



Cognitive Science 40 (2016) 5–50

Copyright © 2015 Cognitive Science Society, Inc. All rights reserved.

ISSN: 0364-0213 print / 1551-6709 online

DOI: 10.1111/cogs.12250

A Computational Evaluation of Sentence Processing Deficits in Aphasia

Umesh Patil, Sandra Hanne, Frank Burchert, Ria De Bleser,
Shravan Vasishth

Department of Linguistics, University of Potsdam

Received 20 January 2012; received in revised form 6 March 2014; accepted 9 June 2014

Abstract

Individuals with agrammatic Broca's aphasia experience difficulty when processing reversible non-canonical sentences. Different accounts have been proposed to explain this phenomenon. The Trace Deletion account (Grodzinsky, 1995, 2000, 2006) attributes this deficit to an impairment in syntactic representations, whereas others (e.g., Caplan, Waters, Dede, Michaud, & Reddy, 2007; Haarmann, Just, & Carpenter, 1997) propose that the underlying structural representations are unimpaired, but sentence comprehension is affected by processing deficits, such as slow lexical activation, reduction in memory resources, slowed processing and/or intermittent deficiency, among others. We test the claims of two processing accounts, slowed processing and intermittent deficiency, and two versions of the Trace Deletion Hypothesis (TDH), in a computational framework for sentence processing (Lewis & Vasishth, 2005) implemented in ACT-R (Anderson, Byrne, Douglass, Lebiere, & Qin, 2004). The assumption of slowed processing is operationalized as slow procedural memory, so that each processing action is performed slower than normal, and intermittent deficiency as extra noise in the procedural memory, so that the parsing steps are more noisy than normal. We operationalize the TDH as an absence of trace information in the parse tree. To test the predictions of the models implementing these theories, we use the data from a German sentence–picture matching study reported in Hanne, Sekerina, Vasishth, Burchert, and De Bleser (2011). The data consist of offline (sentence-picture matching accuracies and response times) and online (eye fixation proportions) measures. From among the models considered, the model assuming that both slowed processing and intermittent deficiency are present emerges as the best model of sentence processing difficulty in aphasia. The modeling of individual differences suggests that, if we assume that patients have both slowed processing and intermittent deficiency, they have them in differing degrees.

Keywords: Aphasia; Non-canonical sentences; Sentence-picture matching; Eye movements; Computational modeling; Cognitive architecture; Individual differences

Correspondence should be sent to Umesh Patil, Institute of Cognitive Science, University of Osnabrück, 49069 Osnabrück, Germany. E-mail: umesh.patil@uni-osnabrueck.de

1. Introduction

Patients with aphasia frequently experience difficulties in comprehending reversible non-canonical sentences (for reviews see Grodzinsky, 2000; Mitchum & Berndt, 2008). They can comprehend canonical sentences without much difficulty. Irreversible non-canonical sentences are also less problematic for them, mainly because world knowledge and pragmatically based comprehension strategies help in inferring the meaning of the sentence. However, when reliance on syntactic structure is necessary to derive the meaning of the sentence, patients tend to experience comprehension difficulties (Caramazza & Zurif, 1976). This difficulty is reflected in the processing of reversible non-canonical structures like passives, object relatives, object clefts, object questions, and object-topicalized sentences (Mitchum & Berndt, 2008). This deficit has been observed cross-linguistically (Burchert, De Bleser, & Sonntag, 2003; Caplan & Hildebrandt, 1988; Caplan, Waters, & Hildebrandt, 1997; Caplan et al., 2007; Gavarro & Romeu, 2010; Grodzinsky, 2000; Grodzinsky, Piñango, Zurif, & Draï, 1999; Schwartz, Linebarger, Saffran, & Pate, 1987; Thompson, Choy, Holland, & Cole, 2010). In addition to non-canonical structures, patients may also experience difficulties in comprehending other syntactic dependencies like binding relations in reflexives and pronouns (e.g., Choy & Thompson, 2010; Edwards & Varlokosta, 2007; Grodzinsky, Wexler, Chien, Marakovitz, & Solomon, 1993; Thompson & Choy, 2009).

Although syntactic comprehension deficits are often associated with agrammatic Broca's aphasia (e.g., Grodzinsky & Santi, 2008), syntactic impairments have also been documented in Wernicke's and global aphasia (see Caplan et al., 1997; Dronkers, Wilkins, Van Valin, Redfern, & Jaeger, 2004). Grodzinsky (2000) and Grodzinsky et al. (1999) provide overviews of studies that compared aphasics' close to accurate performance on canonical sentences and chance-level performance on non-canonical sentences. In spite of the divergence in the interpretation of chance-level performance in non-canonical structures, the behavioral evidence is quite consistent (see, e.g., Burchert, Hanne, & Vasishth, 2013; Caplan, 2001; Caplan et al., 2007; Caramazza, Capasso, Capitani, & Miceli, 2005; De Bleser, Schwarz, & Burchert, 2006).

1.1. Theories and models of sentence comprehension deficits

Theories of sentence processing deficits in aphasia have been traditionally categorized as either representational deficit accounts or processing deficit accounts (for an overview, see Caplan, 2009). Representational or specific deficit accounts ascribe the impairment to disturbances in underlying syntactic representations—that is, they assume that patients suffer from a breakdown in their knowledge of the grammar; in traditional linguistic terms, the deficit is in their competence as opposed to performance. For example, the Trace Deletion Hypothesis (TDH; Grodzinsky, 1995, 2000, 2006) proposes that aphasic patients have lost the ability to represent traces of syntactic movement. Hence, for non-canonical structures such as passives, they have no information about the theta role of a

moved NP and they need to rely on a default cognitive strategy which assigns the AGENT theta role to the first NP in a sentence. However, the patients' syntactic representation also contains another AGENT, the one assigned to the unmoved subject NP in base position, which forces them to simply guess which of the two NPs is assigned the role of AGENT in a reversible non-canonical sentence. Besides the TDH, other accounts exist that relate syntactic comprehension disorders to disruptions in constructing fully intact syntactic representations (e.g., Beretta & Munn, 1998; Hickok & Avrutin, 1995; Mauner, Fromkin, & Cornell, 1993). These accounts ascribe patients' sentence comprehension deficits to a breakdown in syntactic chain formation. In summary, representational deficit accounts generally equate "breakdown in knowledge of grammar" with the loss in the ability to track head-dependent relations (such as the relation between a noun phrase and its trace) or the ability to keep track of the derivational history of a syntactic structure, where derivational history is construed in classical Chomskyan terms. It is by no means clear that such a breakdown is not due to processing difficulty; that is, the failure to keep track of dependencies (which may be equivalent to deleting traces) may be a processing problem arising due to stochastic variability or less efficient retrieval processes in parsing (Engelmann & Vasishth, 2014).

We turn next to theories that have traditionally been termed processing deficit accounts. These assume that the underlying grammatical knowledge of patients is preserved, but the syntactic processing system is affected by processing (or capacity) limitations. Thus, these theories ascribe patients' difficulties with non-canonical sentences to a processing breakdown in parsing.

There exist various processing accounts that differ in how exactly they conceptualize processing limitations in syntactic parsing: (a) timing deficits; (b) reduction in memory (c) intermittent deficiency; (d) weakened syntax; (e) slow syntax; (f) lexical integration deficit; and (g) lexical access deficits.

Timing deficit accounts have been articulated in some detail in computationally implemented models; the others remain verbally stated theories. We discuss each of these theories below.

1.1.1. Timing deficit

A representative of the timing deficit accounts is the work by Haarmann and Kolk (1991). They proposed a computational model of aphasic language breakdown, called SYNCHRON. This model implements the hypothesis by Kolk and Van Grunsven (1985) that parsing fails in agrammatic aphasics because syntactic representational elements that need to be simultaneously active in working memory are often not coactive because of disturbances in timing due to brain damage. The model is provided with a predefined phrase-structure representation of a sentence and it determines whether the complete construction of this phrase-structure representation is possible, given a set of temporal constraints. The model constructs the phrase-structure representation as a bottom-up chain of retrievals—input words cause the retrieval of word forms, word forms cause the retrieval of associated lexical categories, and lexical categories cause the retrieval of phrasal categories. The retrieval of a phrasal category is possible only if all its constituent

categories are available in memory, which is the *computational simultaneity* constraint in the model.

SYNCHRON assumes that aphasics have a temporal disorder—either the retrieval time, the time required to retrieve an element, is longer than normal, or the memory time, the amount of time a retrieved element remains available for further processing, is shorter than normal. A different way to characterize these constraints is in terms of slowed retrieval and faster than normal decay of items in memory. This temporal disorder disrupts computational simultaneity among elements of the phrasal category, causing parsing failures. Haarmann and Kolk (1991) showed that assuming a temporal disorder is sufficient to model the combined effects of the degree of severity and sentence complexity in agrammatic aphasics described in Schwartz, Saffran, and Marin (1980) and in a replication study by Kolk and Van Grunsven (1985). Although SYNCHRON was successful in modeling aphasic behavior on simpler sentence types, its capabilities are limited due to the absence of a parsing process. It also lacked a mechanism for thematic role assignment, which is a crucial issue in sentence processing.

1.1.2. Reduction in memory

In later work, Haarmann et al. (1997) proposed an enhanced model of aphasic sentence processing, the Capacity Constrained Resource Deficit (CCRD) model. It is implemented in the 3CAPS architecture (Just & Carpenter, 1980) and is derived from the Resource Reduction Hypothesis (Miyake, Carpenter, & Just, 1994). This hypothesis proposes that the impairment in aphasia is an extension of sentence processing limitations in low working memory capacity of unimpaired individuals. CCRD focuses on deriving thematic roles assigned by the verb in the sentence. CCRD is composed of three main subsystems that accomplish thematic role assignment by carrying out three different subtasks in sentence processing: performing lexical access, constructing the parse tree, and mapping thematic roles. The functionality of each component subsystem is achieved through a set of production rules. Production rules temporarily activate the working memory elements that lead to various sentence representations. The rules in the thematic role component use the parse tree representation of a sentence to generate thematic roles between words in the sentence. Once the processing of a sentence is completed, the levels of activation of the working memory elements representing the thematic role bindings are recorded. Sentence comprehension accuracy is indicated by the average activation of these memory elements.

Both storage and computation of information need to draw from available activation. If enough activation is not available, this leads to a breakdown in sentence comprehension. The hypothesis for aphasic patients is that they share a deficit of pathologically reduced working memory capacity. A more complex sentence has higher storage and computational demands, and the reduction in the available activation in aphasics induces a breakdown in processing. The model was shown to reproduce the sentence complexity effect obtained by Caplan, Baker, and Dehaut (1985) across nine sentence types, as well as the interaction between the sentence complexity effect and the degree of severity of aphasia in the data from Kolk and Van Grunsven (1985). All simulations involved

modeling the offline measure of sentence comprehension accuracy by fitting the memory capacity parameter along with several other parameters.

Apart from SYNCHRON and CCRD, other attempts at modeling aphasic sentence comprehension are the HOPE model proposed by Gigley (1986), the UNIFICATION SPACE model proposed by Kempen and Vosse (1989), Vosse and Kempen (2000) and the ACT-R based model proposed by Crescentini and Stocco (2005). While these models differ considerably in their details, they are consistent with the assumption that aphasic sentence comprehension is not a result of any breakdown in the knowledge of the grammar, but rather a deficit in the processing of this knowledge. As observed by Haarmann et al. (1997, p. 82), all these previous models share the common assumptions that “(a) knowledge representation and processing are activation driven, (b) successful sentence comprehension requires the co-activation of certain critical representational elements, and (c) in aphasia, co-activation is disturbed by an immediate or emergent timing deficit.”

1.1.3 *Intermittent deficiency*

Support for intermittent deficiency comes from recent online studies with aphasics. In a self-paced listening study combined with a sentence-picture matching and grammaticality judgement task, Caplan et al. (2007) found normal *online* performance for patients when they provided correct offline responses. In contrast, incorrect offline responses were associated with abnormal online performance. This result is unexpected under the TDH, which does not predict systematic differences in online processes underlying correct and erroneous responses. Caplan and colleagues concluded that patients cannot be suffering from constant impairments in an underlying grammatical structure (e.g., deleted traces), or from a total breakdown in specific parsing operations (e.g., associating a trace with its filler). Instead, they argued that sentence comprehension deficits should better be conceptualized as reflecting intermittent deficiencies in resources necessary for syntactic parsing. These intermittent reductions are then seen in divergent self-paced listening data and lead the patient to end up with an erroneous sentence interpretation. These claims are consistent with the results of former sentence processing studies by Caplan and Waters (2003, 1995). In recent work, Hanne et al. (2011) also provided evidence for the systematic differences between correct versus incorrect parses, pointing to intermittent deficiencies.

1.1.4 *Weakened syntax*

Weakened syntax has received support from sentence comprehension studies using eyetracking (e.g., Choy & Thompson, 2010; Dickey & Thompson, 2009; Dickey, Choy, & Thompson, 2007; Hanne et al., 2011; Meyer, Mack, & Thompson, 2012; Thompson & Choy, 2009, for eye tracking studies on sentence production in aphasia see also Cho & Thompson, 2010; Lee & Thompson, 2011a, 2011b). Through a series of studies in the visual world paradigm, Thompson and colleagues explored patients' online parsing abilities for structures like yes-no questions, wh-questions, and object clefts (Dickey et al., 2007; Thompson, Dickey, & Choy, 2004). For correctly answered wh-questions, aphasics showed the same eye movement patterns as controls—anticipatory eye movements to a potential filler (the moved object) for the gap when they heard the verb. When the offline

response was incorrect, they showed increased looks to the subject competitor toward the end of the sentence. According to the authors, the anticipatory eye movements reflect the participants' incremental, automatic gap filling during sentence comprehension. This suggests that "resolving wh-dependencies was relatively unimpaired in the patients" (Dickey et al., 2007, p. 14). Moreover, because patients' eye movements during correct responses were similar to controls in speed and the overall pattern, the results are inconsistent with a slowdown in aphasics' online processing. Referring to Avrutin (2006), the authors suggested a weakened-syntax view of sentence comprehension disorders, which holds that syntactic representations in aphasia are (largely) undamaged and processing operations such as gap filling function with the same speed as in controls, but the resulting syntactic structures are not strong enough to inhibit competition from other sources (such as competing extra-linguistic heuristics; Dickey et al., 2007). To replicate and extend these initial results, Dickey and Thompson (2006, 2009) evaluated patients' online processing in sentences with two different types of syntactic movement—wh-movement in object relative clauses and NP-movement in passives. Although the results showed that patients can successfully resolve wh-movement dependencies, gap filling in object relative clauses was slightly delayed. These results are different from earlier findings involving wh-questions. Hence, the process of gap filling may be delayed in patients at least for some syntactic structures involving movement. Moreover, this process was disrupted by the late-emerging influence of syntactically unlicensed competitor interpretations. This position is closely related to the idea of slow syntax (Burkhardt, Mercedes Piñango, & Wong, 2003), discussed next.

