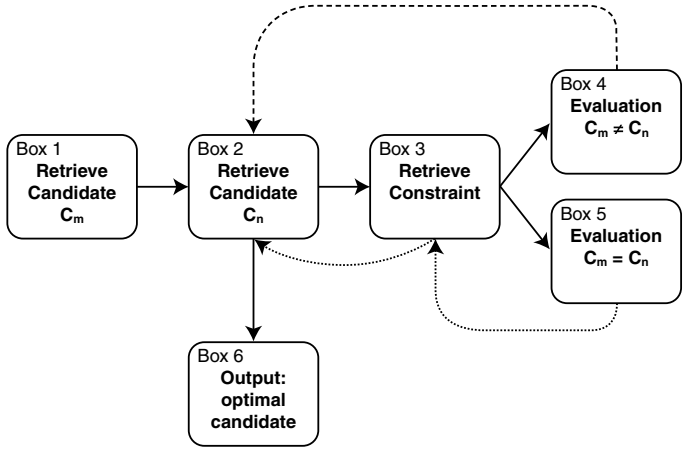

Fig. 4. Structure of the ACT-R model of learning to optimize bidirectionally (adapted from Hendriks *et al.*, 2007).

In comprehension, the input for the DPBE/ACT-R model consists of a pronoun or a reflexive for which the optimal interpretation has to be determined. The first step is to retrieve two of the possible candidates from declarative memory. After the retrieval of the two candidates, a chunk is retrieved representing a constraint that can be used to evaluate the two candidates. Because chunks are ordered based on their activation value, the system will retrieve the highest-ranked constraint first. If one of the candidates violates the constraint, that candidate is replaced by another candidate from declarative memory, 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 between the two candidate meanings, then one of the meanings is randomly selected as the optimal meaning. This process is similar to a recursive optimization process that finds the optimal candidate by evaluating the candidates against the highest-ranked constraint and evaluating the candidates against lower-ranked constraints only if necessary. A complete recursive optimization process would continue until all potential candidates are evaluated. However, the optimization process can be interrupted in the simulations because of cognitive constraints.

In the optimality theoretic analysis of Hendriks & Spenader (2005/2006), PRINCIPLE A is the stronger of the two constraints. When the input is a reflexive, the application of PRINCIPLE A is already sufficient to select the co-referential meaning as the optimal meaning. However, when the input is

https://doi.org/10.1017/S0305000909990560 Published online by Cambridge University Press

a pronoun, the application of both PRINCIPLE A and AVOID PRONOUNS is insufficient to distinguish between the two candidate meanings. Therefore, one of these candidates is randomly selected as the optimal meaning. At this stage, the model performs at chance on pronoun comprehension, whereas it shows almost correct performance on reflexive comprehension. Because the model requires more steps to arrive at an interpretation for the pronoun (first applying two constraints and then selecting a candidate at random) than for the reflexive (just applying one constraint), it is predicted that it takes the model more time to process a pronoun than a reflexive.

Modeling bidirectional optimization

Bidirectional optimization can be thought of as two processes of unidirectional optimization to be performed during on-line sentence processing (Hendriks *et al.*, 2007). In a processing account of bidirectional OT, a straightforward implementation of bidirectional optimization is to have the second step of optimization follow the first step of optimization, as the second step requires the output of the first step. Therefore, we have implemented bidirectional optimization as two unidirectional optimization processes that are performed in sequence. This has already been schematically described in Figure 3.

In the computational simulation new inputs arrive at a fixed rate. As in the situation in which an external speaker determines the speaking rate, the model as a hearer cannot influence this rate. Therefore, the amount of time available for selecting an optimal meaning is limited. Because bidirectional optimization consists of two sequential processes of unidirectional optimization, bidirectional optimization takes more time than unidirectional optimization. In the DPBE/ACT-R model, initially the model can only perform a single process of unidirectional optimization within the limited time and is therefore unable to perform bidirectional optimization during on-line sentence processing. This results in performance that is similar to the performance of young children. As discussed earlier, a unidirectional process results in the correct interpretation of reflexives. However, because two steps of unidirectional optimization are needed for correct pronoun interpretation, performance on pronouns remains at chance level. Only if more time is provided, or when processing efficiency improves, does bidirectional optimization become possible. This account of the DPBE implies that children are in principle capable of applying bidirectional optimization but do not succeed because of limited resources.

In the ACT-R architecture, higher processing efficiency is obtained through the mechanism of production compilation (Taatgen & Anderson, 2002). Production compilation is a learning mechanism that combines two

production rules that are repeatedly executed in sequence into one new production rule. By means of this learning mechanism, cognitive processing becomes much faster, since the new rule has the same functionality as the two production rules before compilation. For example, the model contains two production rules that retrieve candidates from declarative memory. The first production rule requests the retrieval of a candidate on the basis of the received input. The next production rule processes that retrieval and requests another candidate that is not the same as the first retrieved candidate. After repeatedly using these two rules in sequence, the production compilation mechanism creates a new production rule that stores the information of the two candidates at the same time. As this production rule is much more efficient than the two original production rules, this new rule will be preferred by the model. This new rule can again be combined with other production rules in exactly the same way. For example, a production rule may be created that not only retrieves two candidates, but also evaluates these two candidates on the basis of the highest-ranked constraint. Eventually, sentence processing is performed fast enough for bidirectional optimization to succeed within the available amount of time. Note that the time course of learning depends on the frequencies of the input forms, as the compilation of production rules is a function of the number of times a set of rules has been executed in sequence.

In summary, we modeled bidirectional optimization as two sequential processes of unidirectional optimization. If the model cannot perform both steps within the allotted time, pronouns remain ambiguous and a guessing pattern emerges. However, when the model is given more time for interpretation, it will show increased performance on pronoun comprehension. To test this prediction, we performed the psycholinguistic study described in the next section.

STUDY 1: EXPERIMENTAL STUDY

In this section, we present the results of a psycholinguistic study that we carried out to test the predictions of the cognitive model discussed in the previous section. Based on the properties of the DPBE/ACT-R model, we predict that performance of children displaying the DPBE increases when they are given more time for interpretation. We allowed children more time for interpretation by slowing down the speech rate. In contrast to the predicted increase in performance on pronoun interpretation, we predict children's performance on reflexive interpretation to remain level.

