

Can the resource reduction hypothesis explain sentence processing in aphasia? A visual world study in German

Dorothea Pregla, Shravan Vasishth, Paula Lissón, Nicole Stadie, Frank Burchert

University of Potsdam, Germany

ARTICLE INFO

Keywords:

Aphasia
Sentence comprehension
Visual world eye-tracking
Resource reduction
Relative clauses
Subject- and object-control structures

ABSTRACT

Resource limitation has often been invoked as a key driver of sentence comprehension difficulty, in both theories of language-unimpaired and language-impaired populations. In the field of aphasia, one such influential theory is Caplan's resource reduction hypothesis (RRH). In this large investigation of online processing in aphasia in German, we evaluated three key predictions of the RRH in 21 individuals with aphasia and 22 control participants. Measures of online processing were obtained by combining a sentence-picture matching task with the visual world paradigm. Four sentence types were used to investigate the generality of the findings, and two test phases were used to investigate RRH's predictions regarding variability in aphasia. The processing patterns were consistent with two of the three predictions of the RRH. Overall, our investigation shows that the RRH can account for important aspects of sentence processing in aphasia.

1. Introduction

In sentence processing research, it is well-established that limitations in resource capacity can affect sentence comprehension (Just and Carpenter, 1992). The idea of a limited resource capacity has also been implemented to explain the performance of individuals with aphasia (IWA) in sentence comprehension tasks, e.g., sentence-picture matching (Caplan, 2012; Miyake et al., 1994). The resource reduction approach predicts the following performance pattern for IWA: Resource reduction should impair sentence comprehension across different types of sentence structures (e.g., relative clauses, or sentences with pronouns, Caplan et al., 2015; Caplan and Hildebrandt, 1988). Furthermore, resource reduction should generate a variable impairment in sentence comprehension depending on the amount of available resources. These predictions can be tested experimentally by comparing comprehension performance of the same IWA across different tasks and sentence structures. This approach has been taken by Caplan et al. (2006), Caplan et al. (2015), Caplan et al. (2013), Caplan et al. (2007) for English and more recently by Pregla et al. (2021) for German. The tasks consisted in different versions of sentence-picture matching (Caplan et al., 2006; Caplan et al., 2015; Caplan et al., 2013; Caplan et al., 2007; Pregla et al., 2021), grammaticality judgement (Caplan et al., 2006; Caplan et al., 2007), and object manipulation (Caplan et al., 2006; Caplan et al., 2013; Caplan et al., 2007; Pregla et al., 2021). These studies showed that IWA had a variable degree of difficulty comprehending the same sentence structures in different tasks (Caplan et al., 2006; Caplan et al., 2015; Caplan et al., 2013; Caplan et al., 2007; Pregla et al., 2021). Furthermore, comprehension difficulty was not restricted to a specific sentence

structure but affected complex sentences in general. Both the variability in performance and the general impairment for complex sentences speak for the view that the sentence comprehension impairment seen in IWA is brought about by resource reduction.

This paper will examine the resource reduction approach more closely. More specifically, this paper will investigate one influential instantiation of this approach, the *resource reduction hypothesis* (RRH, e.g., Caplan, 2012; Caplan et al., 2006; Caplan et al., 2007; Caplan et al., 2015). Below, we introduce the RRH, and examine whether previous findings relating to online sentence processing in aphasia are consistent with this account.