1.1.5. Slow syntax

As mentioned above, there is some evidence that gap filling is delayed in patients. Further evidence consistent with slow syntax comes from Hanne et al. (2011), who investigated online processing of Broca's aphasics on German reversible canonical (SVO) sentences and their non-canonical counterparts (OVS) using a classical sentence-picture matching task in the visual world paradigm. Online (eye movements) and offline (accuracy and response time) data were collected simultaneously during the task. Patients' accuracy reflected the expected pattern. On average, they performed worse than controls, and comprehension for non-canonical sentences was significantly lower than for canonical sentences. Reaction times were significantly longer in patients than in controls, and non-canonical sentences elicited longer latencies than canonical ones. Fixation patterns showed systematic differences in correct versus incorrect offline responses. For correctly answered trials, patients' eye movement patterns were very similar to controls' (in terms of relative fixation probabilities). For incorrectly answered trials, patients' eye movements were clearly deviant from controls'. Interestingly, patients' eye movement patterns were delayed compared to controls, which is suggestive of a slowdown in online sentence processing. Following Caplan and colleagues, they also came to the conclusion that these preserved processing routines are not always available because of intermittent deficiencies of parsing operations. Thus, the data of Hanne et al. are consistent with the independently motivated idea of a slowdown in syntactic processing in aphasia and with intermittent

deficiency, although there may well be other underlying factors that lead to a slowdown, such as increased uncertainty under noisy memory representations. Using eyetracking with sentence-picture matching, Meyer et al. (2012) investigated the processing of English active and passive sentences in aphasics and age-matched controls. They found that, in active sentences as well as correctly comprehended passive sentences, on average aphasics' eye movements converged to the correct picture a little bit later than in controls. Such delays could be interpreted as slow syntactic processing or, as Meyer et al. interpreted them, as delayed lexical integration, which is discussed next.

1.1.6. Lexical integration deficit

Thompson and colleagues are the main proponents of lexical integration difficulties. In a comprehensive summary of their eyetracking experiments, Thompson and Choy (2009) concluded that sentence comprehension impairments in aphasia are unlikely to be "related to an inability to form, or compute, syntactic representations" (p. 278). They further emphasized that although slight delays in gap filling were seen for some syntactic structures, no delayed syntactic processing was found for patients' dependency resolution in pronominal constructions as well as in *wh*-questions, making a general delay in syntactic computation unlikely. Instead, given the consistent finding of the late-emerging influence of competitor interpretations in patients' incorrect responses, the authors argued for a lexical integration deficit in aphasia, that is, an impairment in the ability to integrate already accessed lexical information into a syntactic or a higher level semantic representation. Work by Hagoort, Brown, and Swaab (1996) and Swaab, Brown, and Hagoort (1997) makes similar claims.

1.1.7. Lexical access deficits

Some authors attribute patients' syntactic difficulties to an earlier stage in language processing—lexical access. Prather, Zurif, Love, and Brownell (1997), using a list priming paradigm, found slower than normal activation of word meanings in Broca's patients. They suggested that this effect directly connects to reduced and/or absent activation effects found at gap sites during real-time sentence processing (Zurif, Swinney, Prather, & Love, 1994). In subsequent studies using a cross-modal lexical priming paradigm with Broca's aphasics, Love and colleagues (Love, Swinney, & Zurif, 2001; Swinney, Zurif, Prather, & Love, 1996) found that priming of a filler at its gap site in syntactic movement structures is not absent, but delayed. The late reactivation of the antecedent was taken as evidence that patients can associate a moved element with a trace; however, this process is pathologically slow. In another study (Love, Swinney, Walenski, & Zurif, 2008), when the speech rate of the auditory input was slowed, patients showed immediate priming effects at the gap site. They also showed delays in lexical activation when a moved NP was first overtly encountered in a sentence. Love and colleagues interpreted their findings as evidence that "the formation of a syntactic dependency involving a moved constituent is selectively vulnerable, not because it's a syntactic operation, but because if lexical reactivation is not accomplished within a normal time frame, a non-grammatical heuristic kicks in to provide a conflicting interpretation" (Love et al., 2008, p. 216). Further,

Ferrill, Love, Walenski, and Shapiro (2012) showed that in patients, lexical activation is slower not only in syntactic structures containing movement dependencies but also in canonical sentences without dislocated NPs. However, it remains unclear why comprehension of canonical structures is less (or even not at all) affected in aphasia.

It is also worth noting that the lexical access account and the lexical integration deficit account may be difficult to disentangle. The lexical access account is in principle distinct from Thompson and colleagues' proposal of a deficit in lexical integration. Thompson and colleagues observed no or only slight delays in the reactivation of antecedents at their gap sites (at least for correct trials) across experiments involving different movement structures (Choy & Thompson, 2010; Meyer et al., 2012; Thompson & Choy, 2009). This absence of an effect, together with the observation of aberrant sentence-end effects of lexical competitors, led Thompson and colleagues to propose an impairment in integrating already accessed lexical information into the syntax or a higher level semantic representation. According to them, this account could explain deficits in comprehending both pronominal and movement structures. However, Thompson and colleagues' findings could also be interpreted as evidence toward a delay in lexical access. For example, in experiment one in Thompson and Choy (2009; see also Choy & Thompson, 2010), patients showed delayed looks to overtly mentioned (unmoved) nouns in a sentence compared to controls. Dickey et al. (2007) observed similar effects at the subject in yes-no-questions (see also experiment two reported in Thompson & Choy, 2009). In addition, Love et al. (2008) pointed out that the auditory sentences used, for example, in Dickey et al. (2007) were spoken at a slower than normal speech rate, which might have confounded the results because the slow input could have compensated for the delay in aphasic's lexical activation. Furthermore, Yee, Blumstein, and Sedivy (2008) showed evidence for reduced lexical activation in Broca's aphasia rather than delays in reaching a certain activation threshold value; this points to impairments in lexical activation levels rather than the time course of this activation. Finally, Blumstein et al. (1998) found no delays but successful priming of a filler at its gap site using a within-modality priming paradigm (auditory–auditory lexical decision). In this study, Broca's patients even patterned with unimpaired participants. Given the diverging results, it is currently still unclear whether and how impairments in sentence comprehension are caused by failures at the stage of lexical access or lexical integration.

We have summarized above the various theories about sentence comprehension deficits in aphasia; but it may be helpful to see the connections, similarities, and differences between these theories by trying to identify some of the key proposals in these theories. We present such a comparison next.

1.2. A comparison of theories of impaired processing, and their relation to theories of unimpaired processing

The theories of sentence processing deficits mentioned above address essentially the same issues that theories of unimpaired populations address (one difference is that the effect of the various determinants of processing difficulty may be amplified in impaired

Table 1

A matrix showing how the models relate to each other along dimensions of the three working-memory-related events—delays, forgetting (or failure to retrieve), mis-retrieval—that have been investigated in sentence comprehension research

Model	Delays	Forgetting	Mis-retrieval
TDH	×	×	
SYNCHRON	×	×	
CCRD		×	
Intermittent deficiency		×	×
Weakened syntax		×	
Slow syntax	×	×	
Lexical integration deficit		×	×
Delayed lexical access	×	×	

populations). This becomes clear when we consider how theories of unimpaired sentence processing that focus on the effect of working memory are characterized; here, comprehension difficulty (i.e., delays) can arise in the integration of lexical items due to decay or interference (Van Dyke & Lewis, 2003), or working memory capacity differences (Just & Carpenter, 1992); dependencies may be forgotten (Frank, Trompenaars, & Vasishth, 2015; Tabor, Galantucci, & Richardson, 2004; Vasishth, Suckow, Lewis, & Kern, 2010), which may or may not lead to parse failure; and there may be occasional mis-parses (Badecker & Straub, 2002; Cunnings & Felser, in press; U. Patil et al., unpublished data; Vasishth, Brüßow, Lewis, & Drenhaus, 2008; Wagers, Lau, & Phillips, 2009), due to interference effects or stochastic noise. As Table 1 shows, classifying the theories mentioned above along these three dimensions—delay, forgetting, and mis-retrieval—demonstrates that while all the theories of sentence comprehension deficits in aphasia try to characterize forgetting in different ways, some try to also develop a theory of why processes are delayed, and why mis-retrievals happen:

1. In TDH, trace deletion has the effect that the relationship between a filler and a gap, which originally was present, is forgotten. Possibly, delays could also occur if the parser carries out extra steps to complete a heuristic strategy to decide on thematic roles for arguments.
2. SYNCHRON implements delays and forgetting by inducing timing deficits that make retrieval slower and that make items in memory decay faster.
3. CCRD induces capacity limitations, which lead to forgetting.
4. Intermittent deficiency, as discussed by Caplan and others, is mainly concerned with occasional forgetting and mis-retrieval, although the precise nature of the deficiency is not defined.
5. Weakened syntax assumes that syntactic structures do not have strong enough representations, which may be a way to implement forgetting.
6. Slow syntax assumes slowed down parsing processes, which would cause delays, and occasional parsing failures.

7. The lexical integration deficit proposal, as developed by Thompson and colleagues, assumes a failure to retrieve a lexical item into a higher level representation; this could be seen as implementing forgetting and possibly also mis-retrieval.
8. The delayed lexical access model assumes that accessing an item in memory in the service of completing a dependency will lead to delays and failures.

Framing existing theories of sentence comprehension deficits in the context of delay, forgetting, and mis-retrieval also highlights the fact that (a) no one theory seems to cover all three events, and (b) one could reclassify theories as either being about delays (more generally, slowed processing), occasional failures to retrieve, or mis-retrievals. As an aside, note that none of the theories have any formalization of prediction cost (e.g., Hale, 2001; Levy, 2008); researchers in aphasia have largely neglected this topic in the past, but it is likely to become a focus of research in the coming years (see Clark, 2012, for an interesting recent attempt using the storage cost metric from the Dependency Locality Theory of Gibson, 2000).

In summary, a useful way to understand the various theories of sentence comprehension deficits is in terms of their attempt to characterize delays, forgetting, and mis-retrievals. The fact that theories of unimpaired sentence comprehension that depend on working memory concepts are also focused on these same events suggests a natural classification of theories of impairment that makes contact with a more general theory of unimpaired processing.

1.3. Deriving quantitative predictions: The importance of computational modeling

As discussed above, the experimental evidence and theoretical proposals in the aphasia literature attribute the comprehension difficulty to deficits at different stages in lexical processing, to processing deficits at the sentence level, or to impaired grammatical representations. How can we derive quantitative predictions of these different theoretical proposals with regard to online and offline processing? A dominant approach has been to derive qualitative predictions of such verbally stated models based on introspection. This methodology, however, has important limitations when the claims are about dynamical processes with non-deterministic properties; in such cases, inferences about model predictions can lead to many surprising errors. For example, Logačev and Vasishth (2015) showed through a set of simulations that a central claim of the Unrestricted Race Model (Traxler, Pickering, Clifton, & Clifton, 1998), that ambiguous sentences are read faster than unambiguous ones (the ambiguity advantage), held only when the finishing times of two attachment processes were very similar to each other; when the difference between the finishing times is large, the ambiguity advantage essentially disappears. Sentence comprehension involves non-deterministic dynamical processes; there is therefore an obvious need to computationally implement the alternative proposals in order to better understand their predictions.

Implementing theories of impairment also has the advantage that it will help us understand what exactly the claim is in various theories. For example, what does intermittent

deficiency mean? There are many ways to realize such a deficiency; whatever the realization, a natural way to understand intermittent deficiency is in the context of a model of unimpaired processing. A further important reason to implement the various proposals is that the theoretical claims should be examined with reference to individual patients—because patients have lesions in different locations, it may be an oversimplification to study average patient behavior (in the case of controls this may make more sense, since they arguably belong to a more homogeneous population, cf. Badecker & Caramazza, 1986; Caramazza & McCloskey, 1988). In order to study behavior at the individual level, models with different parameter settings, and/or with different assumptions, need to be developed.

In recent years, the idea of implementing computational models to inform our understanding of neuropsychological impairments has gained more and more attention (for an overview, see Dell & Caramazza, 2008). As Martin (2006, p. 91) points out, “attempts to model patient performance would lead to important predictions and new empirical tests regarding the nature of the patients’ deficits.” Some models of patient performance do already exist for single-word processing and the respective impairments. Among these are modeling accounts which concentrate on impairments of auditory single-word comprehension and production (Dell, Schwartz, Martin, Saffran, & Gagnon, 1997; Mirman, Yee, Blumstein, & Magnuson, 2011; Rogers et al., 2004), on acquired dyslexia (e.g., Coltheart, 2006; Plaut, McClelland, Seidenberg, & Patterson, 1996; Woollams, Ralph, Plaut, & Patterson, 2007), on optic aphasia (Plaut, 2002), and on single-word processing in bilingual aphasia (Kiran, Grasemann, Sandberg, & Miikkulainen, 2012). Penke and Westermann (2006) presented an implementation of a model designed to account for impairments in morphological processing. Although sentence processing research has made progress in implementing theories computationally (Christiansen & Chater, 2001; Hale, 2001; Vasishth & Lewis, 2006b), research in aphasia has lagged behind.

Previous computational models relating to sentence processing in aphasia have served an important purpose, given the data they considered; but they have several limitations. First, all the computational models mentioned above have attempted to model only offline measures such as sentence comprehension accuracy. Recently, it has been observed that since measures of online processing have the potential to reveal more detail about incremental processing in aphasics, they can help differentiate between various accounts of syntactic deficits in aphasia (Dickey et al., 2007; Hanne et al., 2011). Second, there exists no computational evaluation of the TDH. It is quite possible that a model implementing the TDH can accurately account for aphasic behavior. Finally, most of the approaches have aimed at modeling only aphasics’ responses. A more principled approach would be to first model the responses from unimpaired participants, and then induce the assumed impairment in that model and predict aphasics’ behavior for that task and sentence type. This approach allows us to determine more clearly which properties of the model lead to impairment, and it directly allows us to compare the characteristics of impaired versus unimpaired processing.

We present a set of computational models that aims at modeling offline as well as online responses from controls and aphasics. We compare predictions of the TDH and

three alternative accounts: (a) slowed syntactic processing, (b) intermittent deficiencies of the parser, and (c) a combination of both these impairments; and evaluate the predictions of these two classes of theories of aphasia. We implement the models in a parsing architecture proposed by Lewis and Vasishth (2005). The parsing architecture is grounded within a general computational architecture of human information processing, ACT-R (Adaptive Control of Thought-Rational, Anderson et al., 2004), and it has already been shown to account for several key sentence comprehension phenomena in healthy individuals (Boston, Hale, Vasishth, & Kliegl, 2011; Dillon, Mishler, Sloggett, & Phillips, 2013; Engelmann & Vasishth, 2014; Engelmann, Vasishth, Engbert, & Kliegl, 2013; Patil, Vasishth, & Lewis, 2011; Vasishth & Lewis, 2006a; Vasishth et al., 2008; Wagers et al., 2009; for other parsing architectures grounded in ACT-R, see Dubey, Keller, & Sturt, 2008; and Reitter, Keller, & Moore, 2011). As a target for modeling, we chose data from a visual world paradigm study reported in Hanne et al. (2011). The data consist of the eye movement patterns of aphasics and age-matched controls during a sentence-picture matching task for German reversible canonical and non-canonical sentences. We also model the offline measures, accuracy, and response time, for the sentence-picture matching task. We go beyond looking at aggregate behavior in patients by modeling individuals separately and comparing these to the aggregate behavior of controls. In the next section, we briefly describe the architecture and then present the modeling results.

2. Cue-based retrieval theory

The complete details of the cue-based retrieval theory are described in Lewis and Vasishth (2005); Lewis, Vasishth, and Van Dyke (2006); and Vasishth and Lewis (2006a). Here we only describe the important features of the theory. The theory is derived from (a) the architectural assumptions of ACT-R (Anderson et al., 2004), (b) assumptions about the parsing process based on psycholinguistic evidence, and (c) representational assumptions from syntactic theory.

ACT-R is a generic cognitive architecture consisting of two main components—the declarative memory system and the procedural memory system. Declarative memory holds the contents of long-term memory (semantic and episodic memory), as well as new memories created at run-time during processing. Each element in declarative memory, called a *chunk*, is a set of feature-value pairs. The procedural memory describes procedural knowledge in terms of *production rules*, which are condition-action pairs. These two memory systems also form the core of the architecture. Declarative memory maintains the lexical knowledge, and procedural memory maintains the grammatical knowledge and parsing rules (the control structure) in terms of production rules.