METHOD AND MATERIALS

A Truth Value Judgment Task was carried out to test children's comprehension of pronouns and reflexives in Dutch. Participants were shown a

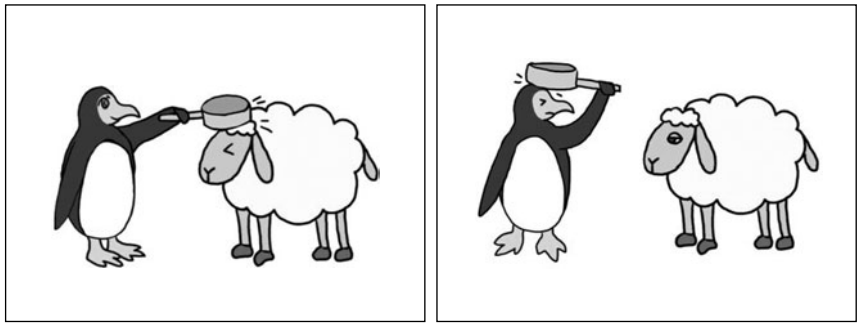

Fig. 5. An example of a picture showing a non-reflexive action (left) and a reflexive action (right).

picture on a computer screen (see Figure 5), and had to judge whether a prerecorded sentence presented to them was a correct description of the picture.

All pictures contained two animals, one of which was depicted as the actor. Both animals were drawn in approximately the same size, to avoid a difference in saliency that may have influenced earlier experiments (for a discussion, see Elbourne, 2005). Each test sentence contained either the reflexive *zichzelf* 'himself' or the pronoun *hem* 'him':

(7) Kijk, een pinguïn en een schaap zijn op de stoep.
De pinguïn slaat hem / zichzelf met een pan.
'Look, a penguin and a sheep are on the sidewalk.
The penguin is hitting him / himself with a pan.'

To allow for the experimental manipulation of processing time, the pronouns and reflexives were always followed by a prepositional phrase. The verbs that were used are *bijten* 'to bite', *kietelen* 'to tickle', *schminken* 'to make up', *wijzen naar* 'to point at', *slaan* 'to hit', *vastbinden* 'to tie up', *zien* 'to see', *schilderen* 'to paint' and *tekenen* 'to draw'. All verbs are typically used for describing an other-directed action, thus avoiding a bias towards a co-referential interpretation (Spenader *et al.*, 2009). The same verbs were used in both speech rate conditions, but the sentences differed in the choice of actors and prepositional objects. Half of the sentences were combined with a matching picture and the other half were presented with a non-matching picture. In addition to the test sentences, four control sentences per condition were included to measure the participants' general performance on the task.

All sentences were prerecorded at normal speech rate (mean speech rate 4·0 syllables per second). Sentences for the Slow Speech Rate condition were then digitally slowed down, while keeping the pitch constant. Using

the software Adobe Audition 1.5, the audio files were stretched 1·5 times, resulting in a reduction of the speech rate with a factor 2/3 (mean speech rate 2·7 syllables per second) (cf. Love, Walenski & Swinney, 2009; Montgomery, 2004; Weismer & Hesketh, 1996). Native Dutch-speaking adults did not report perceiving the slowed-down sentences as disfluent or unnatural. They described the slowed-down sentences as utterances from a slow speaker. During the experiment, the child participants never commented on the speech rate of the test sentences. So there is no indication that slowing down the sentences resulted in an artificial test situation.

Procedure

Every participant was tested in normal and slow speech rate condition. The order of conditions was counterbalanced over participants. Participants were tested individually in a room by two experimenters. A laptop was used to present the pictures and the prerecorded sentences. The sentences started half a second after the picture appeared on the screen. The participants were instructed to press a button with a green smiley face when they considered the sentence a correct description of the picture, and a button with an orange frowning face when they thought the sentence was not a correct description of the picture. Before the test phase, participants practiced the task with two trial items that were presented in the same speech rate as the following condition. They could take as much time as needed to give a response and they were allowed to hear the prerecorded sentence once more when they asked for it. The conditions were presented as blocks of twenty sentences, i.e. eight pronoun sentences, eight reflexive sentences and four control sentences, with a short break in between the two blocks.

Participants

Seventy-five children between the ages of 4;1 and 6;3 were tested. They were all recruited from a Dutch local elementary school. From these 75 children, 13 were excluded from further analysis (4 children were bilingual or non-native Dutch speakers, 5 did not finish the task and 4 responded incorrectly to more than two out of eight control items). The data of the remaining 62 children (35 boys and 27 girls), ranging in age from 4;1 to 6;2, were used for statistical analysis.

Results

Looking at the data of all participants, the percentage of correct interpretations was found to be higher for reflexive sentences than for

pronoun sentences (90% for reflexives, 60% for pronouns; repeated-measures ANOVA: ($F(1,61) = 125{\cdot}968$, $p < 0{\cdot}001$), replicating the results of earlier studies (a.o. Chien & Wexler, 1990; Spenader *et al.*, 2009). Our main question was whether there is a difference in performance between the two speech rate conditions. Statistical analysis of all data showed no significant effect of speech rate on either pronoun comprehension or reflexive comprehension (repeated measures ANOVA: main effect of Speech Rate ($F(1, 61) < 1$); interaction effect of Speech Rate and Expression ($F(1, 61) < 1$)).

However, a closer look at the individual data reveals that a possible effect of speech rate may have been masked, because the participants showed different, and sometimes even opposite, behavior on the task as a function of speech rate. In order to investigate the effect of speech rate on pronoun comprehension in more detail, the participants were classified into different developmental stages on the basis of their performance. This division in groups is crucial for the purposes of our study, as only those children who display the DPBE are predicted to show increased performance with slowed-down speech. First, the criteria used for classifying the participants are described. This is followed by more detailed analyses of the effect of speech rate on the different developmental groups.

Classification of different developmental stages

We divided the 62 participants in our study into three different groups, based on their task behavior.[5] For our classification, we defined (almost) correct performance as more than or equal to 80% correct. Incorrect performance was defined as less than 80% correct.

(i) Children who showed incorrect performance on pronouns as well as reflexives at Normal Speech Rate were categorized as belonging to the INCORRECT PERFORMANCE GROUP ($n = 5$: 3 boys, 2 girls; age 4;3–4;7; mean 4;5).