1.1. The resource reduction hypothesis

According to the RRH, sentence comprehension depends both on the resource capacity of a participant and the amount of resources a particular sentence comprehension task demands from the available resources of that given participant (Caplan, 2012). Although the exact nature of the resource capacity is not defined in the RRH, Caplan et al. (2013) enumerate a number of different resource types that might cause sentence comprehension impairments in case they do not function as expected. The authors suggest that sentence comprehension impairments might for example be caused by a reduced processing efficiency, by "slowed or otherwise disrupted activation of lexical representations", by "a more general disturbance affecting skilled performances, such as an across-the-board slowing of processing speed", by a reduction in "operations that are needed to perform a task to which comprehension is directed, such as planning actions, inspecting pictures, etc.", or by a

<https://doi.org/10.1016/j.bandl.2022.105204>

Received 29 April 2022; Received in revised form 21 October 2022; Accepted 7 November 2022

Available online 23 November 2022

0093-934X/© 2022 The Authors. Published by Elsevier Inc. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

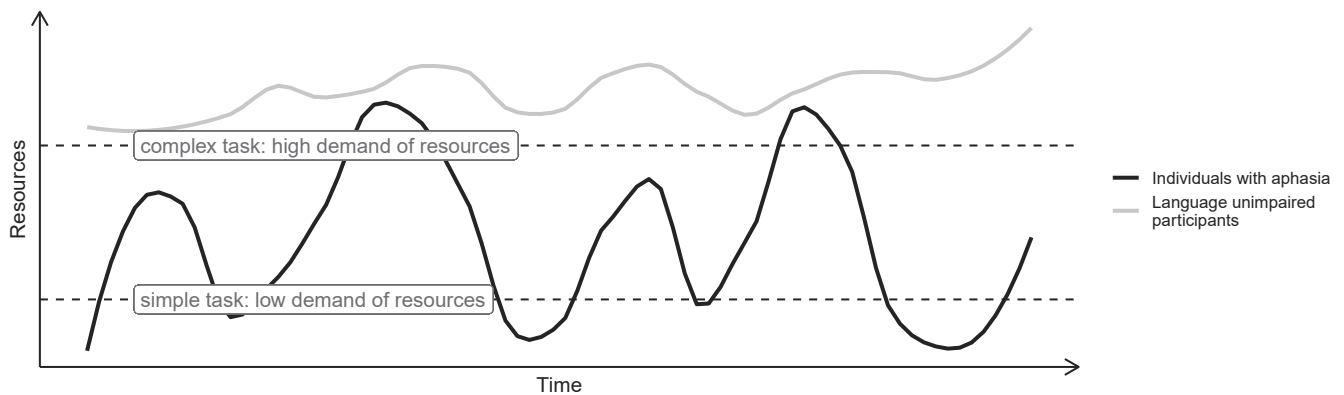


Fig. 1. A schematic illustration of the assumed fluctuation of resources according to the resource reduction hypothesis. Solid lines represent the resource capacity of language-impaired participants (black) and language-unimpaired control participants (grey). Resources randomly fluctuate over arbitrary units of time due to noise in the comprehension system of the participant. Dashed lines represent the resource demand of a simple task (low demand) and of a complex task (high demand). Processing is impaired if the task demand exceeds the resource capacity, otherwise processing is normal.

reduction in "executive functions in the form of deployment of attention, maintenance of task goals, uploading mechanisms that support task performance [...], executing those mechanisms, response selection, assessment of success on a trial, and other processes." (Caplan et al., 2013, p. 28–29). While the RRH is undetermined with respect to what type of resource is affected, the RRH assumes that the capacity of this resource is reduced in IWA in comparison to control participants. Furthermore, the RRH assumes that the resource capacity is subject to random fluctuations caused by noise inherent to a participant. This means that the resources in the processing system can vary from participant to participant and in the same participant from moment to moment. This fluctuation is assumed to be larger in IWA than in control participants. The resource demands depend on the complexity of a task and are stable. Task complexity can be determined by the average performance of IWA or control participants in a sentence comprehension task, i.e., tasks that are difficult for a group of participants are said to be complex (Caplan, 2012). The RRH assumes that tasks with high complexity impose greater resource demands than tasks with lower complexity.