What do production rules correspond to in terms of human sentence parsing? They simply encode the structure that should be built when a word is seen. For example, psycholinguists routinely assume that, upon reading the word “The,” a Determiner Phrase (DP) is built, leading to the expectation that a noun is coming up, and that a verb is also going to appear. This is an informal version of the left-corner parsing algorithm (Aho &

Ullman, 1972). The model simply implements this algorithm for the specific sentences we are interested in modeling (for a more general architecture, see Boston et al., 2011). The incremental syntactic structure created during parsing gets stored in declarative memory as chunks. These chunks are X-bar structures (Chomsky, 1986a) representing maximal projections with features corresponding to X-bar positions (specifier, complement, head) and other grammatical features such as case marking, and person, number, gender agreement. Traces are implemented as co-indexation inside maximal projections. The parse tree is updated at each input word by creating new chunks and attaching them to the existing parse tree representation. Sentence parsing takes place as an iterative sequence of production rule firing, retrieval of memory chunks, and update of the current parse tree.

In addition to the symbolic system (i.e., procedural and declarative memory), the behavior of a model is modulated by a set of subsymbolic computations. These computations impose constraints on the retrieval of chunks from memory and the selection of production rules at each stage. The constraints on retrieval are specified in terms of activation of chunks. The activation value of a chunk determines the probability of its retrieval and the latency of the retrieval. The frequency and recency of retrievals determine the activation of a chunk. Eqs. 1–3 give the details of computing the activation value for each chunk i at every retrieval.

$$A_i = B_i + \sum_{j=1}^m W_j S_{ji} + \sum_{k=1}^p PM_{ki} + \varepsilon_i \quad (1)$$

$$B_i = \ln \left(\sum_{k=1}^n t_k^{-d} \right) \quad (2)$$

$$S_{ij} = S - \ln(\text{fan}_j) \quad (3)$$

Eq. 1 specifies the total activation of a chunk i (A_i), which is the sum of the base-level activation (B_i), the spreading activation received through retrieval cues (first summation component), the activation received due to partial match between retrieval cues and corresponding feature values in the chunks (second summation component), and stochastic noise (ε_i). The base-level activation of a chunk is calculated using Eq. 2. Here, t_k is the time since the k^{th} successful retrieval of chunk i , and d is the decay parameter. The base-level activation function captures the power law of forgetting and the power law of practice for a memory representation (Anderson et al., 2004). The spreading activation that a chunk i receives (first summation component in Eq. 1) is computed using W_j and S_{ji} values. W_j is normally equal to $1/n$, where n is the number of retrieval cues. S_{ji} is the strength of association from an element (typically a retrieval cue) j to chunk i and it is computed using Eq. 3. S is a parameter defining the maximum associative strength and fan_j is the number of items associated with cue j . The strength of association from a cue is reduced as a function of the “fan” of the retrieval cue, resulting in associative retrieval interference. The second summa-

tion component in Eq. 1 specifies the activation received through a partial match. It is computed over p retrieval cues using P and M_{ki} values. P is the match scaling parameter and M_{ki} refers to the similarity between the retrieval cue k and the corresponding value in chunk i .

The mapping from activation A_i to retrieval latency T_i for a chunk i is obtained using Eq. 4. F is the scaling parameter. The higher the activation of the chunk the faster the retrieval.

$$T_i = Fe^{-A_i} \quad (4)$$

Another subsymbolic computation that affects the behavior of a model is the calculation of the utility value for each production. If there are multiple productions that can fire at a certain stage, then the production with the highest utility value is the one that gets selected. The utility value of each production is set depending on the significance of that production in the given context. Just like chunk activations, utilities have stochastic noise added to them; the distribution of this noise is controlled by the utility noise parameter. If there are a number of productions competing with expected utility values U_j , the probability of choosing production i is described by Eq. 5; s is the noise parameter that controls the variance of the noise distribution. Once a production is selected it takes a constant amount of time for it to “fire” and accomplish the actions assigned to it. The value of this constant time can be modified using a parameter called *default action time* (DAT).

$$Probability(i) = \frac{e^{U_i/\sqrt{2}s}}{\sum_j e^{U_j/\sqrt{2}s}} \quad (5)$$

From the perspective of modeling aphasic sentence processing, the ACT-R architecture provides a framework with a set of well-defined constraints on the memory processes involved in parsing sentences, while at the same time offering the flexibility, in terms of possible memory representations and the control structure, for extending the architecture to model other tasks. The set of modifiable parameters makes it possible to extend the framework to impaired sentence processing. The next section describes how we harness this flexibility to extend the cue-based retrieval model of unimpaired sentence processing to the modeling of a sentence-picture matching task with controls and aphasics.

3. Cue-based retrieval model of Hanne et al. (2011)

Here, we present the details of the cue-based retrieval models of the data from Hanne et al. (2011). The study reported in Hanne et al. (2011) was a sentence-picture matching task in a visual world paradigm. Participants listened to German reversible canonical and non-canonical sentences as in (1) while they were presented with two pictures (see

Fig. 1) on the screen. Each trial consisted of listening to a sentence in one of the two word orders. One of the pictures on the screen matched the sentence and the other did not. At the end of the trial, participants were asked to select the correct picture. Participants' response and response time were recorded. Participants' eye movements were also recorded during the whole trial. The experiment was carried out with seven individuals with agrammatic Broca's aphasia and seven age-matched controls without any history of neurological impairment. As a result, Hanne et al. (2011) provides one online and two offline measures of sentence processing of both controls and individuals with aphasia.

1. (a) **Canonical:**
 Der Sohn fängt den Vater
 the_{NOM} son is_catching the_{ACC} father
 "The son is catching the father"
- (b) **Non-canonical:**
 Den Sohn fängt der Vater
 the_{ACC} son is_catching the_{NOM} father
 "The father is catching the son"

In the past, the cue-based retrieval theory has been predominantly used for modeling reading time data (Boston et al., 2011; Dillon, 2011; Dillon et al., 2013; Engelmann & Vasishth, 2014; Engelmann et al., 2013; Lewis & Vasishth, 2005; Patil, Vasishth, & Kliegl, 2009; Vasishth & Lewis, 2006a; Vasishth et al., 2008; Wagers et al., 2009). The predictions of the model are in terms of retrieval times, which have been used as the model's estimate of processing difficulty. We retain this assumption in the current model. The model also makes predictions in terms of retrieval failures and parsing failures. These predictions can be mapped to errors in sentence processing (Vasishth et al., 2008) or comprehension question-response accuracies (Patil et al., 2011). However, there exists no model using the ACT-R architecture which targets online (eye movement patterns)



Fig. 1. Sample pictures used in the Hanne et al. (2011) sentence-picture matching task. For the canonical sentence *Der Sohn fängt den Vater* ("The son is catching the father"), (A) is the correct picture and (B) is the distractor.

and offline (picture matching accuracies and response times) data in a sentence-picture matching task. To model the sentence-picture matching data from Hanne et al. (2011), we must first make new assumptions in terms of a linking hypothesis between predictions of the ACT-R-based theory and cross-modal tasks.

3.1. Linking hypothesis: Implementing the sentence-picture matching task in the ACT-R-based architecture

A sentence-picture matching task typically involves listening to a sentence while gazing at two or more pictures—a target and one or several distractor pictures. In the version used in Hanne et al. (2011), eye movements were recorded during sentence presentation, and at the end of each sentence, participants selected the picture (out of two pictures) matching the sentence. Hence, modeling the sentence-picture matching task entails modeling eye movements across the two pictures during sentence presentation, modeling response accuracy, and modeling response time. To accomplish this, we make a set of new assumptions in the architecture.

First, we assume that, while processing a sentence, the model creates two separate semantic representations of the two pictures on the screen after the first noun is processed. These semantic representations are stored as chunks in declarative memory. The picture chunks are created after the first noun is processed in order to keep the model assumptions as close as possible to the actual experimental procedure. In Hanne et al. (2011), the picture presentation and (auditory) presentation of the sentence began at the same time. Hence, we assume that the creation of the picture chunk does not happen right at the beginning of the sentence. However, as revealed by separate modeling, creating picture chunks at the beginning of the sentence does not lead to significantly different predictions.

As in earlier ACT-R models, at each input word, the parser incrementally updates the partial representation of the sentence. Algorithm 1 and 2 in Appendix A list the steps followed by the model to create syntactic and semantic representations of sentences in (1). The assumption that the semantic representations of pictures are stored as chunks is necessary for two reasons: first, chunks are the only representational units in declarative memory; and, second, the task of matching a sentence to a picture can be accomplished only if they have comparable representations.

Next, we assume that, as the model processes new sentence input, it selects the picture that matches the partial representation of the sentence up to that point. The picture selection is performed by means of a retrieval request for a matching picture chunk. As a consequence of a retrieval request, the two picture chunks receive varying amounts of activation boost. The amount of activation boost received by each chunk depends on its match with the partial sentence representation.

The difference in the activation of the correct and incorrect picture chunk at the time of retrieval determines the probability of fixating on the correct picture. The difference in activation is calculated by subtracting the activation of the incorrect picture from the activation of the correct picture. A positive sign on the difference denotes a fixation on

the correct picture and a negative sign denotes a fixation on the incorrect picture; the higher the value of the difference the more likely it is that the correct picture is fixated. For simplicity, we assume that the mapping between activation difference and fixation probability is linear; this is a reasonable approximation since the probabilities for patients are not far from 50%. This set of assumptions for fixation probabilities in terms of activations is based on the architectural constraints on the model and reflects earlier claims about the link between activations of memory objects and fixation probabilities in the visual world paradigm (Allopenna, Magnuson, & Tanenhaus, 1998; Altmann & Kamide, 2007; Dahan, Magnuson, Tanenhaus, & Hogan, 2001). Allopenna et al. (1998) and Dahan et al. (2001) have proposed a similar linking hypothesis for spoken word recognition in the visual world paradigm. They predicted fixation probabilities for the target and distractors using activations computed from the TRACE model (McClelland & Elman, 1986). Altmann and Kamide (2007) proposed a linkage between language processing and eye movements in the visual world paradigm. On the basis of studies reported in Dahan and Tanenhaus (2005); Huettig and Altmann (2005); Myung, Blumstein, and Sedivy (2006); Altmann and Kamide (2007) proposed that a conceptual overlap between the linguistic input and visual objects results in an increase in activation of the memory representations of those objects. The increase in the activation of an object constitutes a shift in the attentional state of the cognitive system, and this shift in attention increases the probability of eye movements toward the spatial location of that object. We implement the effect of activation boost from linguistic input in terms of multiple retrievals of the picture chunks; in the ACT-R architecture, retrieval is the only process that induces an increase in the activation of an existing memory chunk. Memory retrievals cause an increase in activation through activation spreading and a boost in the base-level activation (see the explanation for Eqs. 1 and 2 in the earlier section).

Finally, we assume that the picture that is retrieved at the end of the sentence is the picture that is finally selected as the response to the sentence-picture matching task, and that the duration between the processing of the first word of the sentence and the retrieval of the picture at the end of the sentence is the response time for the picture matching task.

3.2. *Details of the model of sentence-picture matching*

Modeling the sentence-picture matching task is dependent on accurately creating a representation of the input sentence and retrieving the picture chunk from memory (refer to the algorithms in the Appendix A for the sequence of steps followed by the model to perform the task in Hanne et al., 2011). Creating a syntactic and semantic representation of the sentence involves cue-based retrieval of existing syntactic structures such as a verb phrase (VP) or a noun phrase (NP). Retrievals of these structures and the picture chunks from memory are subject to the constraints decay, interference, and partial match. These constraints arise from Eqs. 1–4. For example, at the input object noun *Sohn* in the OVS sentence, the model sets a retrieval request such as “retrieve a syntactic structure with CATEGORY = DP and CASE = accusative” to retrieve an existing determiner phrase

(DP) marked with accusative case. The accuracy and latency of retrieving the requested DP are affected by the decay of activation (Eq. 2) of the DP, spreading of activation due to the fan of the retrieval cues (the first summation component in Eqs. 1 and 3), the partial match of the cues (the second summation component in Eq. 1), and the noise component (the term ϵ_i in Eq. 1). The picture chunks are retrieved with cues specifying the thematic roles—AGENT and THEME—in the sentence; for example, a picture retrieval request is of the form: “retrieve a picture representation with AGENT = Sohn and THEME = Vater.” The picture retrieval requests that are carried out before the second noun is processed (e.g., steps 7 in Algorithm 1 in Appendix A) have an empty value for one of the features—AGENT or THEME—depending on the word order. As in syntactic structure retrievals, the accuracy and latency of picture retrievals are subject to the same retrieval constraints.

3.3. *Mapping deficit assumptions to the ACT-R models*

Our goal in the current study is to determine what kind of deficit induced in the model for unimpaired individuals can be used to model responses from impaired individuals. For this reason, using the cue-based retrieval architecture discussed earlier, we test some of the proposals put forward by several alternative accounts.

3.3.1. *Models for the Trace Deletion Hypothesis*

Our discussion about the TDH’s predictions for German canonical and non-canonical order is based on an email discussion with Yosef Grodzinsky, dated July 4, 2012. The predictions of the TDH for the comprehension of non-canonical sentences in German and other languages with free word order depend on certain assumptions about the underlying syntactic structure, specifically, regarding the underlying position of the subject. Understanding the predictions of the TDH under different assumptions requires an unavoidable digression into syntactic theory. We attempt below to unpack the assumptions in very general terms, and from first principles.

In syntactic theory, a common assumption is to represent a sentence in terms of a tree with a mother node labeled S (for sentence), and daughter nodes constituting the phrases that make up the sentence. A simple sentence that has a subject and a verb can be represented as an S node with daughters NP (noun phrase) and VP (verb phrase), each of which contains the noun phrase and the verb phrase that make up the sentence. A more sophisticated version of such a tree would add more intermediate nodes that represent, for example, inflectional morphology on the verb. One such node is the inflectional phrase (IP). Thus, a simple monoclausal sentence would have an S node as the mother node, with daughters NP (for the subject) and IP, and the IP would contain the VP as a daughter. A more general syntactic tree would have as mother node not an S node, but rather a CP (complementizer phrase). Syntacticians have also proposed a general theory of the internal structure of such phrases as the NP, VP, IP, and CP; this is generally known as X-bar theory (Chomsky, 1970). Under this proposal, any phrase XP (where X can be instantiated as N, I, P, etc.) is assumed to have two daughters, a Spec(ifier)

node and an intermediate X' (X-bar) node, and the X' node has at least one daughter, the X node, but may have a further adjunct or (in the case of a verb) an argument as a daughter.

For German, one could in principle assume, following standard assumptions in the syntax of English (Chomsky, 1981), that the subject originates inside the Specifier of the IP (Spec-IP), where it is assigned the thematic role of AGENT and case by the finite verb. In canonical order sentences, the subject would then be assumed to “move” to the Specifier of the CP, leaving behind a trace at Spec-IP; the trace would be co-indexed with the NP that it is linked to. In order to arrive at an object-first word order, the object NP, which originates inside the VP, would be assumed to move to the Spec-CP position, leaving behind a trace of its movement inside the VP; importantly, the subject can no longer occupy the Spec-CP position, because the object already occupies that position. This would have the consequence that non-canonical sentences contain exactly one trace that is deleted (the one that is created by movement of the object from inside the VP). As a result of trace deletion, the sentence-initial object would be thematically empty and, hence, interpreted by the default strategy, which assigns an AGENT role to it. If we were to adopt the hypothesis of subjects originating inside the IP, the TDH would predict chance performance because there is an AGENT role (the one assigned to the subject in IP by the finite verb) that competes with the AGENT role assigned by the default strategy to the object.

However, an influential alternative assumption in syntactic theory, called the verb-internal subject hypothesis (VISH, e.g., Koopman & Sportiche, 1991), assumes that the subject is originally located inside Spec-VP (where it receives its theta role) and then moves to Spec-IP for case assignment. The VISH is the only reasonable assumption for German. This becomes evident when we consider an example of VP-fronting, which shows that the subject is contained in the fronted VP: [_{VP} Mädchen geküsst] haben ihn viele ([_{VP} Girls kissed] have him many, “Many girls have kissed him”). This example shows that the subject is fronted along with the verb (we are grateful to Gisbert Fanselow for this example, also see Diesing, 1992; Meinunger, 2000).