[5] In earlier papers (Van Rij, Hendriks, Spenader & Van Rijn, 2009a; 2009b), we distinguished four different groups: (i) the No DPBE group ($n = 5$); (ii) the Extra-Linguistic Strategy group ($n = 9$); (iii) the DPBE group ($n = 34$); and (iv) the Correct Performance group ($n = 14$). Participants who were classified as belonging to the Extra-Linguistic Strategy group used the extralinguistic strategy of answering 'yes' to all pronoun mismatch items in both speech rate conditions, while their performance on reflexive items was correct. Participants who were classified as belonging to the DPBE group did not seem to make use of a particular strategy for answering the pronoun items, sometimes giving a correct response while at other times giving an incorrect response, although they showed a general bias to say 'yes'. For simplicity, we combined the Extra-Linguistic Strategy group with the DPBE group in this paper. In addition, we changed the name of the No DPBE group into Incorrect Performance group, because this name better reflects the behavior of its members.

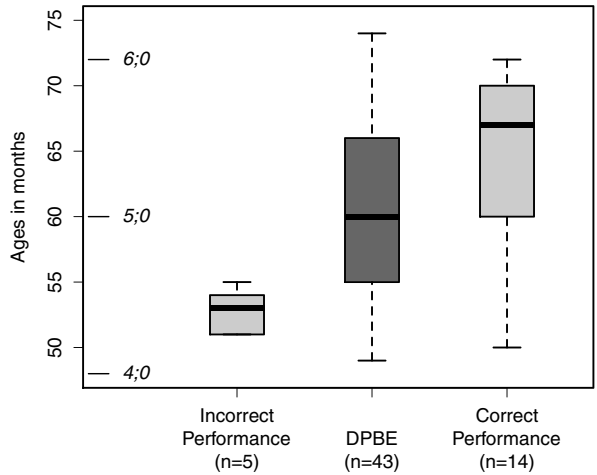

Fig. 6. Mean age in months for the different groups in the psycholinguistic study.

(ii) Children who showed incorrect performance on pronouns but (almost) correct performance on reflexives at Normal Speech Rate were categorized as the DPBE GROUP ($n=43$: 23 boys, 20 girls; age 4;1–6;2; mean 5;1).

(iii) Children who showed (almost) correct performance on both reflexives and pronouns at Normal Speech Rate were categorized as the CORRECT PERFORMANCE GROUP ($n=14$: 9 boys, 5 girls; age 4;2–6;0; mean 5;5).

On the basis of the criteria mentioned above, none of the children showed the fourth conceivable pattern of (almost) correct performance on pronouns but incorrect performance on reflexives. Figure 6 shows the distribution of the ages for the three different groups.

Children's scores were analyzed using (logistic) linear mixed-effect models (Baayen, Davidson & Bates, 2008; Bates, 2005). This type of analysis is more suited for the data than repeated-measures ANOVAs, because several assumptions for using ANOVAs are not met (see Baayen (2008) for a discussion on this topic). As our DPBE/ACT-R model starts out from the situation in which knowledge of the linguistic constraints and their ranking is already in place, we will only discuss the results of the DPBE group and the Correct Performance group.

Results of the DPBE group

Figure 7 shows the mean percentage of correct interpretations of the 43 children displaying the DPBE. The left plot presents the mean performance

https://doi.org/10.1017/S0305000909990560 Published online by Cambridge University Press

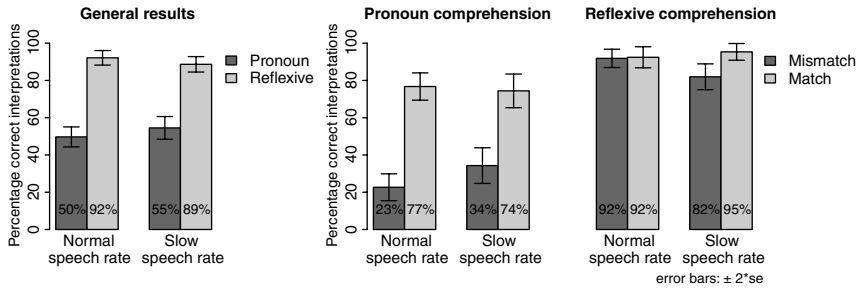

Fig. 7. Mean percentage of correct interpretations of sentences with a pronoun or a reflexive in the two speech rate conditions, for the children showing the DPBE ($n=43$).

on sentences with pronouns and reflexives. The right plot distinguishes between performance on sentences matching the picture and sentences not matching the picture.

Figure 7 shows a clear difference in performance on match items (Normal Speech Rate 77%; Slow Speech Rate 74%) versus mismatch items (Normal Speech Rate 23%; Slow Speech Rate 34%), probably caused by a yes-bias (see also Chien & Wexler, 1990; Grimshaw & Rosen, 1990). To determine the relative contribution of a number of factors on performance, logistic linear mixed-effects models (Bates, 2005) were fitted to the data by Laplace approximation. The factors included as fixed effects were: BLOCK, a between-subjects factor defining the order of presentation of the two conditions; EXPRESSION, a within-subjects factor specifying type of anaphor (pronoun or reflexive); EXPECTEDANSWER, a binary within-subjects factor specifying whether the sentence matched the picture or not (yes or no), and a within-subject binary factor SPEECHRATE specifying speech rate (normal or slow). The interactions between Expression, ExpectedAnswer and SpeechRate were included as well. SUBJECT and a by-subject effect for ExpectedAnswer were included as random effects, to account for individual differences of the participants and for individual answer biases of participants. Separate sets of models were constructed with pronouns and reflexives as dependent variables.

For the sentences with PRONOUNS, we compared the mixed-effects model that included SpeechRate as a factor with the model that did not, to measure whether manipulation of SpeechRate significantly affected the participants' performance. A comparison was conducted on the basis of the models' log-likelihoods (Baayen, 2008). The comparison showed that the model including SpeechRate explains significantly more variance ($\chi^2(2)=7{\cdot}1796$, $p=0{\cdot}028$) than the model without SpeechRate. Thus, slowed-down speech has a significant effect on pronoun comprehension. The following factors contributed to the participants' score on the pronoun

https://doi.org/10.1017/S0305000909990560 Published online by Cambridge University Press