Fig. 1 illustrates the interplay between the resource capacity inherent to participants (solid lines) and task demands (broken lines) according to the RRH. The figure displays the randomly fluctuating resource capacity of IWA (black) and control participants (grey) over an arbitrary period of time. When the resources of a given participant meet the task demands, sentence processing proceeds in a normal-like fashion, resulting in a correct response. However, if the task demands exceed the available resources of a given participant, sentence processing is impaired, resulting in an incorrect response. According to the RRH, processing is more impaired in complex sentences (e.g., object relative clauses) than in simple sentences (e.g., subject relative clauses) because the resource demands of complex sentences are more likely to exceed the participant's resource capacity. However, since noise randomly affects resource capacity, the resources of the participant can sometimes be high enough to process a complex sentence correctly, and sometimes too low to even process a simple sentence correctly. From these assumptions of the RRH, we derived the novel prediction that performance in sentence comprehension tasks should be variable both within sessions and between sessions because of the noise that randomly affects processing in IWA.

The RRH explains the offline comprehension performance of IWA, but it also makes predictions regarding the online processing mechanisms in sentence comprehension in aphasia. So far, the RRH's predictions regarding online performance have only been investigated with the self-paced listening paradigm (Caplan et al., 2007; Caplan et al., 2015). In the present study, the RRH's predictions were investigated using the visual world paradigm (Cooper, 1974). In this paradigm,

participants are simultaneously presented with pictures on a visual display and auditory speech while their proportion of fixations to each picture is recorded. The visual world paradigm is well-established as a means to study sentence processing as it unfolds (Tanenhaus et al., 1995). Next, the predictions of the RRH regarding online processing in aphasia will be presented and it is shown whether previous results are consistent with the predictions. Since the RRH is undetermined with respect to what type of resource is affected in IWA, three options were taken into account that are all discussed in Caplan et al. (2015), namely random fluctuations in resources leading to intermittent deficiencies, slowed processing speed, and syntactic proficiency as expressed by sentence comprehension accuracy.

1.1.1. Prediction 1: Normal-like processing in correct trials

The RRH predicts that correct responses in sentence comprehension should result predominantly from normal syntactic processing (as opposed to accidentally correct responses because of guessing in every trial) while incorrect responses should result from impaired syntactic processing.¹ In line with this prediction, Caplan et al. (2007) found that the self-paced listening times in IWA differed between correct and incorrect trials, and that the listening times were qualitatively similar to control participants in correct trials. Visual world studies also found that the proportion of fixations to the target picture (henceforth target fixations) differed between correct and incorrect trials in IWA (Arantzeta et al., 2017; Choy and Thompson, 2010; Dickey et al., 2007; Dickey and Thompson, 2009; Hanne et al., 2011; Hanne et al., 2012; Hanne et al., 2015; Hanne et al., 2016; Meyer et al., 2012). In correct trials, target fixations of IWA and control participants were qualitatively similar (Arantzeta et al., 2017; Dickey et al., 2007; Dickey and Thompson, 2009; Hanne et al., 2011; Hanne et al., 2015; Hanne et al., 2016; Meyer et al., 2012). Thus, as predicted by the RRH, syntactic processing in IWA tends to proceed normal-like in correct trials.

Besides the normal-like pattern, a number of visual world studies reported delayed target fixations in IWA in comparison to control participants in correct trials (Hanne et al., 2011; Hanne et al., 2016; Meyer et al., 2012; Schumacher et al., 2015). These delays were ascribed to a processing slowdown (Hanne et al., 2011; Hanne et al., 2016; Meyer et al., 2012). The RRH does not make a prediction about processing speed. However, it can account for slowed processing under the assumption that the reduced capacity is processing speed, which Caplan

¹ However, Caplan et al. (2015) point out that severely impaired IWA could be indeed guessing when they answer correctly. Caplan et al. (2015) define severely impaired IWA as those with accuracies below chance level in sentence-picture matching.

et al. (2015) consider likely. Thus, the finding that syntactic processing in IWA in correct trials seems to proceed normal-like but slower than in control participants is compatible with the RRH.