Once we adopt the VISH, the TDH makes different predictions. Non-canonical sentences such as (1-b) then contain two traces that are deleted, that is, the trace created by the movement of the object to Spec-CP (as discussed above), and the trace created by subject movement from Spec-VP to Spec-IP. As a consequence, both the sentence-initial object and the subject are thematically empty and their interpretation is subject to the default strategy. This strategy, however, assigns thematic roles to NPs according to their linear positions, and since the object in non-canonical sentences comes first, it would be assigned the AGENT role and the next argument in line, the subject, the THEME role. This would result into a consistent misinterpretation of non-canonical sentences. The TDH, therefore, will predict below chance performance, a prediction that is inconsistent with the predominant reports of chance-level performance with non-canonical sentences in German patients with aphasia (Burchert & De Bleser, 2004; Burchert et al., 2003).

Since there are good independent arguments in favor of VISH in German, a linguistically informed theory such as the TDH would consistently make the wrong prediction. However, it is possible to salvage the TDH by assuming that subjects in non-canonical sentences receive their AGENT role in Spec-IP after they have been moved and are assigned NOMINATIVE case by the finite verb in the head of IP and, as a result, have become visible for theta role assignment (visibility requirement, Chomsky, 1986b). Note that NOMINATIVE case is assigned by the finite verb or its trace in verb-second sentences where the verb is moved from VP via IP to CP. Note also that verb-traces are not subject to trace deletion (cf. Grodzinsky & Finkel, 1998). Deletion of the subject trace in VP, therefore, does not have an impact on the thematic interpretation of moved subjects and the TDH would correctly predict chance performance due to competition between the AGENT role assigned to the moved object by the default strategy and the AGENT role assigned to the subject in IP by the verb (or its chain).

Since the VISH is difficult to abandon for German syntax, we evaluated the TDH using the two variants above that assume the VISH:

(TDH-1) The subject originates in Spec-VP, then moves to Spec-IP and it is assigned AGENT theta role in IP by the trace of the verb.

(TDH-2) The subject originates in Spec-VP, then moves to Spec-IP and it is assigned AGENT theta role through its trace inside VP.

Although the details of the TDH's assumptions are admittedly rather intricate, and assume a great deal of syntactic machinery, the essential element of modeling the TDH lies in assuming that the trace information for the moved element (the subject and/or the object, depending on whether we assume TDH-1 or TDH-2) is deleted. One can question whether deletion of the trace is related to being unable to represent grammatical constructs, or whether this is inextricably related to processing (e.g., the inability to complete a dependency). This is, however, beside the point for us; whatever kind of deficit the TDH in its various versions represents, we can test its consequences within a given parsing architecture.

3.3.2. *Models for slowed processing and intermittent deficiency*

In the literature, slowed processing is considered to be a pathological slowdown in the processing system, and intermittent deficiency a reduction in the resources available for carrying out syntactic, semantic and task-related computations. We do not investigate all the other theoretical proposals in the present work due to limitations of time. Note that it is not at all clear, especially in the verbally stated theories, (a) how to distinguish one theory from another in a computational setting, and (b) what exactly the theoretical claim amounts to. As an example of (a), consider the proposal about delays in lexical access, and lexical integration failure. The lexical-level delays could be a consequence of a slightly slower syntactic parser, which is not slow enough to be detected in experimental data, but is slow enough to indirectly cause a reduction in the lexical-level activation of items in memory; and in an architecture like ACT-R, a reduction in lexical-level activation can also lead to failure in lexical integration. Since these theories do not make any

detailed commitments as regards the parsing algorithm, the timing of parsing events, and the underlying constraints on retrieval, these are effectively free parameters in these theories, and many, many different permutations of commitments about the underlying parser are consistent with these theories. Slowed processing and intermittent deficiency are interesting for us because they have been suggested as plausible explanations for the data (Hanne et al., 2011) that we model in the present paper, and because they have a natural realization in the computational architecture we used for modeling.

Consistent with the processing accounts, we assume that there is no impairment in the grammatical knowledge of aphasics. This means that the set of production rules (the procedural memory) is the same across models for controls and aphasics. Since it is possible that only one of the two assumptions, slowed processing and intermittent deficiency, is enough to explain the data from Hanne et al. (2011), we evaluated these assumptions using three separate models:

(M1) Only slowed processing

Slowed processing is implemented in terms of a higher value for the DAT parameter. DAT controls the amount of time required for one production rule to fire and hence the amount of time required to processing.

(M2) Only intermittent deficiency

Intermittent deficiency is implemented in terms of a higher value for the *utility noise* parameter. The utility noise value determines the variance in the utility values of productions. The utility value of productions determines the probability of selecting a particular production when multiple productions are competing to fire.

(M3) Slowed processing and intermittent deficiency

This model is affected by both factors.

Note that it is possible to evaluate alternatives to intermittent deficiency. We are hoping to carry out such a comprehensive model comparison in future work.

3.4. Estimation of parameter values

For controls, all ACT-R parameters, except for the three mentioned below, were set to the values that have been used in the earlier models from Lewis and Vasishth (2005), Vasishth and Lewis (2006a), Vasishth et al. (2008), and Patil et al. (2011). The value of the *maximum difference* parameter was changed from the earlier models and it was set to its default ACT-R value; this parameter defines the range of the scale of similarity between a retrieval cue and the corresponding value of a chunk in declarative memory. The value of the DAT parameter was lowered from its value used in earlier models; DAT specifies the amount of time needed for each production rule to fire. Finally, the value of the *latency factor* parameter was also lowered; this factor works as a scaling parameter in the equation for retrieval latency (see Eq. 4 above). All parameter values were adjusted (by using a brute force search) to optimize the fit for the sentence-picture matching task (measures of fit are described below). Table B1 lists parameter values in earlier models and those used here for controls.

For aphasics, we estimated values for the *utility noise* and DAT for each patient; other parameters had the same value as in controls. Differences in patients' lesions and also in their responses call for estimating parameters for each individual separately (Dilkina, McClelland, & Plaut, 2008; Nickels, Biedermann, Coltheart, Saunders, & Tree, 2008). We estimated the individual parameter values only for the offline data from aphasics. Once parameter values were estimated for the offline data, predictions for the eye movement data were generated without any changes to these values. The parameters for each patient were estimated by doing an exhaustive search (Nievergelt, 2000) through a range of parameter values for utility noise (the range: 0.01–0.22) and DAT (range: 0.01–0.10). The search chose the best parameter values using the normalized root-mean-square deviation measure—the difference between the values predicted by the model and the observed data. The parameter values estimated for each subject are listed in Table B4 in Appendix B and the corresponding normalized root-mean-square deviations are listed in Tables 2–4.

Table 2

The normalized root-mean-square deviation between observed and predicted response accuracies across five models of aphasia

Participant	TDH-1	TDH-2	M1	M2	M3
P1	0.66	1.55	1.37	0.13	0.08
P2	0.7	2.2	1.33	0.14	0.07
P3	0.29	1.17	0.93	0.11	0.15
P4	0.62	0.59	1.17	0.42	0.4
P5	0.66	1.55	1.41	0.13	0.09
P6	0.12	0.63	0.67	0.22	0.21
P7	0.73	1.14	1.43	0.29	0.26
Mean	0.54	1.26	1.19	0.21	0.18

Note. For TDH-1 and TDH-2, the predictions are derived from the same parameter values as those for controls. For M1, M2, and M3, the predictions are derived from the best parameters for each patient.

Table 3

The normalized root-mean-square deviation between observed and predicted response times across five models of aphasia

Participant	TDH-1	TDH-2	M1	M2	M3
P1	1.14	1.13	0.35	1.13	0.36
P2	1.23	1.23	0.47	1.21	0.49
P3	0.83	0.81	0.4	0.82	0.39
P4	1.08	1.07	0.33	1.07	0.33
P5	1.62	1.59	0.5	1.59	0.5
P6	0.9	0.9	0.36	0.9	0.37
P7	2.55	2.54	0.45	2.54	0.46
Mean	1.34	1.33	0.41	1.32	0.41

Note. For TDH-1 and TDH-2, the predictions are derived from the same parameter values as those for controls. For M1, M2, and M3, the predictions are derived from the best parameters for each patient.

Table 4

The averaged normalized root-mean-square deviation between observed and predicted response accuracies and response times across five models of aphasia

Participant	TDH-1	TDH-2	M1	M2	M3
P1	0.9	1.34	0.86	0.63	0.22
P2	0.97	1.72	0.9	0.68	0.28
P3	0.56	0.99	0.66	0.46	0.27
P4	0.85	0.83	0.75	0.75	0.36
P5	1.14	1.57	0.96	0.86	0.3
P6	0.51	0.77	0.52	0.56	0.29
P7	1.64	1.84	0.94	1.41	0.36
Mean	0.94	1.29	0.8	0.76	0.3

Note. For TDH-1 and TDH-2, the predictions are derived from the same parameter values as those for controls. For M1, M2, and M3, the predictions are derived from the best parameters for each patient.

For model M1, we estimated the values for the DAT parameter and for model M2, we estimated the values for the utility noise parameter. For model M3, we used the newly estimated values of both the parameters simultaneously (see assumptions for M1, M2, & M3 in the preceding section). As an alternative, we also checked the predictions of M3 by estimating both parameters again for M3; this did not lead to substantial changes in the predictions.

No parameters values were estimated for TDH-1 and TDH-2 because these two models do not assume any impairment in processing time or activation.

In the current model, we also adjusted an ACT-R parameter that was not used in the previously published models, the utility values of productions. Note that the utility value and utility noise value are different parameters. The utility value is associated with each production and can be different for different productions. By contrast, the utility noise value is globally defined for the procedural memory system and it determines the variance in the utility value of each production.

As described earlier, the utility value of a production determines the probability of selecting that production. The data from Hanne et al. (2011) (see Fig. 2) and other psycholinguistic studies (Ferreira, 2003; Matzke, Mai, Nager, Rüsseler, & Münte, 2002) show that non-canonical sentences are harder to process even for healthy individuals. The difficulty in processing non-canonical sentences is reflected in slow response time and increased error in the sentence-picture matching task. We adjusted the utility values of productions that assign thematic roles to noun phrases in the sentence. We assume here that the error in processing non-canonical sentences is a consequence of error in assigning thematic roles. When the structural information in the parse tree is ignored, the parser fails to correctly mark the AGENT and THEME roles of the sentence, which, in turn, leads to an error in comprehending the sentence. We estimated the utility values of the productions that process nouns in the sentence. There were two alternative productions that could process nouns—one that uses the syntactic information and the other that does not.

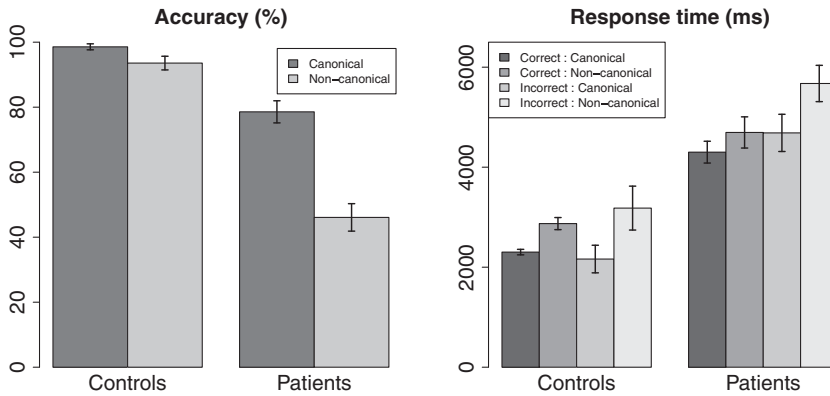


Fig. 2. Data from Hanne et al. (2011); accuracy and response time for the sentence-picture matching task. The error bars in the response time plot show 95% confidence intervals.

For each simulation, one of the two productions gets selected depending on its utility value. The probability of selecting a production i , given its utility value U_i , is calculated using Eq. 5. We set the utility values of these productions such that non-canonical sentences were processed with the accuracy observed for controls in the Hanne et al. (2011) data. The utility values that we used for these productions are listed in Table B2. These values were estimated *only* for the data from controls and then kept constant while modeling the data from aphasics. This means that the utility values of these productions were the same for controls' model and all aphasics' models (TDH-1, TDH-2, M1, M2, and M3). Next, we report the predictions of these models. The predictions for both offline and online responses are generated by averaging across 1,000 runs of each model for each condition.

3.5. Modeling results

3.5.1. Accuracy

Here, we compare the sentence-picture matching accuracy (the accuracy panel in Fig. 2) with the predictions of each model of aphasia (the accuracy panel in Fig. 3). The model TDH-1 captures chance-level performance (close to the 50% mark) in aphasics for non-canonical sentences, but for canonical sentences it predicts accuracy to be as high as for controls. TDH-2 does not capture the pattern in the accuracy data for either condition—it predicts below chance performance in non-canonical sentences and high accuracy for canonical sentences. M1 (slowed processing) also does not capture the pattern in the accuracy data for either condition—it predicts high accuracy values for both sentence types. In contrast, models M2 (intermittent deficiency) and M3 (slowed processing and intermittent deficiency) are able to reproduce the accuracy patterns in aphasics' data quite well. These models predict chance-level performance for non-canonical sentences; for canonical sentences, they predict above chance performance which is not as high as in controls. We return to a comparison between models M1–M3 below, when we consider individual differences.

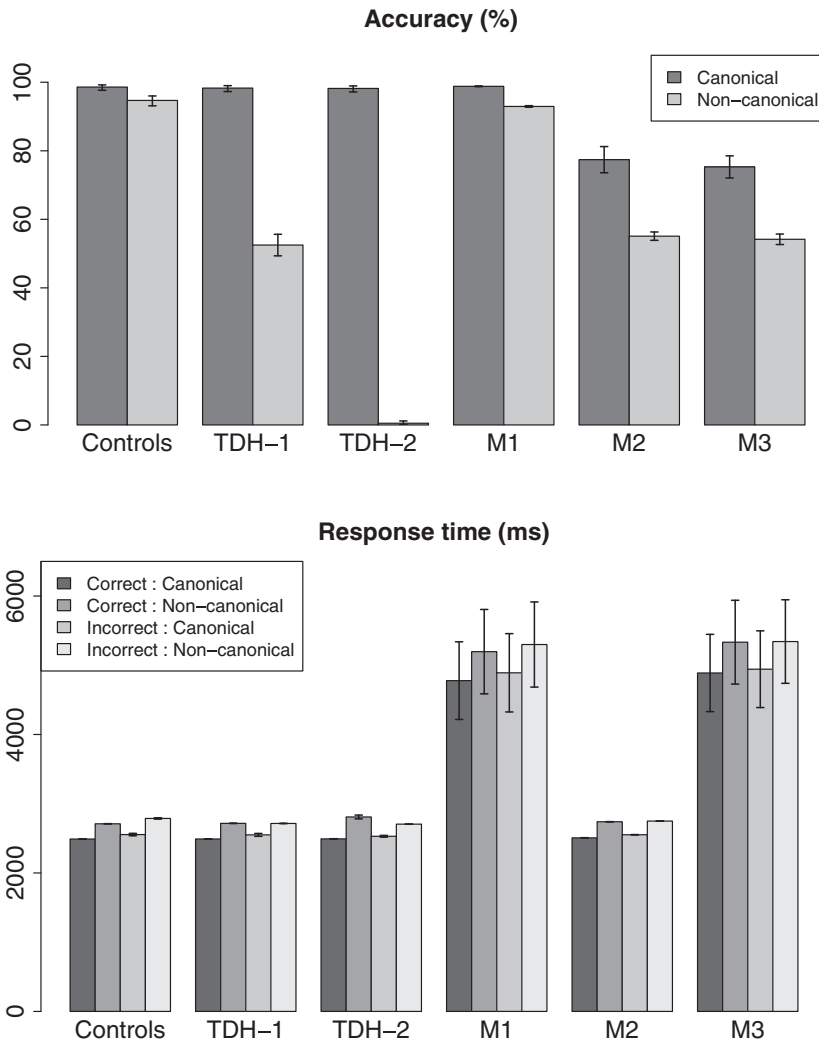


Fig. 3. Accuracy and response time predictions of the model for controls and the models for patients—TDH-1, TDH-2, M1, M2, and M3. The predictions of accuracy are in terms of percentages of retrievals of the correct picture at the end of the sentence. The predictions of response time are equal to the models’ processing time starting from the first word of the sentence until the retrieval of a picture at the end of the sentence. The error bars show 95% confidence intervals. For accuracy plots for controls, TDH-1 and TDH-2, the confidence intervals, are calculated using the Clopper-Pearson (or the “exact”) method since the predicted accuracy values are the same for all participants, unlike M1, M2, and M3 where each individual has a different parameter value and hence different accuracy values.