items: ExpectedAnswer (yes) ($\beta = 2{\cdot}964$; $z = 8{\cdot}24$; $p < 0{\cdot}001$), SpeechRate (Slow) ($\beta = 0{\cdot}689$; $z = 2{\cdot}67$; $p = 0{\cdot}008$), Block (Slow Speech Rate condition first) ($\beta = 0{\cdot}242$; $z = 0{\cdot}82$; $p = 0{\cdot}412$) and the interaction between ExpectedAnswer and SpeechRate ($\beta = -0{\cdot}841$; $z = -2{\cdot}25$; $p = 0{\cdot}024$). The yes-bias, as illustrated in Figure 7 (right plot), is reflected in the significant effect of ExpectedAnswer. The positive β-value of Speech Rate (0·689) indicates that slowed-down speech has a positive effect on pronoun comprehension, although this effect is reduced in the match items, as suggested by the negative coefficient of the interaction effect between Expected Answer and Speech Rate (−0·841). Further analysis of the interaction between ExpectedAnswer and SpeechRate confirmed that there is a significant positive effect of slowed-down speech on the mismatch items (23% correct interpretations in the Normal Speech Rate condition versus 34% correct interpretations in the Slow Speech Rate condition (paired $t(42) = 2{\cdot}457, = 0{\cdot}018$). However, no significant difference was found for the match items (paired $t(42) < 1$). The main conclusion from these analyses is that slowed-down speech has a positive effect on pronoun comprehension for children that show the DPBE, as predicted by the DPBE/ACT-R model.

Similar linear mixed-effects models were used to analyze the performance on sentences with REFLEXIVES. Figure 7 shows almost correct performance on reflexive comprehension with match items as well as mismatch items in the Normal Speech Rate condition. However, in the Slow Speech Rate condition, the percentage of correct responses decreases on the mismatch items, but not on the match items, suggesting a small yes-bias. This decrease in performance on mismatch items also suggests a detrimental effect of slowed-down speech.

Again, we compared a model including the factor SpeechRate with a model without SpeechRate. The model including the factor SpeechRate explains significantly more variance ($\chi^2(2) = 9{\cdot}757$, $p = 0{\cdot}008$) than the simpler model. Although this shows that slowed-down speech has a significant effect on reflexive comprehension, the effects are not as straightforward as with pronouns. The effects of the included factors on the reflexive items are: Block (Slow Speech Rate condition first) ($\beta = -1{\cdot}696$; $z = -4{\cdot}29$; $p = 0{\cdot}000$), ExpectedAnswer (yes) ($\beta = 1{\cdot}827$; $z = 2{\cdot}61$; $p = 0{\cdot}009$), SpeechRate (Slow) ($\beta = -0{\cdot}967$; $z = -2{\cdot}79$; $p = 0{\cdot}005$) and the interaction between ExpectedAnswer and SpeechRate ($\beta = 1{\cdot}666$; $z = 2{\cdot}47$; $p = 0{\cdot}013$). The negative estimated effect of SpeechRate (–0·967) might be due to interaction between ExpectedAnswer and SpeechRate (1·666). Further analysis revealed that slowed-down speech indeed has a significant effect only in the mismatch (no) items (paired $t(42) = -2{\cdot}418$, $p = 0{\cdot}020$), and not in the match (yes) items (paired $t(42) < 1$).

In the pronoun analyses, the estimate of Block was not significantly different from zero, but in these reflexive analyses Block has a negative

effect on the percentage of correct interpretations. Children who started the experiment with the Slow Speech Rate condition performed worse on reflexive comprehension in slowed-down speech than children who first participated in the Normal Speech Rate condition. It might be that starting the experiment in the Slow Speech Rate condition triggers other processing strategies, causing additional effects in comprehension. As the effects are more pronounced in the pronoun sentences, a similar effect could be hidden in the variance of that dataset. However, the current data does not allow for testing this.

To summarize, if the child displays the DPBE, slowed-down speech has a positive effect on children's comprehension of pronouns. In contrast, slowed-down speech has a negative effect on children's comprehension of reflexives.

Results of the Correct Performance group

The computational model discussed above predicts that if children are able to take into account both their own perspective and the speaker's perspective under normal conditions, they are also able to do so when they have more time for interpretation. Therefore, the model predicts no effects of speech rate on pronoun or reflexive comprehension in the Correct Performance group. To test this prediction, performance on pronoun comprehension was analyzed, again using linear mixed-effect model comparisons. The factors Block, ExpectedAnswer and SpeechRate were included as fixed effects, as well as the interaction effects of ExpectedAnswer and SpeechRate. In addition, Subject and a by-subject effect for ExpectedAnswer were included as random effects. The factor SpeechRate was found to have a significant effect on pronoun comprehension ($\chi^2(2) = 17{\cdot}450$, $p < 0{\cdot}001$, with as estimated effect of SpeechRate: $\beta = -1{\cdot}618$; $z = -2{\cdot}99$; $p = 0{\cdot}003$). In particular, slowed-down speech has a negative effect on pronoun comprehension. Because the effect of SpeechRate is significant both for mismatch items (paired $t(13) = -3{\cdot}647$, $p = 0{\cdot}003$) and match items (paired $t(13) = -2{\cdot}687$, $p = 0{\cdot}019$), slow speech may have a general negative effect on linguistic performance. Support for this idea comes from the observation that a marginally significant effect is also found for the factor Block ($\beta = -1{\cdot}273$; $z = -1{\cdot}85$; $p = 0{\cdot}064$). Because slow speech is especially detrimental at the start of the experiment, this suggests that the negative effects of slow speech pertain to task performance in general rather than to performance on particular items.

DISCUSSION

The experiment investigated whether children's errors in pronoun interpretation are caused by their limited processing speed. The results show

https://doi.org/10.1017/S0305000909990560 Published online by Cambridge University Press

that slowed-down speech has a beneficial effect on pronoun comprehension, but only if the child displays the DPBE. This supports the hypothesis that children showing the DPBE do not have sufficient time to take into account the speaker's perspective, causing pronouns to remain ambiguous. The results of the children who already perform correctly on pronouns suggest that in other cases slowed-down speech has an overall negative effect on performance, making the positive effects of slowed-down speech in the DPBE group even more striking.

STUDY 2: SIMULATION STUDY

We constructed a computational cognitive model to test whether the mechanism of bidirectional optimization can account for children's behavior in the experiment discussed above. To this end, we combined the DPBE/ACT-R model with a computational model of sentence processing. The resulting model, which we refer to as the Speech Rate model, is able to process incoming sentences on a word-by-word basis. With this model we simulated the performance of a group of child participants on sentences with normal and slowed-down speech rate.