1.1.2. Prediction 2: Processing difficulty in complex vs. simple sentences, and a complexity-capacity interaction

The RRH predicts processing differences between syntactically simple and complex sentences. In line with this prediction, Caplan et al. (2007) and Caplan et al. (2013) found complexity effects in the form of lower accuracy scores and slower response times for syntactically complex versus simple sentences. Additionally, the RRH predicts a super-additive interaction of resource capacity and resource demands, i. e., increased demands should affect participants with a lower capacity level far more than participants with a higher capacity level (e.g., Caplan et al., 2007). Caplan et al. (2007) and Caplan et al. (2015) investigated this prediction in two self-paced listening experiments. As a measure of resource capacity, the authors used the accuracy of each IWA in non-canonical sentences.² As a measure of task complexity, the authors used the listening times in simple and complex sentences. In line with the RRH, Caplan et al. (2007) and Caplan et al. (2015) found a super-additive effect, i.e., the difference in listening times between simple and complex sentences was larger for IWA with lower accuracy.

A number of visual world studies have investigated the influence of sentence complexity on fixations to a target picture (Hanne et al., 2015; Mack et al., 2016; Meyer et al., 2012; Sheppard et al., 2015). Both IWA and control participants showed more target fixations in simple canonical versus complex non-canonical sentences (Hanne et al., 2015; Mack et al., 2016; Meyer et al., 2012). However, participant groups differed in the sentence region where complexity influenced the target fixations. Control participants showed increased target fixations in canonical versus non-canonical sentences before the region that disambiguated the sentence's reading (e.g., *The man was*). The differences vanished directly after disambiguation (e.g., *shaving/ shaved by the boy*). This fixation behavior was interpreted as an agent-first processing pattern, i.e., a tendency to process the first noun of a sentence as the agent followed by a revision in non-canonical sentences (Hanne et al., 2015; Mack et al., 2016; Meyer et al., 2012). In contrast to control participants, IWA showed increased target fixations in canonical versus non-canonical sentences only after the disambiguating region (Hanne et al., 2015; Mack et al., 2016; Mack and Thompson, 2017; Meyer et al., 2012). The fixation pattern in IWA is compatible with the assumption of the RRH that the processing difficulty is larger in complex versus simple sentences but arises more slowly than in control participants.

1.1.3. Prediction 3: Unsystematic variability in the performance between test and retest

The RRH predicts that sentence comprehension varies unsystematically over time within the same IWA. This is because noise should randomly affect the resources available for sentence processing. Little is known about the nature of this variability in online processing. Only one visual world study (Mack et al., 2016) has investigated this issue so far. Mack et al. (2016) tested the processing of active and passive sentences (*The man visited the woman/ was visited by the woman*) in a group of 12 IWA and 21 control participants in two sessions spaced one week apart. The authors investigated the test-retest reliability of the eye-tracking measures and found that the reliability was generally strong in IWA (intra-class correlation coefficients between 0.59 and 0.75), and overall stronger in the IWA than in the control participants. Therefore, Mack et al. (2016) concluded that eye-tracking measures can be reliably used to investigate changes over time in the performance of IWA.

² To determine capacity, the authors used the accuracy across sentence types (including passives, object clefts and object relative clauses Caplan et al., 2007) or the accuracy for each separate sentence type (for passives, object relative clauses, reflexives and pronouns Caplan et al., 2015).

Furthermore, the authors investigated the intra-individual variability in the eye-tracking measures and observed that it did not differ between the language-impaired and language-unimpaired groups. Thus, Mack et al. (2016) tentatively suggest that day-to-day variability in online sentence processing is not larger in IWA than in language-unimpaired individuals. Finally, both participant groups showed increased target fixations in the second compared to the first session independent of the sentence type. Mack et al. (2016) interpreted the increase in target fixations in the retest as a practice effect. The practice effect might indicate that variable processing between sessions is not just random fluctuation but reflects systematic changes. Such changes are currently not accounted for by the RRH.