3.5.2. Response time

Next, we compare the response time data for the sentence-picture matching task (see the response time panel in Fig. 2) with the predictions of each model of aphasia (see the response time panel in Fig. 3). For the response time data, all models capture the pattern

of slow response times for non-canonical sentences in correct and incorrect responses compared to the responses for canonical sentences. The important effect that was observed in aphasics' data was that all of their responses are slower than controls' responses. This effect is captured only by M1 and M3; TDH-1, TDH-2, and M2 predict response times similar to those of controls. Although none of the models can capture the fact that, in the patients' data, non-canonical correct response times are about the same as canonical incorrect response times, this effect does not appear to be robust; a subsequent experiment by Hanne and colleagues with a comparable design did not replicate this effect (Hanne, Burchert, De Bleser, & Vasishth, 2014). The wide confidence intervals for the predictions of M1 and M3 resulted from the assumption that each participant has a different value of the DAT parameter (the assumption of slowed processing); for details regarding separate model fits for each patient, see the section on individual differences in the deficits below.

3.5.3. *Eye movements*

The data and predictions for eye movements for controls and aphasics are illustrated in Figs. 4 and 5. The eye movements show percentage of fixations on the correct picture at various points in the sentences—"NP1," "verb," "NP2," "silence." For canonical sentences, "NP1" is the grammatical subject and "NP2" is the grammatical object, and vice versa for the non-canonical sentences. "Silence" is the time after presentation of the spoken sentence; during this time participants still look at both pictures until they respond by pressing a button to select the correct picture. Controls' data consisted of correct responses in the sentence-picture matching task, and since they performed close to ceiling, this constitutes most of their data. Aphasics' data were partitioned into correct trials (those that led to a correct response) and incorrect trials (those that led to an incorrect response). We model the eye movement data using the linking hypothesis proposed earlier. According to this hypothesis, the models carry out retrievals of picture chunks from memory based on the semantic representation of the sentence at various stages of the input. The models' predictions about the percentages of fixations on the correct picture are derived from the difference in activations of the correct and incorrect pictures at the time of these retrievals.

Controls' eye movements for both sentence types show a gradual increase in fixations on the correct picture. The model for controls reproduces this pattern in the data (see plots (A) and (B) in Fig. 4); although the model seems to perform better than humans toward the end of the sentence (in the "silence" region), this is merely a consequence of the linear linking function assumed between activation differences and fixation probabilities. We also estimated a more sophisticated non-linear linking function; but for the number of data points being modeled, this would have been an overfitted linking function (cf. "Summary of results and discussion" section for more details about a possible non-linear linking function).

For non-canonical sentences, controls start to fixate on the correct picture earlier, at the "verb"; the model does not capture this early certainty. Aphasics, in their correctly answered trials, show similar behavior in eye fixations—gradual increase in fixations on the correct picture. For canonical sentences, their responses are delayed—they fixate on

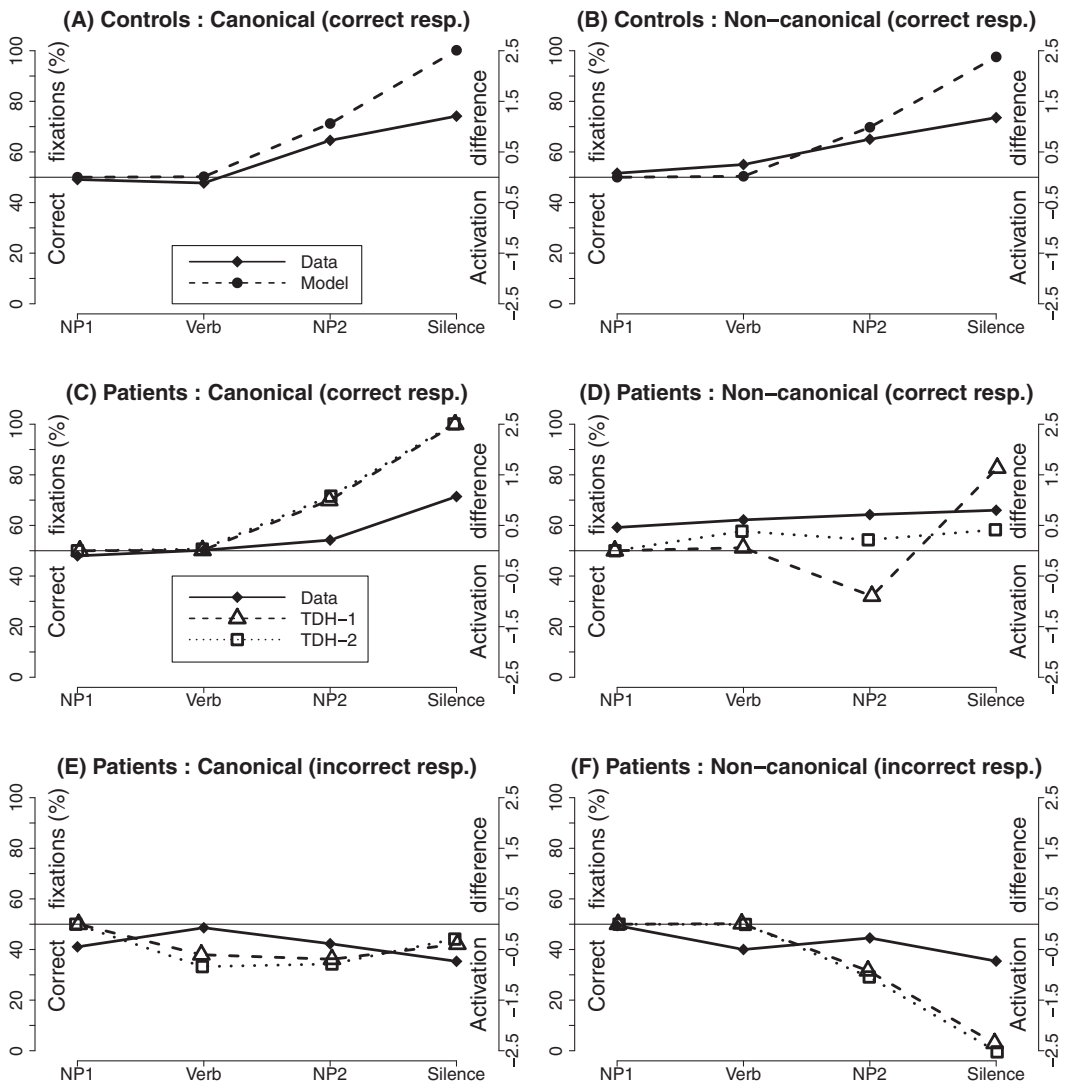


Fig. 4. Eye movement data from Hanne et al. (2011) and the predictions of the model for controls and the models TDH-1 and TDH-2 for patients. The x-axis shows different regions in the sentences. NP1 and NP2 denote the two noun phrases in the sentences. For sentences in canonical word order, they denote subject and object respectively, and for sentences in non-canonical word order they denote object and subject, respectively. The y-axis on the left-hand side of each plot shows percentages of correct fixations observed in the data and the y-axis on the right-hand side shows differences in the activations of the two pictures in the models. An activation difference of zero is aligned with the 50% mark in the fixation percentages; the horizontal line represents this mark. Note that in panel D, even though TDH-2 appears to predict the fixation proportions correctly, these predictions are not reliable because TDH-2 seldom processes non-canonical sentences correctly.

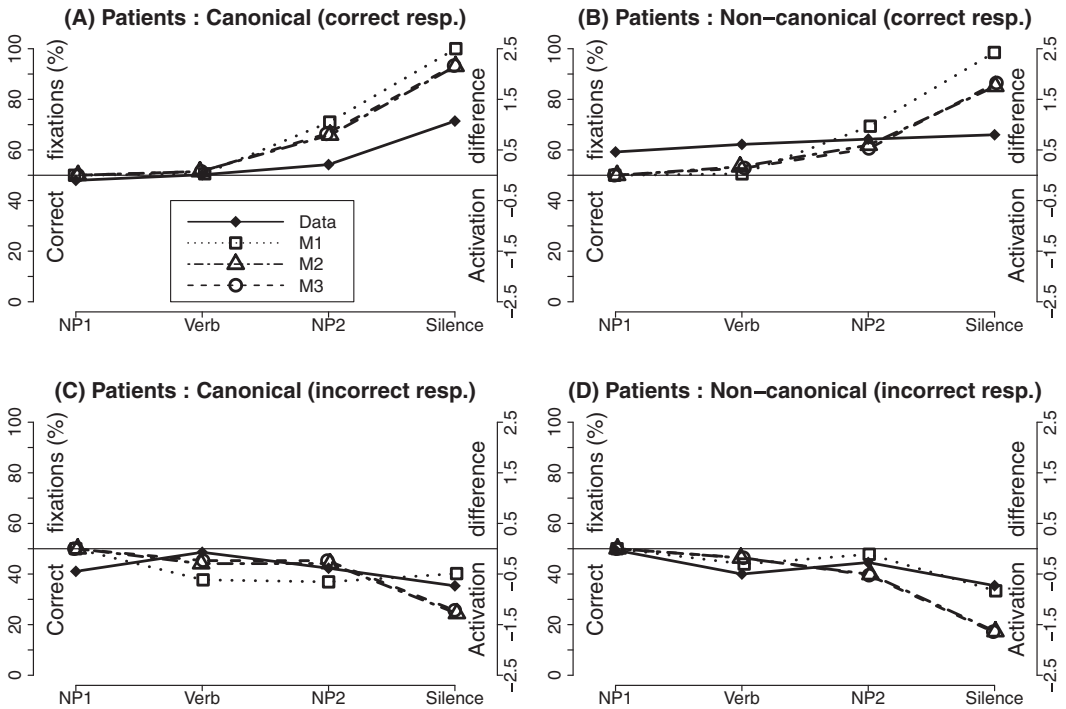


Fig. 5. Eye movement data from Hanne et al. (2011) and the predictions of models M1, M2, and M3. The x -axis shows different regions in the sentences. NP1 and NP2 denote the two noun phrases in the sentences. For sentences in canonical word order, they denote subject and object, respectively, and for sentences in non-canonical word order they denote object and subject, respectively. The y -axis on the left-hand side of each plot shows percentages of correct fixations observed in the data, and the y -axis on the right-hand side shows differences in the activations of the two pictures in the models. An activation difference of zero is aligned with the 50% mark in the fixation percentages; the horizontal line represents this mark.

the correct picture only in the “silence” region, whereas in non-canonical sentences, just like controls, they are certain about the correct picture from the “verb” region onward.

All models for aphasics show similar behavior in correctly answered canonical sentences (see plot (C) in Fig. 4 and plot (A) in Fig. 5). They capture the pattern of gradual increase in correct eye fixations. However, all models show early certainty (at “NP2”) in fixating the correct picture and they consistently show a higher percentage of correct picture fixations than patients’ data. For non-canonical sentences, TDH-1 predicts that patients remain uncertain about the correct picture till the “verb” region, and fixate on the incorrect picture more often in the “NP2” region; in the “silence” region, the patients are predicted to fixate on the correct picture (see plot (D) in Fig. 4). This behavior of TDH-1 stems from initially using the subject-first strategy and then using the trace of the VP to assign AGENT theta role to the second NP (cf. assumptions for TDH-1 above and the Algorithm 2 in Appendix A). TDH-2 predicts an early bias (at “NP1”) toward fixating the correct picture and this bias remains until the “silence” region, but this pattern is not

reliable because TDH-2 predicts that patients would process non-canonical sentences correctly very rarely (see the accuracy panel in Fig. 3); the pattern seen for TDH-2 in non-canonical sentences is based on a very small proportion of simulation trials. In contrast to the TDH implementations, models M1, M2, and M3 predict a more consistent and similar behavior across regions in non-canonical sentences (see plot (B) in Fig. 5). These models fixate on the correct picture from the “NP2” region onward, and predict a higher number of correct picture fixations toward the end. However, in these non-canonical sentence trials, they fail to capture the early certainty (at “NP1”) in aphasics’ fixations.

For incorrect trials of aphasics, the main pattern observed in the data was that the proportion of fixations on the correct picture was mostly below the 50% mark (see plots (E) and (F) in Fig. 4 and (C) and (D) in Fig. 5). All the models capture this behavior, for the most part. The only exception is model M1 in the “NP2” region for non-canonical sentences, but model M1 seldom processes non-canonical sentences incorrectly, which means that a very small proportion of simulation trials lead to this prediction.

3.5.4. Individual differences in the deficits

As mentioned earlier, we estimated the two parameters—utility noise and DAT—for each patient separately using their offline responses. This is because aphasics’ responses show high variance between participants (see the data in Table B3 in Appendix B), implying that the impairment in aphasics varies between individuals to a greater extent than in controls. The parameter estimates for each patient are listed in Table B4 in Appendix B and the corresponding normalized root-mean-square deviations (NRMSD) are listed in Tables 2–4; the smaller the NRMSD value, the better the model fit.

For model simulations represented as a vector *sim* and observed values as a vector *obs*, both of length *n*, the NRMSD is calculated using Eq. 6. The normalization renders the measure of fit comparable, independent of scale.

$$\text{NRMSD} = \frac{\sqrt{(\sum(sim - obs)^2/n)}}{\max(obs) - \min(obs)} \quad (6)$$

Table 2 shows that the models M2 and M3 outperform M1 and the TDH models, and Table 3 shows that the models M1–M3 outperform the TDH models. But across these two tables no one M* model seems to clearly outperform the other two. The superior performance of M3 over M1 and M2 becomes clear only when we take the mean NRMSD scores over accuracy and response time. Note that since the RMSD values are normalized, they are on the same scale; and since NRMSD is a normalized standard deviation, it is reasonable to take the mean of the by-participant NRMSDs to get an estimate of the mean quality of fit of each model for both dependent measures. The averaged NRMSD is shown in Table 4; it is clear from this table that M3 is the best model when both accuracy and response time fits are considered.