SENTENCE PROCESSING

Words are presented to the model in a serial fashion, with an interval between the consecutive words that is derived from the speech rate. The same sentences are used as in the experiment described above. Two different speech rates were used: a normal speech rate of 4·0 syllables per second (resulting in an inter-word interval of 0·31 seconds) and a slow speech rate of 2·1 syllables per second (inter-word interval 0·62 seconds). To simulate the differences between utterances in naturally occurring speech, normally distributed noise ($m=0$, $SD=0{\cdot}01$) is added to each inter-word interval.

A typical trial commences as follows. As soon as the model detects an audio-event, it focuses its attention on that sound. A word is then retrieved from declarative memory on the basis of the properties of the perceived stimulus. After retrieving the word, its syntactic category is retrieved (for a more extensive description of how concept and lemma information is represented, see Van Maanen & Van Rijn, 2007). After these retrievals, the word's lexical information is attached to the syntactic goal category that represents the syntactic structure of the sentence (see Lewis & Vasishth, 2005). As a complete simulation of parsing is not required for investigating the effects of speech rate on the DPBE, this part of the process is implemented in a similar fashion as in the model of reading and dictating of Salvucci & Taatgen (2008).

As soon as the model identifies, on the basis of the retrieved syntactic category, the current word as a pronoun or a reflexive, the model starts the optimization process described earlier (see also Figure 5). So the model does not wait with the process of bidirectional optimization until the sentence is completed but starts the process of bidirectional optimization immediately when it encounters a pronoun or reflexive.

SELECTING THE RESPONSE

After the sentence is processed, the model has to decide whether the sentence is a correct description of the picture. The interpretation of the picture is given to the model from the onset of the trial, as it was also available on the screen before the participants in the experiment heard the sentence. Therefore, the response of the model depends on the outcome of the optimization process, which is the model's interpretation of the anaphor. If the optimization process results in a single interpretation, the model uses that interpretation in its response. However, if the model cannot settle on a single interpretation, it will randomly select a response (with a 80/20 yes–no distribution to reflect the yes-bias, cf. Chien & Wexler, 1990). Note that this random selection process only takes place when the model cannot settle on an interpretation, that is, when the input is a pronoun and no bidirectional optimization took place. Because the successful use of bidirectional optimization will increase with time, the effect of the yes-bias will gradually decrease. After selecting an answer, the model generates a response by pressing the appropriate button.

MODELING THE ACQUISITION OF BIDIRECTIONAL OPTIMIZATION

Because we assume that bidirectional optimization is in principle available, the model develops the ability to perform this process by mere exposure to sentences with pronouns or reflexives. Hereto, we presented the model with randomly selected sentences containing either a pronoun or a reflexive. By means of production compilation, over time the model learns to perform the required operations quicker and with fewer errors.

To simulate the differences in frequency between pronouns and reflexives in natural language, the model was presented with pronouns in 90% of the training trials and reflexives in the remaining 10%. The model was given about 0·32 seconds to determine the optimal meaning for the input, comparable to the time frame in normal speech. As in earlier work on developmental modeling (e.g. McClelland, 1995; Van Rijn, Van Someren & Van der Maas, 2003), the model was presented with experimental sessions at regular intervals (every 50 trials) to assess the current stage of development. This way, each simulation resulted in thirteen simulated

experimental datasets. During the 'experimental sessions', learning was turned off. This testing scheme was chosen to prevent too much influence of the repeated presentation of the experimental sentences on the outcome.

PERFORMANCE OF THE MODEL ON THE EXPERIMENT

For assessing the performance of the model, we ran the model for sixteen simulations, resulting in 208 simulated datasets. This way, the effect of speech rate is compared over different simulated participants, who received different amounts of training, thus making the dataset comparable to the human dataset discussed earlier. The same criteria were used to classify the simulated participants into different groups. Of the simulated participants, 97 showed the DPBE (mean number of training trials 177, $SD = 156$) and 110 showed correct performance (mean number of training trials 408, $SD = 149$).[6] None of the simulated participants showed similar behavior as the children in the Incorrect Performance group, because the model is already able to perform unidirectional optimization from the start.

Similar to the analysis of the experimental data, we analyzed the performance of the simulated participants who showed the DPBE by fitting separate mixed-effect models on the performance on pronoun and reflexive comprehension. The first model contains a random variable to account for the effects of the different simulated participants, and ExpectedAnswer (yes or no) to account for the introduced yes-bias. The second model contains the same variables, but also contains the variable SpeechRate. A significant difference was found between the mixed-effect models of pronoun comprehension ($\chi^2(2) = 47{\cdot}801$, $p < 0{\cdot}001$), but no difference was found between the models of reflexive comprehension ($\chi^2(2) = 0$). Thus, slowed-down speech has a similar effect on comprehension for the model as for the participants of the experiment. A follow-up analysis on the models' performance on pronoun sentences showed that slowed-down speech did have a beneficial effect on both match (paired $t(96) = 2{\cdot}672$, $p = 0{\cdot}009$) and mismatch trials (paired $t(96) = 5{\cdot}010$, $p < 0{\cdot}001$).

Model fit

Figure 8 shows the fit of the model with the experimental data of the DPBE group on pronoun sentences (Pearson $r^2 = 0{\cdot}96$, $RMSSD = 1{\cdot}74$). The model accounts for the two general trends earlier discussed: the increase in performance on the mismatch items caused by the slowed-down speech,

[6] Only 207 simulated participants are reported (97 DPBE and 110 Correct Performance), because the final experiment of one of the simulations was interrupted, resulting in only 12 experimental datasets for that particular simulation.

https://doi.org/10.1017/S0305000909990560 Published online by Cambridge University Press

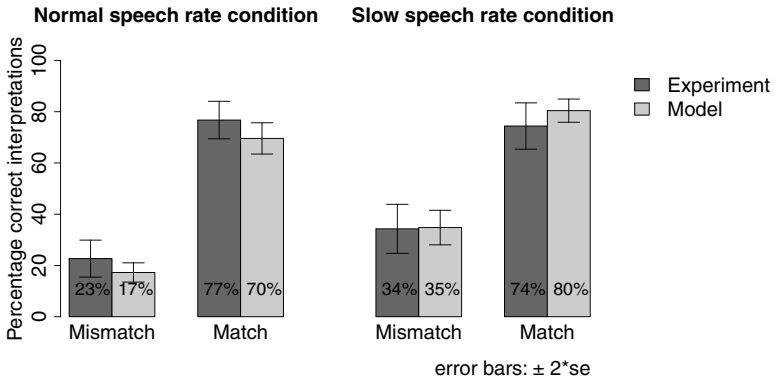

Fig. 8. Comparison between mean percentages of correct interpretations of sentences containing a pronoun in the two speech rate conditions, for children (Experiment: $n=43$) and simulated participants (Model: $n=97$) showing the DPBE.

and the large yes-bias. However, one aspect is not captured. The model predicts a significant increase in performance under slowed-down speech for mismatch AND match items, whereas the experimental data did not show an increase in performance in the match trials. This might suggest that our implementation of the yes-bias is not sensitive enough to capture the details of the child data.