As mentioned above, the estimated parameter values (see Table B4 in Appendix B) that lead to the NRMSD values reflect the variability between patients—the values for

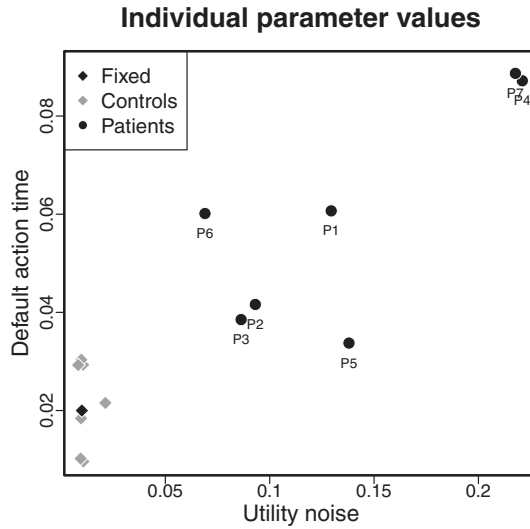


Fig. 6. The parameter values estimated for different participants. The parameter along the x-axis, utility noise, represents intermittent deficiency, and the parameter along the y-axis, default action time, represents slowdown in processing. Each black circle represents a patient (the patient IDs are printed below the circles) and each gray diamond represents a control participant. Note that the values estimated for controls are only for the sake of comparison between the two groups; the predictions for controls were generated with a fixed set of parameter values for all controls, which are denoted by a black diamond in the plot.

aphasics fall in a wider range than for controls. Fig. 6 shows the range of values estimated for the two parameters; we also estimated values of each control (this was done only to facilitate comparison of the different amounts of variability within each of the two groups; the predictions for controls reported earlier were generated with a fixed set of parameter values for all participants). The values for controls form a tighter cluster than those for aphasics. The figure also shows that aphasics have different degrees of intermittent deficiency and slowdown in processing: While patients P1 and P6 are almost equally affected by slowdown in processing (due to similar values for the DAT parameter), they are affected differently by intermittent deficiency (due to dissimilar values for the utility noise parameter); on the other hand, while P1 and P5 are almost equally affected by intermittent deficiency, they are affected differently by slowed processing.

By contrast, the fits to the eyetracking data of the different models provide no useful information on the relative quality of fits; this is clear from Figs. 4 and 5, and also from an inspection of NRMSD values (across regions, conditions, and patients) for each model: TDH-1: 2.8, TDH-2: 2.77, M1: 2.75, M2: 2.76, and M3: 2.76.

3.5.5. Summary of results and discussion

The NRMSDs for accuracies and response time, taken together, suggest that M3, which assumes slowed processing and intermittent deficiency, is the best model among the candidates considered.

Model M3 captures the main effect of participant type in the accuracy data: Aphasics are predicted to have significantly lower accuracy than controls. It captures the main effect of word order: Non-canonical sentences result in lower accuracy (*chance level*) in the sentence-picture matching task. It captures the main effect of participant type in the response time data: Aphasics are predicted to have higher response times than controls. And it also captures the main effect of word order: Response times for non-canonical sentences are higher. For the eye movement data, the model M3 captures the divergent eye movement patterns in correct versus incorrect responses. However, M3 fails to model the early looks to the correct picture that was seen in the non-canonical order with aphasics.

An important conclusion is that models assuming only slowed processing or only intermittent deficiency or trace deletion fail to capture several of the crucial patterns in the aphasics' data: TDH-2 and M1 fail to model the chance-level performance in non-canonical sentences; TDH-1, TDH-2, and M2 fail to model the high response times in aphasics; TDH-1 fails to model the eye movements in the correctly answered non-canonical trials. The demonstration that TDH-1 fails to capture aphasics' online behavior in non-canonical structures has important implications. From the offline predictions, it appears that TDH-1 accurately captures chance-level behavior in non-canonical sentences, which makes it difficult to rule out TDH-1 as a possible explanation for offline data on sentence comprehension deficits in aphasia. But the predictions of TDH-1 for online behavior diverged drastically from the observed behavior. Thus, although the offline behavior is consistent with the TDH, the online behavior is not. This is an independent validation of the observation by, among others, Burchert et al. (2013) that "... the consideration of behavioral offline data alone may not be sufficient to evaluate a performance in language tests and draw theoretical conclusions about language impairments. Rather it is important to call on additional data from online studies looking at language processing in real time to understand a performance at the behavioral level and the nature of eventual underlying deficits." In sum, when the predictions for both, online and offline (grouped and individual), measures are considered, model M3 emerges as the best model of aphasic sentence processing among the five models considered.

Although the models—the model for controls and the best model for aphasics, M3—capture the major patterns in the data, some effects were not consistent with the predictions. First, toward the end of the sentence, the models outperform humans in the proportion of fixations to the correct picture. This divergence is only an artifact of the relationship we assume between activation difference and fixation probabilities. For simplicity, we assumed a linear relation between activation difference and fixation probabilities. A further issue with modeling fixation probabilities using activation differences could be that the probabilities are bounded, but activations in ACT-R are not. Using a non-linear linking function with asymptotic growth (e.g., a sigmoid function) can also put an upper bound on the predicted fixation probabilities.

We also tried a non-linear function to define the mapping between activation difference and fixation probabilities. We fit the Boltzmann Sigmoid A function, which has four parameters:

$$y = \frac{(a - b)}{(1.0 + \exp((x - c)/d))} + b \quad (7)$$

The first two parameters, a and b , were kept at 100 and 0, respectively, as they mark the upper and lower bound of the values we want to map activation difference to. The parameters c and d were estimated using the `optim` function in R, and the optimization interface available from `zunzun.com`. The parameter estimates were comparable for `optim` and `zunzun` ($c = -0.139$, $d = -2.327$). Once more data become available, this function could be a good candidate for modeling the mapping from activation differences to fixation probabilities. This non-linear function gave us a better fit and also provided an upper bound on the predicted fixation probabilities. However, such a mapping would artificially overfit the sparse data we have here (eight data points). We, therefore, chose to retain the linear linking function.

Finally, although model M3 does not do a particularly good job of fitting patients' eye movements in non-canonical sentences that resulted in a correct response, it does outperform TDH-1. What is important here is relative fit. If the data reflect an early ability in patients to build the correct parse, this aspect of non-canonical word order processing is not reflected in any of the models. We note in passing here that Hanne and colleagues have conducted follow-up studies with canonical and non-canonical structures using a slightly different design with picture preview, and did not consistently find such an early increase in looks to the correct picture.

4. General discussion

We carried out a computational investigation of several instantiations of two sets of competing theories concerning sentence comprehension deficits in aphasia: the TDH, and theories assuming intermittent deficiencies and/or slowed processing. The two sets of theories have been traditionally categorized as representational deficit accounts and processing deficit accounts. Intermittent deficiency and slowed processing generally attribute difficulty to events in parsing, whereas the TDH assumes that the underlying cause of these disorders is an impairment in representing traces of syntactic movement.

We employed the cue-based retrieval architecture (Lewis & Vasishth, 2005) for modeling the sentence-picture matching study reported in Hanne et al. (2011). We implemented five models of sentence processing deficits in aphasia, comparing two instantiations of the TDH with three alternative accounts—slowed processing, intermittent deficiency, and both these impairments together. Slowed processing was operationalized as slowed procedural memory, and intermittent deficiency as extra noise in the utility values of the parsing rules (productions in ACT-R), and the TDH as an absence of the trace information in the syntactic tree built by the models. Modeling results revealed that the model assuming slowed processing *and* intermittent deficiency fits the aphasics' offline data better than the other models considered. A model of individual aphasics suggests that it may be reasonable to assume that the two impairments are present to differing degrees in

individuals. The TDH models implemented in ACT-R failed to capture crucial patterns in the offline data, such as reduced accuracy in sentence-picture matching for canonical sentences, chance-level performance for non-canonical sentences, and elevated response times in aphasics. Models assuming only one of the two deficits (slowed processing *or* intermittent deficiency) also failed to capture either the chance-level performance or slowed response times.

For modeling online data, that is, the incremental eye movement patterns, we proposed a linking hypothesis. The hypothesis states that a model's predictions about fixating a picture are derived from the activation values of the two pictures presented on the screen. The difference in the activation of the correct and incorrect picture is assumed to predict the probability of eye fixations on the pictures. Importantly, the predictions of eye movements did not involve any parameter fitting; only the offline data were used for parameter estimation. In the data, aphasics' eye movements show that correct offline responses are associated with normal-like online processing and incorrect offline responses are associated with aberrant online processing (aberrant relative to controls' trials and patients' correct trials). Only the model that assumed both intermittent deficiency and slowed processing, and the model assuming only intermittent deficiency, could capture this divergent pattern in the eye movement data. Note that even though the model assuming only intermittent deficiency captured the eye movement data, it could not capture the slow response times in aphasics. For the models based on the TDH, one of the implemented versions of the TDH accurately predicted chance-level behavior in offline accuracies for non-canonical sentences. However, this model failed to predict the observed online behavior in eye movements, especially in correctly answered non-canonical sentences.

Thus, when results for offline and online data are considered together, the model assuming slowed processing as well as intermittent deficiency emerges as the best model (among the models considered) of the aphasics' data from Hanne et al. (2011). Consequently, the results are consistent with the class of hypotheses that ascribe aphasics' sentence processing deficits to intermittent deficiency and slowed processing.

Although the models that implemented the TDH failed to predict crucial patterns in the offline as well as online data, one important caveat here is that we have only investigated two instantiations of the TDH, and we have had to make assumptions about how the TDH predictions would play out in a very specific parsing architecture. It is entirely possible that there exists a particular set of assumptions compatible with the TDH which, when computationally implemented, would yield predictions that fit the data better. In other words, we do not (and cannot) claim that accounts such as the TDH are wrong; we have only shown that, in the specific context of the present simulations, holding all other assumptions about the parsing environment constant, the processing deficit accounts fare better than the versions of TDH we implemented to the best of our knowledge.

How does our model compare with existing computational models? Existing computational models of aphasia (Crescentini & Stocco, 2005; Haarmann & Kolk, 1991; Haarmann et al., 1997) have also evaluated processing or resource reduction-based explanations of aphasia. However, these models differ from our architecture in terms of the precise processing deficit assumed: Crescentini and Stocco (2005) assume slow lexical

activation, Haarmann and Kolk (1991) assume temporal disruption, and Haarmann et al. (1997) assume reduction in memory resources for aphasics. In contrast to these proposals, the present model posits intermittent deficiency and slowed processing as an explanation of sentence comprehension deficits in aphasic patients. An obvious direction for future research is to evaluate the relative merits of these different assumptions, or combinations of these assumptions, within a single architecture. We intend to take this up in future work.

A major contribution of the present work is that we not only propose an implemented model of aphasic sentence processing deficits but also evaluate the predictions of several competing accounts. Further, the model presented goes beyond existing work in delivering predictions for response times as well as incremental eye movements. Finally, the modeling architecture we use here is unique in that a well-developed model of unimpaired sentence comprehension is “damaged” in order to evaluate impairment, and both unimpaired and impaired processing are investigated within a fixed architecture. To our knowledge, this is the first attempt to computationally evaluate the predictions of slowed processing, intermittent deficiency, and the trace deletion account, and to model offline as well as online measures (fixation probabilities in the visual world paradigm) of aphasic sentence processing using the same model for impaired and unimpaired processing.

The linking hypothesis proposed here for modeling sentence-picture matching data opens new possibilities for modeling data from the visual world paradigm, which is an important methodology in psycholinguistics. Here, we modeled mainly the task of sentence-picture matching, but the hypothesis can be extended for modeling other tasks in the visual world paradigm such as the task described in Altmann and Kamide (1999), where participants listen to sentences while they are shown a scene containing mentioned and distractor objects; the task described in Choy and Thompson (2010) where participants are shown an array of mentioned and distractor objects instead of a scene; or the task described in Huettig and McQueen (2007) where participants are shown objects as textual words on the screen. For modeling these tasks, the linking hypothesis can be generalized as “The chunk in declarative memory that receives higher activation as a consequence of processing the current sentence fragment is the object from the visual stimuli that is more likely to be fixated by the participants.” This claim is consistent with the proposal about the linkage between language processing and visual attention presented in Altmann and Kamide (2007).

We must, of course, acknowledge that our modeling attempt also has several limitations. Specifically, we limited our scope to testing the assumptions of the slowed processing and intermittent deficiency accounts on aphasic sentence comprehension for non-canonical sentences in German. There are several other proposals in the literature that assume a processing deficit in aphasia; examples are the timing disorder proposed in Haarmann and Kolk (1991), delayed lexical activation in Love et al. (2008), and lexical integration failure in Thompson and Choy (2009). Additionally, experimental work in aphasia examines data for a wide range of linguistic structures like passives, clefts, subject/object relative clauses, etc. We recognize that a more detailed investigation of the model is needed; a complete investigation of sentence processing in aphasia should

compare predictions of all existing hypotheses across a range of linguistic structures and languages, and perhaps also competing architectures. Such an extensive study is needed in order to determine how general the conclusions of the present work are. Nevertheless, one achievement of the present work is that it demonstrates that the cue-based retrieval architecture—an independently motivated theory of sentence comprehension—provides a flexible environment for testing the different assumptions about deficits in aphasia, in comparison to controls.

There are several interesting implications from our work for modeling sentence processing deficits in aphasia. For example, in recent work Caplan, Michaud, and Hufford (2013) have investigated 61 persons with aphasia on different syntactic structures and different tasks (sentence-picture matching, sentence-picture matching with auditory moving window presentation, and object manipulation). Based on this large dataset, they concluded that the underlying deficit may be a reduction in processing resources; they also found an interaction between task demands and parsing difficulty (a topic that is of great interest in unimpaired sentence processing; see Logačev & Vasishth, 2015; P. Logačev and S. Vasishth (2014), unpublished data). Reductions in resources can be operationalized in several different ways within ACT-R; some examples are intermittent deficits, increased utility noise, or less efficient retrieval processes (Engelmann & Vasishth, 2014). Since one of the features of ACT-R is that different tasks, such as those involving fairly elaborate hand movements, can be directly modeled in ACT-R, our particular modeling approach opens up new possibilities for understanding what it means for a patient to have reductions in resources, and how their performance interacts with tasks. The recent integration of eye-movement control and the cue-based retrieval model described here (Engelmann et al., 2013) also creates new opportunities to model the eye-parser connection directly when modeling visual world data.

A further contribution of the present work is that patients' behavior can be studied in the same architecture that is used to model unimpaired processing; an obvious implication is that models of unimpaired processing may also show occasional impairments that are due to occasional retrieval failures, under conditions of processing overload. This implication should be investigated in future work both experimentally and through computational modeling.

5. Conclusions

There are four major achievements of the present work. First, aphasic patients' impaired sentence processing, reflected in their offline and online behavior during the sentence-picture matching task, is best captured by a model assuming two processing deficits—intermittent deficiency and slowed processing. Second, the model presented here suggests that individual patients may be affected by the two deficits in differing degrees; this necessitates moving beyond evaluating average behavior among patients (Caramazza & McCloskey, 1988). Third, we show that the two models of aphasia based on the TDH fail to capture the effects of reduced accuracy, delays in offline responses, and normal

online performance in correctly answered trials. Finally, we show that a well-developed model of unimpaired sentence comprehension, the cue-based retrieval model grounded in ACT-R, can be extended to model both unimpaired and impaired sentence processing in both offline and online modalities.

Acknowledgments

We thank Florian Jäger, three other anonymous reviewers, and the action editor, Delphine Dahan, for their detailed comments; these greatly helped to improve the paper. We thank David Caplan for discussions of the work reported here, Richard Lewis for helpful suggestions, and Ruben van de Vijver for comments on the earlier version of the manuscript. Earlier versions of this work were presented at various meetings such as the Academy of Aphasia 2010, 2011, 2013; CUNY 2011, 2013; and the Science of Aphasia conference 2011; comments from the audiences at these conferences are gratefully acknowledged.

References

- Aho, A. V., & Ullman, J. D. (1972). *The theory of parsing, translation and compiling volume I: Parsing*. Upper Saddle River, NJ: Prentice-Hall.
- Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language*, 38(4), 419–439.
- Altmann, G. T., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition*, 73(3), 247–264.
- Altmann, G. T., & Kamide, Y. (2007). The real-time mediation of visual attention by language and world knowledge: Linking anticipatory (and other) eye movements to linguistic processing. *Journal of Memory and Language*, 57(4), 502–518.
- Anderson, J. R., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, 111(4), 1036–1050.
- Avrutin, S. (2006). Weak syntax. In Y. Grodzinsky & K. Amunts (Eds.), *Broca's region* (pp. 49–62). New York: Oxford University Press.
- Badecker, W., & Caramazza, A. (1986). A final brief in the case against agrammatism: The role of theory in the selection of data. *Cognition*, 24(3), 277–282.
- Badecker, W., & Straub, K. (2002). The processing role of structural constraints on the interpretation of pronouns and anaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(4), 748–769.
- Beretta, A., & Munn, A. (1998). Double-agents and trace-deletion in agrammatism. *Brain and Language*, 65(3), 404–421.
- Blumstein, S. E., Byma, G., Kurowski, K., Hourihan, J., Brown, T., & Hutchinson, A. (1998). On-line processing of filler-gap construction in aphasia. *Brain Language*, 61(2), 149–168.
- Boston, M. F., Hale, J. T., Vasishth, S., & Kliegl, R. (2011). Parallel processing and sentence comprehension difficulty. *Language and Cognitive Processes*, 26(3), 301–349.
- Burchert, F., & De Bleser, R. (2004). Passives in agrammatic sentence comprehension: A German study. *Aphasiology*, 18(1), 29–45.

- Burchert, F., De Bleser, R., & Sonntag, K. (2003). Does morphology make the difference? Agrammatic sentence comprehension in German. *Brain and Language*, 87(2), 323–342.
- Burchert, F., Hanne, S., & Vasisht, S. (2013). Sentence comprehension disorders in aphasia: The concept of chance performance revisited. *Aphasiology*, 27, 112–125.
- Burkhardt, P., Mercedes Piñango, M., & Wong, K. (2003). The role of the anterior left hemisphere in real-time sentence comprehension: Evidence from split intransitivity. *Brain and Language*, 86(1), 9–22.
- Caplan, D. (2001). The measurement of chance performance in aphasia, with specific reference to the comprehension of semantically reversible passive sentences: A note on issues raised by Caramazza, Capitani, Rey, and Berndt (2001) and Draai, Grodzinsky, and Zurif (2001). *Brain and Language*, 76(2), 193–201.
- Caplan, D. (2009). The neural basis of syntactic processing. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (4th ed., pp. 805–816). Cambridge, MA: MIT Press.
- Caplan, D., & Hildebrandt, N. (1988). *Disorders of syntactic comprehension*. Cambridge, MA: MIT Press.
- Caplan, D., & Waters, G. S. (1995). Aphasic disorders of syntactic comprehension and working memory capacity. *Cognitive Neuropsychology*, 12(6), 637–649.
- Caplan, D., & Waters, G. (2003). On-line syntactic processing in aphasia: Studies with auditory moving window presentation. *Brain and Language*, 84(2), 222–249.
- Caplan, D., Baker, C., & Dehaut, F. (1985). Syntactic determinants of sentence comprehension in aphasia. *Cognition*, 21(2), 117–175.
- Caplan, D., Waters, G., & Hildebrandt, N. (1997). Determinants of sentence comprehension in aphasic patients in sentence-picture matching tasks. *Journal of Speech, Language, and Hearing Research*, 40, 542–555.
- Caplan, D., Waters, G., Dede, G., Michaud, J., & Reddy, A. (2007). A study of syntactic processing in aphasia I: Behavioral (psycholinguistic) aspects. *Brain and Language*, 101(2), 103–150.
- Caplan, D., Michaud, J., & Hufford, R. (2013). Dissociations and associations of performance in syntactic comprehension in aphasia and their implications for the nature of aphasic deficits. *Brain and Language*, 127(1), 21–33.
- Caramazza, A., & McCloskey, M. (1988). The case for single-patient studies. *Cognitive Neuropsychology*, 5(5), 517–527.
- Caramazza, A., & Zurif, E. B. (1976). Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and Language*, 3(4), 572–582.
- Caramazza, A., Capasso, R., Capitani, E., & Miceli, G. (2005). Patterns of comprehension performance in agrammatic Broca's aphasia: A test of the Trace Deletion Hypothesis. *Brain and Language*, 94(1), 43–53.
- Cho, S., & Thompson, C. K. (2010). What goes wrong during passive sentence production in agrammatic aphasia: An eye-tracking study. *Aphasiology*, 24(12), 1576–1592.
- Chomsky, N. (1970). Readings in English transformational grammar. In R. Jacobs, & P. Rosenbaum (Eds.), *Remarks on nominalization* (pp. 184–221). Waltham: Ginn.
- Chomsky, N. (1981). *Lectures on government and binding*. Berlin: Mouton de Gruyter.
- Chomsky, N. (1986a). *Barriers*. Cambridge, MA: MIT Press.
- Chomsky, N. (1986b). *Knowledge of language: Its nature, origins, and use*. New York, NY: Praeger.
- Choy, J. J., & Thompson, C. K. (2010). Binding in agrammatic aphasia: Processing to comprehension. *Aphasiology*, 24(5), 551–579.
- Christiansen, M. H., & Chater, N. (2001). Finite models of infinite language: A connectionist approach to recursion. In M. H. Christiansen, & N. Chater (Eds.), *Connectionist psycholinguistics* (pp. 138–176). Westport, CT: Ablex Publishing.
- Clark, D. G. (2012). Storage costs and heuristics interact to produce patterns of aphasic sentence comprehension performance. *Frontiers in Psychology*, 3, 1–19.
- Coltheart, M. (2006). Acquired dyslexias and the computational modelling of reading. *Cognitive Neuropsychology*, 23(1), 96–109.

- Crescentini, C., & Stocco, A. (2005). Agrammatism as a failure in the lexical activation process. In B. Bara, L. Barsalou, & M. Bucciarelli (Eds.), *Proceedings of the 27th Annual Conference of the Cognitive Science Society*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Cunnings, I., & Felser, C. (2013). The role of working memory in the processing of reflexives. *Language and Cognitive Processes*, 28(1–2), 188–219.
- Dahan, D., & Tanenhaus, M. (2005). Looking at the rope when looking for the snake: Conceptually mediated eye movements during spoken-word recognition. *Psychonomic Bulletin & Review*, 12, 453–459.
- Dahan, D., Magnuson, J. S., Tanenhaus, M. K., & Hogan, E. M. (2001). Subcategorical mismatches and the time course of lexical access: Evidence for lexical competition. *Language and Cognitive Processes*, 16(5–6), 507–534.
- De Bleser, R., Schwarz, W., & Burchert, F. (2006). Quantitative neurosyntactic analyses: The final word? *Brain and Language*, 96(2), 143–146.
- Dell, G. S., & Caramazza, A. (2008). Introduction to special issue on computational modelling in cognitive neuropsychology. *Cognitive Neuropsychology*, 25(2), 131–135.
- Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, 104(4), 801–838.
- Dickey, M. W., & Thompson, C. K. (2006). Automatic processing of wh- and NP-movement in agrammatic aphasia: Evidence from eyetracking. *Brain and Language*, 99(1–2), 63–64.
- Dickey, M. W., & Thompson, C. K. (2009). Automatic processing of wh- and NP-movement in agrammatic aphasia: Evidence from eyetracking. *Journal of Neurolinguistics*, 22(6), 563–583.
- Dickey, M. W., Choy, J. J., & Thompson, C. K. (2007). Real-time comprehension of wh- movement in aphasia: Evidence from eyetracking while listening. *Brain and Language*, 100(1), 1–22.
- Diesing, M. (1992). *Indefinites*. (Linguistic Inquiry Monograph 20.) Cambridge, MA, MIT Press.
- Dilkina, K., McClelland, J. L., & Plaut, D. C. (2008). A single-system account of semantic and lexical deficits in five semantic dementia patients. *Cognitive Neuropsychology*, 25(2), 136–164.
- Dillon, B. (2011). Structured access in sentence comprehension. Unpublished doctoral dissertation, Maryland.
- Dillon, B., Mishler, A., Sloggett, S., & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, 69(2), 85–103.
- Dronkers, N. F., Wilkins, D. P., Van Valin, R. D., Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition*, 92(1–2), 145–177.
- Dubey, A., Keller, F., & Sturt, P. (2008). A probabilistic corpus-based model of syntactic parallelism. *Cognition*, 109(3), 326–344.
- Edwards, S., & Varlokosta, S. (2007). Pronominal and anaphoric reference in agrammatism. *Journal of Neurolinguistics*, 20(6), 423–444.
- Engelmann, F., & Vasishth, S. (2014). Predicting individual differences in underspecification: An integrated model of good-enough processing. *Poster presented at CUNY Conference on Human Sentence Processing*.
- Engelmann, F., Vasishth, S., Engbert, R., & Kliegl, R. (2013). A framework for modeling the interaction of syntactic processing and eye movement control. *Topics in Cognitive Science*, 5(3), 452–474.
- Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, 47(2), 164–203.
- Ferrill, M., Love, T., Walenski, M., & Shapiro, L. P. (2012). The time-course of lexical activation during sentence comprehension in people with aphasia. *American Journal of Speech-Language Pathology*, 21(2), 179–189.
- Frank, S. L., Trompenaars, T., & Vasishth, S. (2015). *Cross-linguistic differences in processing double-embedded relative clauses: Working-memory constraints or language statistics?* *Cognitive Science*.
- Gavarró, A., & Romeu, J. (2010). Agentive versus non-agentive passives in catalan agrammatism. *Procedia Social and Behavioral Sciences*, 6, 200–201.
- Gibson, E. (2000). Dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W. O’Neil (Eds.), *Image, language, brain: Papers from the first mind articulation project symposium* (95–126). Cambridge, MA: MIT Press.

- Gigley, H. M. (1986). Studies in artificial aphasia: Experiments in processing change. *Computer Methods and Programs in Biomedicine*, 22(1), 43–50.
- Grodzinsky, Y. (1995). A restrictive theory of agrammatic comprehension. *Brain and Language*, 50(1), 27–51.
- Grodzinsky, Y. (2000). The neurology of syntax: Language use without Broca's area. *Behavioral and Brain Sciences*, 23(1), 1–71.
- Grodzinsky, Y. (2006). The language faculty, Broca's region, and the mirror system. *Cortex*, 42(4), 464–468.
- Grodzinsky, Y., & Finkel, L. (1998). The neurology of empty categories: Aphasics' failure to detect ungrammaticality. *Journal of Cognitive Neuroscience*, 10(2), 281–292.
- Grodzinsky, Y., & Santi, A. (2008). The battle for Broca's region. *Trends in Cognitive Sciences*, 12(12), 474–480.
- Grodzinsky, Y., Wexler, K., Chien, Y., Marakovitz, S., & Solomon, J. (1993). The breakdown of binding relations. *Brain and Language*, 45(3), 396–422.
- Grodzinsky, Y., Piñango, M. M., Zurif, E., & Draai, D. (1999). The critical role of group studies in neuropsychology: Comprehension regularities in Broca's aphasia. *Brain and Language*, 67(2), 134–147.
- Haarmann, H. J., & Kolk, H. H. (1991). A computer model of the temporal course of agrammatic sentence understanding: The effects of variation in severity and sentence complexity. *Cognitive Science*, 15(1), 49–87.
- Haarmann, H. J., Just, M. A., & Carpenter, P. A. (1997). Aphasic sentence comprehension as a resource deficit: A computational approach. *Brain and Language*, 59(1), 76–120.
- Hagoort, P., Brown, C. M., & Swaab, T. Y. (1996). Lexical-semantic event-related potential effects in patients with left hemisphere lesions and aphasia, and patients with right hemisphere lesions without aphasia. *Brain*, 119(2), 627–649.
- Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of the second meeting of the North American Chapter of the Association for Computational Linguistics* (pp. 159–166). Pittsburgh, PA: Association for Computational Linguistics.
- Hanne, S., Sekerina, I. A., Vasishth, S., Burchert, F., & De Bleser, R. (2011). Chance in agrammatic sentence comprehension: What does it really mean? Evidence from eye movements of German agrammatic aphasic patients. *Aphasiology*, 25(2), 221–244.
- Hanne, S., Burchert, F., De Bleser, R., & Vasishth, S. (2014). Sentence comprehension and morphological cues in aphasia: What eye-tracking reveals about integration and prediction. *Journal of Neurolinguistics*, 34, 83–111.
- Hickok, G., & Avrutin, S. (1995). Representation, referentiality, and processing in agrammatic comprehension: Two case studies. *Brain and Language*, 50(1), 10–26.
- Huetting, F., & Altmann, G. T. M. (2005). Word meaning and the control of eye fixation: Semantic competitor effects and the visual world paradigm. *Cognition*, 96(1), B23–B32.
- Huetting, F., & McQueen, J. M. (2007). The tug of war between phonological, semantic and shape information in language-mediated visual search. *Journal of Memory and Language*, 57(4), 460–482.
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4), 329–354.
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99(1), 122–149.
- Kempen, G., & Vosse, T. (1989). Incremental syntactic tree formation in human sentence processing: A cognitive architecture based on activation decay and simulated annealing. *Connection Science*, 1(3), 273–290.
- Kiran, S., Grasemann, U., Sandberg, C., & Miikkulainen, R. (2012). A computational account of bilingual aphasia rehabilitation. *Bilingualism: Language and Cognition*, 16(02), 325–342.
- Kolk, H. H., & Van Grunsven, M. M. (1985). Agrammatism as a variable phenomenon. *Cognitive Neuropsychology*, 2(4), 347–384.
- Koopman, H., & Sportiche, D. (1991). The position of subjects. *Lingua*, 85(2), 211–258.

- Lee, J., & Thompson, C. K. (2011a). Real-time production of arguments and adjuncts in normal and agrammatic speakers. *Language and Cognitive Processes*, 26(8), 985–1021.
- Lee, J., & Thompson, C. K. (2011b). Real-time production of unergative and unaccusative sentences in normal and agrammatic speakers: An eyetracking study. *Aphasiology*, 25(6–7), 813–825.
- Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126–1177.
- Lewis, R. L., & Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29(3), 375–419.
- Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, 10(10), 447–454.
- Logačev, P., & Vasishth, S. (2015). *A multiple-channel model of task-dependent ambiguity resolution in sentence comprehension*. *Cognitive Science*.
- Logačev, P., & Vasishth, S. (2014). *What is underspecification?* (Manuscript submitted for publication)
- Love, T., Swinney, D., & Zurif, E. B. (2001). Aphasia and the time-course of processing long-distance dependencies. *Brain and Language*, 79, 169–170.
- Love, T., Swinney, D., Walenski, M., & Zurif, E. (2008). How left inferior frontal cortex participates in syntactic processing: Evidence from aphasia. *Brain and Language*, 107(3), 203–219.
- Martin, R. C. (2006). The neuropsychology of sentence processing: Where do we stand? *Cognitive Neuropsychology*, 23(1), 74–95.
- Matzke, M., Mai, H., Nager, W., Rüsseler, J., & Münte, T. (2002). The costs of freedom: An ERP-study of non-canonical sentences. *Clinical Neurophysiology*, 113(6), 844–852.
- Maunder, G., Fromkin, V. A., & Cornell, T. L. (1993). Comprehension and acceptability judgments in agrammatism: Disruptions in the syntax of referential dependency. *Brain and Language*, 45(3), 340–370.
- McClelland, J. L., & Elman, J. L. (1986). The trace model of speech perception. *Cognitive Psychology*, 18(1), 1–86.
- Meinunger, A. (2000). *Syntactic aspects of topic and comment* (Vol. 38). Philadelphia, PA: John Benjamins Publishing.
- Meyer, A. M., Mack, J. E., & Thompson, C. K. (2012). Tracking passive sentence comprehension in agrammatic aphasia. *Journal of Neurolinguistics*, 25(1), 31–43.
- Mirman, D., Yee, E., Blumstein, S. E., & Magnuson, J. S. (2011). Theories of spoken word recognition deficits in aphasia: Evidence from eye-tracking and computational modeling. *Brain and Language*, 117(2), 53–68.
- Mitchum, C. C., & Berndt, R. (2008). Comprehension and production of sentences. In R. Chapey (Ed.), *Language intervention strategies in aphasia and related neurogenic communication disorders* (5th ed., pp. 632–653). Philadelphia: Wolters Kluwer.
- Miyake, A., Carpenter, P. A., & Just, M. A. (1994). A capacity approach to syntactic comprehension disorders: Making normal adults perform like aphasic patients. *Cognitive Neuropsychology*, 11(6), 671–717.
- Myung, J.-Y., Blumstein, S. E., & Sedivy, J. C. (2006). Playing on the typewriter, typing on the piano: Manipulation knowledge of objects. *Cognition*, 98(3), 223–243.
- Nickels, L., Biedermann, B., Coltheart, M., Saunders, S., & Tree, J. J. (2008). Computational modelling of phonological dyslexia: How does the drc model fare? *Cognitive Neuropsychology*, 25(2), 165–193.
- Nievergelt, J. (2000). Exhaustive search, combinatorial optimization and enumeration: Exploring the potential of raw computing power. In V. Hlaváč, K. G. Jeffery, & J. Wiedermann (Eds.), *Sofsem 2000 - Theory and Practice of Informatics* (pp. 18–35). London, UK: Springer-Verlag.
- Patil, U., Vasishth, S., & Kliegl, R. (2009). Compound effect of probabilistic disambiguation and memory retrievals on sentence processing: Evidence from an eye-tracking corpus. In A. Howes, D. Peebles, & R. Cooper (Eds.), *9th international conference on cognitive modeling – ICCM2009* (100–283). Manchester, UK: University of Manchester.
- Patil, U., Vasishth, S., & Lewis, R. L. (2011). Early retrieval interference in syntax-guided antecedent search. Talk presented at CUNY Conference on Human Sentence Processing.