The model's fit on reflexives is not as good as the fit on pronouns. Figure 9 shows the mean percentage of correct interpretations on reflexive comprehension for the model and the experimental data. The model correctly predicts the overall performance on reflexive comprehension for children showing the DPBE. However, the model predicts perfect performance, whereas this level of performance is never found in experimental data with children. Despite these differences, the overall fit (i.e. pronoun and reflexive sentences combined) of the model's performance on the data of the DPBE group is very high (Pearson $r^2=0{\cdot}96$, $RMSSD=2{\cdot}68$).

DISCUSSION

The computational simulation captures the main effects of the psycholinguistic experiment with children, such as the difference in performance on pronouns versus reflexives, the yes-bias and the beneficial effect of slowed-down speech on pronoun comprehension in the DPBE group. However, there are also differences between the model and the experimental data. For example, the model predicts perfect performance on reflexive comprehension in the DPBE group. In contrast, the DPBE children in our experiment did not show perfect performance. Another difference between the behavior of the model and children's performance is that the children in

https://doi.org/10.1017/S0305000909990560 Published online by Cambridge University Press

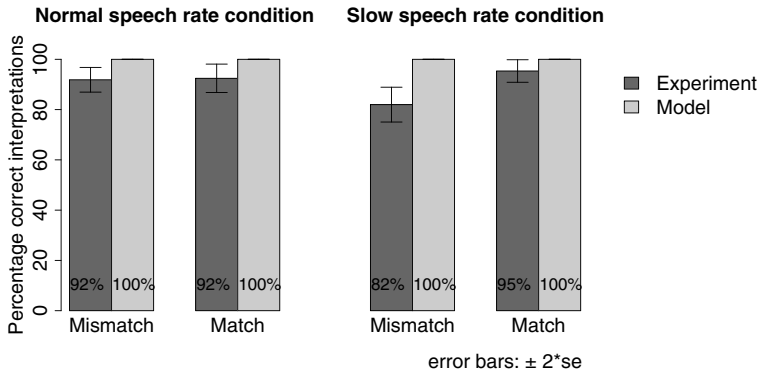

Fig. 9. Comparison between mean percentages of correct interpretations of sentences containing a reflexive in the two speech rate conditions, for children (Experiment: $n=43$) and simulated participants (Model: $n=97$) showing the DPBE.

our experiment showed a significant decrease in performance with reflexives (DPBE group) and pronouns (Correct Performance group) in the slowed-down speech condition. Children probably have to adjust to the unusually slow speech. The model, on the other hand, is not adjusted to normal speech, and as such does not need to readjust. To summarize, although the model does not explain all details of children's performance in the experiment, it does explain the major effects associated with the DPBE as well as adult-like performance on pronoun and reflexive comprehension.

GENERAL DISCUSSION

In this paper we showed how a linguistic explanation of the DPBE that is embedded in a cognitive architecture allowed us to generate and test detailed predictions with respect to linguistic performance. According to Hendriks & Spenader's (2005/2006) optimality theoretic account of the DPBE, pronouns are ambiguous and are disambiguated only if hearers not only select the optimal meaning for the pronoun, but also take into account the speaker's perspective. This allows them to block the co-referential meaning. We modeled this process in the cognitive architecture ACT-R (Anderson *et al.*, 2004). Our DPBE/ACT-R model simulates adult pronoun comprehension as a process consisting of two consecutive steps. The first step involves selecting the optimal meaning for the pronoun, and the second step involves checking whether a speaker would have expressed this meaning with the same form. Our DPBE/ACT-R model predicts that performing the two steps consecutively requires more processing time than performing only the first step. If children are given sufficient time to perform both steps within the available time, they are predicted to be able

https://doi.org/10.1017/S0305000909990560 Published online by Cambridge University Press

to block the co-referential meaning for the pronoun. We tested this prediction by comparing children's comprehension of pronouns at a normal speech rate with their comprehension at a slower speech rate. Our finding confirms the predictions of the DPBE/ACT-R model: slowed-down speech has a significant beneficial effect on pronoun comprehension, but only if the child displays the DPBE.

If the DPBE were caused by children's lack of pragmatic knowledge, as Thornton & Wexler (1999) argue, it remains unexplained how slowing down the speech rate would provide children with the necessary pragmatic knowledge or the ability to use this knowledge to interpret pronouns correctly. Although Reinhart's (2006) explanation of the DPBE in terms of children's limited working memory capacity appears to be related to the explanation presented here, it is unclear how exactly working memory limitations influence children's comprehension, and how this relates to the present findings. It has been argued that slowed-down speech places a greater temporal load on working memory, because information must be retained over a longer duration (e.g. Small, Andersen & Kempler, 1997). If this is true, then slowed-down speech is expected to decrease performance when working memory capacity is limited, in contrast to what Reinhart predicts. However, the results of studies investigating the relation between slowed-down speech and working memory are not very clear. For example, Montgomery (2004) did not find an association between sentence processing at different speech rates and working memory capacity in children. So although Thornton & Wexler (1999) and Reinhart (2006) attribute the DPBE to non-linguistic factors, it is difficult to see how these accounts would explain the present findings. In addition, it remains unclear how these accounts relate to general constraints on cognition, and what predictions they would and would not generate regarding children's and adults' linguistic performance.

The results of the psycholinguistic experiment are predicted by the DPBE/ACT-R model, which was constructed by embedding the optimality theoretic account of Hendriks & Spenader (2005/2006) in the cognitive architecture ACT-R. However, these results do not necessarily follow from the optimality theoretic account in itself. Hendriks and Spenader's optimality theoretic account would also be compatible with an explanation in terms of perspective taking: children may be unable to use bidirectional optimization because they lack the cognitive ability to take into account another person's perspective. In contrast to the explanation implemented in the DPBE/ACT-R model, this explanation would not predict an effect of slowed-down speech, because it is unclear how slowed-down speech would improve children's cognitive skills. Another conceivable explanation of the DPBE that is compatible with an optimality theoretic account is the view that bidirectional optimization is an off-line pragmatic decision process.