- Patil, U., Vasishth, S., & Lewis, R. L. (2015). *Retrieval interference in syntactic processing: The case of reflexive binding in English*. (Manuscript in preparation)
- Penke, M., & Westermann, G. (2006). Broca's area and inflectional morphology: Evidence from Broca's aphasia and computer modeling. *Cortex*, 42(4), 563–576.
- Plaut, D. C. (2002). Graded modality-specific specialisation in semantics: A computational account of optic aphasia. *Cognitive Neuropsychology*, 19(7), 603–639.
- Plaut, D. C., McClelland, J. L., Seidenberg, M. S., & Patterson, K. (1996). Understanding normal and impaired word reading: Computational principles in quasi-regular domains. *Psychological Review*, 103(1), 56–115.
- Prather, P. A., Zurif, E., Love, T., & Brownell, H. (1997). Speed of lexical activation in nonfluent Broca's aphasia and fluent Wernicke's aphasia. *Brain and Language*, 59(3), 391–411.
- R Development Core Team. (2009). R: A Language and Environment for Statistical Computing [Computer software manual]. Vienna, Austria. (ISBN 3-900051-07-0)
- Reitter, D., Keller, F., & Moore, J. D. (2011). A computational cognitive model of syntactic priming. *Cognitive Science*, 35(4), 587–637.
- Rogers, T. T., Ralph, M. A. L., Garrard, P., Bozeat, S., McClelland, J. L., Hodges, J. R., & Patterson, K. (2004). Structure and deterioration of semantic memory: A neuropsychological and computational investigation. *Psychological Review*, 111(1), 205–235.
- Schwartz, M. F., Saffran, E. M., & Marin, O. S. M. (1980). The word order problem in agrammatism: I. comprehension. *Brain and Language*, 10(2), 249–262.
- Schwartz, M. F., Linebarger, M. C., Saffran, E. M., & Pate, D. S. (1987). Syntactic transparency and sentence interpretation in aphasia. *Language and Cognitive Processes*, 2(2), 85–113.
- Swaab, T. Y., Brown, C. M., & Hagoort, P. (1997). Spoken sentence comprehension in aphasia: Event-related potential evidence for a lexical integration deficit. *Journal of Cognitive Neuroscience*, 9, 39–66.
- Swinney, D., Zurif, E. B., Prather, P., & Love, T. (1996). Neurological distribution of processing operations underlying language comprehension. *Journal of Cognitive Neuroscience*, 8, 174–184.
- Tabor, W., Galantucci, B., & Richardson, D. (2004). Effects of merely local syntactic coherence on sentence processing. *Journal of Memory and Language*, 50, 355–370.
- Thompson, C. K., & Choy, J. (2009). Pronominal resolution and gap filling in agrammatic aphasia: Evidence from eye movements. *Journal of Psycholinguistic Research*, 38(3), 255–283.
- Thompson, C. K., Dickey, M. W., & Choy, J. J. (2004). Complexity in the comprehension of wh-movement structures in agrammatic Broca's aphasia: Evidence from eyetracking. *Brain and Language*, 91(1), 124–125.
- Thompson, C. K., Choy, J. J., Holland, A., & Cole, R. (2010). Sentactics®: Computer-automated treatment of underlying forms. *Aphasiology*, 24(10), 1242–1266.
- Traxler, M. J., Pickering, M. J., Clifton, J. C., & Clifton, C. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language*, 39(4), 558–592.
- Van Dyke, J. A., & Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language*, 49(3), 285–316.
- Vasishth, S., & Lewis, R. L. (2006a). Argument-head distance and processing complexity: Explaining both locality and antilocality effects. *Language*, 82(4), 767–794.
- Vasishth, S., & Lewis, R. L. (2006b). Human language processing: Symbolic models. In K. Brown (Ed.), *Encyclopedia of language and linguistics* (Vol. 5, pp. 410–419). Oxford: Elsevier.
- Vasishth, S., Brüssow, S., Lewis, R. L., & Drenhaus, H. (2008). Processing polarity: How the ungrammatical intrudes on the grammatical. *Cognitive Science*, 32(4), 685–712.
- Vasishth, S., Suckow, K., Lewis, R. L., & Kern, S. (2010). Short-term forgetting in sentence comprehension: Crosslinguistic evidence from head-final structures. *Language and Cognitive Processes*, 25(4), 533–567.
- Vosse, T., & Kempen, G. (2000). Syntactic structure assembly in human parsing: A computational model based on competitive inhibition and a lexicalist grammar. *Cognition*, 75, 105–143.

- Wagers, M. W., Lau, E. F., & Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, 61(2), 206–237.
- Woollams, A. M., Ralph, M. A. L., Plaut, D. C., & Patterson, K. (2007). Sd-squared: On the association between semantic dementia and surface dyslexia. *Psychological Review*, 114(2), 316–339.
- Yee, E., Blumstein, S. E., & Sedivy, J. C. (2008). Lexical-semantic activation in Broca's and Wernicke's aphasia: Evidence from eye movements. *Journal of Cognitive Neuroscience*, 20(4), 592–612.
- Zurif, E. B., Swinney, D., Prather, P., & Love, T. (1994). Functional localization in the brain with respect to syntactic processing. *Journal of psycholinguistic research*, 23, 487–497.

Appendix A: Parsing algorithms for the cue-based retrieval models

The models were implemented as simulations in R (R Development Core Team, 2009).

Algorithm 1: Canonical sentence (Der Sohn fängt den Vater).

1. Input: λ (empty string before the 1st word of the sentence)
 - (a) Create CP, IP, VP (attach the VP to comp-IP and the IP to comp-CP)
 - (b) Predict an SVO structure: create a nominative DP with an empty head and attach it to spec-CP; create trace “t_s” in spec-VP and in spec-IP, co-index “t_s” and the DP
2. Input: Der
 - (a) Retrieve the nominative DP and assign “Der” to head-DP
3. Input: Sohn
 - (a) Retrieve the nominative DP; create an NP with “Sohn” as the head and attach it to comp-DP
4. Create picture chunks of the two pictures on the screen
5. Retrieve a picture chunk that matches the (partial) sentence
6. Input: fängt
 - (a) Attach “fängt” to head-CP; create trace “t_v” in head-VP & head-IP; co-index “t_v” and the CP
 - (b) Predict an accusative DP: create an accusative DP and attach it to comp-VP
 - (c) Retrieve the nominative DP

[Note: (d1) & (d2) below lead to two competing productions with different utility values in the models for controls, M1, M2, and M3; and (d3) implements the hypotheses in the models for TDH-1 and TDH-2.]

 - (d1) Use the trace “t_s” in spec-VP and assign an AGENT theta role to the retrieved DP
 - (d2) Fail to use the *available* trace information in spec-VP and assign a THEME theta role to the retrieved DP
 - (d3) Due to the deletion of trace “t_s” in spec-VP assign an AGENT theta role to the retrieved DP following the subject-first strategy
7. Retrieve a picture chunk that matches the (partial) sentence
8. Input: den

- (a) Retrieve the accusative DP and attach “den” to head-DP
[Note: (b1) & (b2) below lead to two competing productions with different utility values in the models for controls, M1, M2, and M3; (b1) implements the hypotheses in the models for TDH-1 and TDH-2, (b2) is absent in the TDH models.]
 - (b1) Use the information that the DP is in comp-VP and assign a THEME theta role to the retrieved DP
 - (b2) Fail to use the information that the DP is in comp-VP and assign an AGENT theta role to the retrieved DP
9. Input: Vater
- (a) Create an NP with “Vater” as the head
 - (b) Retrieve the accusative DP and attach NP to comp-DP
10. Retrieve a picture chunk that matches the sentence

Algorithm 2: Non-canonical sentence (Den Sohn fängt der Vater).

1. Input: λ (empty string before the 1st word of the sentence)
 - (a) Create CP, IP, VP (attach the VP to comp-IP and the IP to comp-CP)
 - (b) Predict an SVO structure: create a nominative DP with an empty head and attach it to spec-CP; create trace “t_s” in spec-VP and in spec-IP, co-index “t_s” and the DP
2. Input: Den
 - (a) Change the prediction of an SVO structure to OVS structure: Delete the DP in spec-CP; delete trace “t_s” in spec-VP and spec-IP
 - (b) Create a DP with “Den” as the head and attach it to spec-CP; create trace “t_o” in comp-VP; co-index “t_o” and the DP
3. Input: Sohn
 - (a) Create an NP with “Sohn” as the head and attach it to comp-DP
4. Create picture chunks of the two pictures on the screen
5. Retrieve a picture chunk that matches the (partial) sentence
6. Input: fängt
 - (a) Attach “fängt” to head-CP; create trace “t_v” in head-VP & head-IP; co-index “t_v” and the CP
 - (b) Predict a nominative DP: create a nominative DP and attach it to spec-IP; create trace “t_s” in spec-VP and co-index “t_s” with the DP
 - (c) Retrieve the accusative DP
[Note: (d1) & (d2) below lead to two competing productions with different utility values in the models for controls, M1, M2, and M3; and (d3) implements the hypotheses in the models for TDH-1 and TDH-2.]
 - (d1) Use the trace “t_o” in comp-VP and assign a THEME theta role to the retrieved DP
 - (d2) Fail to use the *available* trace information in comp-VP and assign an AGENT theta role to the retrieved DP

- (d3) Due to the deletion of the trace “t_o” in comp-VP and assign an AGENT theta role to the retrieved DP following the subject-first strategy
7. Retrieve a picture chunk that matches the (partial) sentence
 8. Input: der
 - (a) Retrieve the nominative DP and attach “der” to head-DP
[Note: (b1) & (b2) below lead to two competing productions with different utility values in the models for controls, M1, M2, and M3; (b2) & (b3) implement the hypothesis in the model for TDH-1; and (b4) implements the hypothesis in the model for TDH-2.]
 - (b1) Use the trace “t_v” of the VP in head-IP and assign an AGENT theta role to the retrieved DP
 - (b2) Fail to use the *available* trace “t_v” of the VP in head-IP and assign a THEME theta role to the retrieved DP
 - (b3) Use the trace “t_v” of the VP in head-IP and assign an AGENT theta role to the retrieved DP and change the AGENT theta role of the earlier DP to THEME
 - (b4) Due to the deletion of the trace “t_s” in spec-VP assign a THEME theta role to the retrieved DP
 9. Input: Vater
 - (a) Create an NP with “Vater” as the head
 - (b) Retrieve the nominative DP and attach NP to comp-DP
 10. Retrieve a picture chunk that matches the sentence

Appendix B: Details of the models

Table B1

The list of parameter values used in the previous studies with ACT-R models and the values used in the current models for controls and the two models implementing the Trace Deletion Hypothesis (TDH-1 and TDH-2)

Parameter	Previous Models	Current Models (Controls, TDH-1, and TDH-2)
Decay (d)	0.50	0.50
Maximum associative strength (S)	1.50	1.50
Retrieval Threshold (T)	-1.50	-1.50
Maximum difference	-0.60	-1
Latency Factor (F)	0.14, 0.46	0.04
Noise (ϵ)	0, 0.15, 0.30, 0.45	0.30
Default action time	0.05	0.02
Utility noise	0	0.01

Table B2

The utility values estimated for the productions that assign theta roles. Each algorithm step corresponds to a production in ACT-R (see Algorithm 1 and 2 in Appendix A)

Algorithm Steps	Utility Values for	
	Controls, M1, M2, & M3	TDH-1
Algorithm 1 step (6)(d1)	0.6	–
Algorithm 1 step (6)(d2)	0.4	–
Algorithm 1 step (8)(b1)	0.6	–
Algorithm 1 step (8)(b2)	0.4	–
Algorithm 2 step (6)(d1)	0.52	–
Algorithm 2 step (6)(d2)	0.48	–
Algorithm 2 step (8)(b1)	0.52	–
Algorithm 2 step (8)(b2)	0.48	0.5
Algorithm 2 step (8)(b3)	–	0.5

Note. A missing value in a cell indicates that that production does not exist in the model(s). The values are listed only for competing productions (algorithm steps); for all other steps the utility value is the default value of zero. Utility values for TDH-2 are not listed because there are no competing productions in TDH-2.

Table B3

The individual aphasics' data from offline measure in the Hanne et al. (2011) study

Participant	Percentage Correct		Response Time in ms (Correct)		Response Time in ms (Incorrect)	
	SVO	OVS	SVO	OVS	SVO	OVS
P1	75	50	4,148 (261)	4,528 (384)	4,731 (584)	6,279 (904)
P2	80	60	3,629 (279)	3,337 (184)	4,342 (1,001)	3,829 (316)
P3	85	53	3,636 (351)	5,642 (887)	3,560 (261)	3,764 (564)
P4	70	25	5,943 (889)	9,613 (1,617)	5,194 (1,386)	7,655 (1,076)
P5	75	50	3,143 (320)	3,538 (534)	3,642 (605)	3,277 (272)
P6	95	45	3,840 (246)	4,144 (457)	5,160 (–)	6,679 (788)
P7	70	40	7,487 (757)	6,017 (735)	6,718 (899)	7,758 (860)

Note. The numbers in parenthesis indicate standard errors.

Table B4

The estimated values of the parameters *default action time* (DAT) and *utility noise* for each aphasic participant

Participant	P1	P2	P3	P4	P5	P6	P7
Default action time	0.05	0.04	0.04	0.09	0.03	0.06	0.09
Utility noise	0.15	0.09	0.08	0.21	0.15	0.05	0.18

Note. The values of DAT were estimated for response times in M1 while keeping the other parameter values constant. Similarly, the values for utility noise were estimated for response accuracies in M2. In M3 the two estimated values were used together.

Table B5

The normalized root-mean-square deviation between observed and predicted response accuracies and response times (RT) for controls

Participant	Accuracy	RT
C1	0.3	0.6
C2	0.9	1.9
C3	0.2	1.3
C4	0.2	1.2
C5	0.7	0.5
C6	Undefined	0.2
C7	0.3	0.3

Note. Note that model predictions for controls were generated with a single set of parameter values for all participants. (Participant C6 had 100% accuracy for both canonical and non-canonical sentences, therefore, NRMSD is undefined; see Eq. 6 for the definition of NRMSD.)