This view contrasts with our DPBE/ACT-R model, as we implemented bidirectional optimization as a process that takes place DURING on-line sentence processing. If bidirectional optimization is only performed after completion of the sentence, slowed-down speech is not expected to have any effect on comprehension. In the two speech rate conditions, the same amount of processing time was available at the end of the sentence: Participants in the psycholinguistic experiment could take as much time as needed to give a response in either condition. However, within the sentence, processing time was limited due to the presentation of the next word of the sentence, as the critical word (i.e. a pronoun or a reflexive) was always followed by further sentence material in the form of a prepositional phrase. Therefore, the results of this study suggest that the process of bidirectional optimization is an on-line process.

In addition to a psycholinguistic study, we also performed a simulation study to investigate the predictions of the DPBE/ACT-R model. We built a new cognitive model that also allowed for incremental sentence processing. This model was shown to capture the main effects of slowed-down speech on comprehension that were seen in the psycholinguistic study. For those simulated participants who displayed the DPBE, the cognitive model showed an increase in performance due to slowed-down speech on the comprehension of pronouns, but no effect of slowed-down speech on the comprehension of reflexives. These results support the hypothesis that difficulties with pronoun comprehension are caused by a limited speed of processing, due to which the process of bidirectional optimization cannot be completed.

In our simulations, the process of bidirectional optimization gradually became more efficient as the number of training items increased, because the production compilation mechanism of ACT-R is dependent on frequency of use. As a consequence, the model predicts that repetitive testing of children showing the DPBE on pronoun sentences in binding contexts will result in an increase of their performance on pronoun comprehension (although we did not simulate this in our model). However, we assume that children only start to perform bidirectional optimization for pronoun comprehension when their cognitive and linguistic capacities are sufficiently developed (cf. Case, 1987; Van Rijn *et al.*, 2003). This is reflected in the starting point of our model, according to which children are in principle able to perform bidirectional optimization, but not yet within the limited amount of time. Therefore, we predict that children will only show a positive effect of repetitive testing and slowed-down speech when they are ready to master the process of bidirectional optimization.

Our simulation study also illustrates some of the considerations and limitations in using cognitive models to study theories of language acquisition. First, cognitive models are necessarily simplifications of reality.

Therefore, choices have to be made as to what aspects of the task should be modeled and what aspects can be left unspecified. For example, we chose not to model the sentence-processing component of the model in detail. One of the effects of this choice was that the performance of the model only increased significantly on the comprehension of pronouns at half the normal speech rate. In contrast, the DPBE children already showed an increase in performance at two-thirds of the normal speech rate. This difference is caused by a simplification of the sentence processing component: in the current version of the model, processing a word takes almost all the time that is available before the next word comes in (about 300 of the 320 ms). Hence, not much time is left for bidirectional optimization. To obtain a significant effect of bidirectional optimization, we had to slow down the speech rate more. However, this simplification of the cognitive model did not result in a qualitative difference between the simulation model and the psycholinguistic study, but only in a quantitative difference. It is left for further study whether a more realistic sentence-processing component would lead to better predictions by the cognitive model.

A difficulty in using cognitive models to study language is the possibility that the linguistic theory and the cognitive architecture may employ different or even conflicting assumptions. For example, Optimality Theory, due to its roots in neural network theory, assumes candidates to be evaluated in parallel, and also assumes the constraints of the grammar to apply in parallel. ACT-R, on the other hand, assumes a central processing bottleneck. This implies that only one production rule can be applied at a time. We chose to adopt the ACT-R assumption, since it imposes the strongest restrictions on cognitive processing. Note that this choice is not incompatible with OT per se, as it preserves the input–output relations predicted by the OT grammar as well as the linguistic knowledge constraining these relations, but merely specifies the process by which these input–output relations are obtained. As a result of this choice, in the DPBE/ACT-R model only two candidates are evaluated at a time and the constraints are applied one by one. The hypothesis that children do not have sufficient time to perform bidirectional optimization follows from this particular property of the DPBE/ACT-R model.

A related issue concerns those cases where a particular effect could in principle be explained by the grammar, but also by the cognitive architecture. In language acquisition research, computational models of grammar typically use corpus data as input and observed patterns in the child's speech as output. As a consequence, frequency distribution patterns in the input and output are of crucial importance to the grammar. In a cognitive modeling approach, the grammar may be non-probabilistic because the cognitive model is already sensitive to frequency distributions. For example, our DPBE/ACT-R model was trained on language input which

consisted of 10% reflexives and 90% pronouns, reflecting the unequal distribution of reflexives versus pronouns found in corpus studies of child-directed speech (e.g. Bloom, Barss, Nicol & Conway, 1994). Because the production compilation mechanism of ACT-R is dependent on frequency of use, this unequal frequency distribution resulted in a faster acquisition of bidirectional optimization for pronouns than for reflexives (although the model assigns a correct interpretation to reflexives faster, because its interpretation is not dependent on bidirectional optimization). So cognitive modeling accounts of language acquisition are not incompatible with frequency-based accounts, but rather provide complementary insights. The exact division of labor between grammar and cognitive architecture may be determined by theory-internal considerations as well as empirical observations.

In conclusion, embedding a theory of linguistic competence in a cognitive architecture is a promising new approach to understanding issues in the domain of language. While linguistic theories may offer an adequate account of children's linguistic competence, cognitively informed models are required to test these competence theories empirically. Because cognitive architectures are based on well-founded theories of cognition and guide the construction of computational simulations that allow us to test the performance of a cognitive system under different conditions, they may help us to gain a better understanding of the process of language acquisition.

REFERENCES

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?* New York: Oxford University Press.

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), 1036–60.

Anderson, J. R. & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science* **2**(6), 396–408.

Baayen, R. H. (2008). *Analyzing linguistic data: A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

Baayen, R. H., Davidson, D. J. & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language* **59**(4), 390–412.

Bates, D. M. (2005). Fitting linear mixed models in R. *R News* **5**, 27–30.

Bloom, P., Barss, A., Nicol, J. & Conway, L. (1994). Children's knowledge of binding and coreference: Evidence from spontaneous speech. *Language* (*Baltimore*) **70**(1), 53–71.

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

Case, R. (1987). The structure and process of intellectual development. *International Journal of Psychology* **22**(5), 571–607.

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), 225–95.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (1981). *Lectures on Government and Binding: The Pisa lectures*. Dordrecht: Foris Publications.

De Villiers, J., Cahillane, J. & Altreuter, E. (2006). What can production reveal about Principle B? In K. U. Deen, J. Nomura, B. Schulz & B. D. Schwartz (eds), *The Proceedings of the Inaugural Conference on Generative Approaches to Language Acquisition – North America, Volume 1*, 89–100. Honolulu, HI: University of Connecticut Occasional Papers in Linguistics 4.

Dijkstra, T. & De Smedt, K. (1996). Computer models in psycholinguistics: An introduction. In T. Dijkstra & K. De Smedt (eds), *Computational psycholinguistics*, 3–23. London: Taylor & Francis.

Elbourne, P. (2005). On the acquisition of Principle B. *Linguistic Inquiry* **36**(3), 333–65.

Grimshaw, J. & Rosen, S. T. (1990). Knowledge and obedience: The developmental status of the Binding Theory. *Linguistic Inquiry* **21**(2), 187–222.

Hendriks, P. & De Hoop, H. (2001). Optimality theoretic semantics. *Linguistics and Philosophy* **24**(1), 1–32.

Hendriks, P. & Spenader, J. (2005/2006). When production precedes comprehension: An optimization approach to the acquisition of pronouns. *Language Acquisition* **13**(4), 319–48.

Hendriks, P., Van Rijn, H. & Valkenier, B. (2007). Learning to reason about speakers' alternatives in sentence comprehension: A computational account. *Lingua* **117**(11), 1879–96.

Jakubowicz, C. (1984). On markedness and binding principles. *Proceedings of the North Eastern Linguistics Society* **14**, 154–82.

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

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

Love, T., Walenski, M. & Swinney, D. (2009). Slowed speech input has a differential impact on on-line and off-line processing in children's comprehension of pronouns. *Journal of Psycholinguistic Research* **38**(3), 285–304.

Matthews, D., Lieven, E., Theakston, A. & Tomasello, M. (2009). Pronoun co-referencing errors: Challenges for generativist and usage-based accounts. *Cognitive Linguistics* **20**(3), 599–626.

McClelland, J. L. (1995). A connectionist perspective on knowledge and development. In T. J. Simon & G. S. Halford (eds), *Developing cognitive competence: New approaches to process modeling*, 157–204. Hillsdale, NJ: Lawrence Erlbaum Associates.

Meyer, D. E. & Kieras, D. E. (1997). A computational theory of executive cognitive processes and multiple-task performance: I. Basic mechanisms. *Psychological Review* **104**(1), 3–65.

Misker, J. M. V. & Anderson, J. R. (2003). Combining Optimality Theory and a cognitive architecture. In F. Detje, D. Dörner & H. Schaub (eds), *Proceedings of the Fifth International Conference on Cognitive Modeling*, 165–70. Bamberg: Universitäts-Verlag Bamberg.

Montgomery, J. W. (2004). Sentence comprehension in children with specific language impairment: Effects of input rate and phonological working memory. *International Journal of Language & Communication Disorders* **39**(1), 115–33.

Newell, A. (1990). *Unified theories of cognition*. Cambridge, MA: Harvard University Press.

Philip, W. & Coopmans, P. (1996). The double Dutch delay of Principle B effect. In A. Stringfellow, D. Cahana-Amitay, E. Hughes & A. Zukowski (eds), *Proceedings of the Annual Boston University Conference on Language Development*, 576–87. Somerville, MA: Cascadilla Press.

Prince, A. & Smolensky, P. (2004). *Optimality Theory: Constraint interaction in generative grammar*. Malden, MA: Blackwell.

Reinhart, T. (2006). *Interface strategies: Optimal and costly computations*: Cambridge, MA: MIT Press.

Salvucci, D. D. & Taatgen, N. A. (2008). Threaded cognition: An integrated theory of concurrent multitasking. *Psychological Review* **115**(1), 101–130.

Small, J. A., Andersen, E. S. & Kempler, D. (1997). Effects of working memory capacity on understanding rate-altered speech. *Aging, Neuropsychology, and Cognition* **4**(2), 126–39.

Smolensky, P. (1996). On the comprehension/production dilemma in child language. *Linguistic Inquiry* **27**, 720–31.
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, Volume 2, Linguistic and philosophical implications*, 307–338. 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), 123–55.
Thornton, R. & Wexler, K. (1999). *Principle B, VP ellipsis, and interpretation in child grammar*. Cambridge, MA: MIT Press.
Van Maanen, L. & Van Rijn, H. (2007). An accumulator model of semantic interference. *Cognitive Systems Research* **8**(3), 174–81.
Van Maanen, L., Van Rijn, H. & Borst, J. P. (2009). Stroop and picture-word interference are two sides of the same coin. *Psychonomic Bulletin & Review* **16**(6), 987–99.
Van Rij, J., Hendriks, P., Spenader, J. & Van Rijn, H. (2009a). 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, Volume 2*, 563–74. Somerville, MA: Cascadilla Press.
Van Rij, J., Hendriks, P., Spenader, J. & Van Rijn, H. (2009b). Modeling the selective effects of slowed-down speech in pronoun comprehension. In J. Crawford, K. Otaki & M. Takahashi (eds), *Proceedings of the 3rd Conference on Generative Approaches to Language Acquisition North America* (GALANA, 2008), 291–302. Somerville, MA: Cascadilla Proceedings Project.
Van Rijn, H. & Anderson, J. R. (2003). Modeling lexical decision as ordinary retrieval. In F. Detje, D. Dörner & H. Schaub (eds), *Proceedings of the Fifth International Conference on Cognitive Modeling*, 207–212. Bamberg: Universitätsverlag.
Van Rijn, H., Van Someren, M. & Van der Maas, H. (2003). Modeling developmental transitions on the balance scale task. *Cognitive Science* **27**(2), 227–57.
Weismer, S. E. & Hesketh, L. J. (1996). Lexical learning by children with specific language impairment effects of linguistic input presented at varying speaking rates. *Journal of Speech, Language and Hearing Research* **39**(1), 177–90.