1.2. Aim of the study

The present study aimed to investigate the RRH's predictions regarding sentence processing in IWA. To this end, the visual world paradigm was used. This paradigm allows us to investigate automatic processing during auditory sentence presentation. The paradigm has the advantage over other online paradigms (e.g., self-paced listening, or cross-modal priming) that it is easy to carry out for IWA (Dickey et al., 2007). Furthermore, the paradigm offers more direct information on syntactic processing than offline analyses, because the data are gathered during sentence presentation and thus can reveal how participants arrive at a sentence interpretation (Dickey et al., 2007). Offline responses also require additional conscious processes that might be impaired in IWA making it difficult to draw conclusions about underlying processing abilities (Caplan et al., 2013). Therefore, the visual world paradigm is suitable to test the predictions of the RRH regarding processing in IWA.

Our experimental design was unique in that sentence processing was investigated across two test phases and four sentence types. This design was chosen to assess the fluctuation in sentence processing in IWA. Furthermore, our study included a relatively large group of 21 IWA. According to a review by Sharma et al. (2021) including 13 visual world studies on sentence comprehension in aphasia, the average number of participants amounts to less than ten IWA (mean = 9 IWA, range = 4 to 16 IWA).³ Furthermore, our study tested sentence comprehension in German while previous studies investigating the RRH focused on English (Caplan et al., 2007; Caplan et al., 2013; Caplan et al., 2015). Given that the RRH is presumably a language-independent theory, it is vital to test its predictions in other languages. There are several reasons why it is interesting to investigate German. In comparison to English, German has a relatively free word order. Furthermore, German allows disambiguating thematic roles based on case marking. Therefore, word order complexity can be varied based on minimal changes in case marking. To our knowledge, our study is the first comprehensive investigation of the RRH for German, and the first to use the visual world paradigm for this purpose.

The following predictions with respect to the target fixations in aphasia were derived from the RRH:

Fixation patterns derived from prediction 1: Normal-like processing in correct trials. In correct trials, the target fixations of the IWA should be similar to those of control participants. That is, both participant groups should show increases in target fixations over the course of a trial (i.e., increases relative to the beginning of a trial where the proportion of target fixations should be 50%). However, target fixations might increase more slowly in IWA than in control participants, as observed in previous visual world experiments (Hanne et al., 2011; Hanne et al., 2016; Meyer et al., 2012; Schumacher et al., 2015). The RRH would be compatible with slow increases in target fixations in IWA

³ Sharma et al. (2021) report a range of 4 to 19 IWA because they did not account for the fact that the study with 19 IWA (Barbieri et al., 2019) had to exclude 3 IWA. This results in the upper bound of 16 IWA reported here.

Table 1
Demographic and neurological data of the individuals with aphasia.

IWA	Gender	Years Age	Years Education	Years P.O.	Etiology	Localization	LEMO ¹ (raw scores)		AAT ²	
							T3 (n = 80)	T11 (n = 20)	Aphasia type	Severity (standard nine)
2	F	72	8	7	IMI	L	77	19	Anomic	6.8 (mild)
3	M	76	20	17	IMI	L/R	61	20	Not-classifiable	7 (mild)
4	F	47	13	21	IMI	L	78	20	Anomic	7.8 (mild)
6	M	55	14	10	IMI	L	67	20	Anomic	6.8 (mild)
8	F	51	19	7	MA	L	74	20	Anomic	7.4 (mild)
9	M	64	15	2	IMI	L	73	20	Anomic	7.4 (mild)
10	M	58	18	1	IMI	L	52	20	Broca	5 (moderate)
11	F	63	12	1	IMI	L	73	20	Broca	6.8 (mild)
12	F	46	12	13	IMI	L	65	20	Broca	4.2 (moderate)
13	M	74	13	8	IMI	L	57	20	Broca	4.4 (moderate)
14	M	66	13	17	IMI	L	75	20	Anomic	6.4 (mild)
15	F	59	21	4	I	L	77	20	Broca	5.2 (moderate)
16	M	67	17	26	VH	R	72	19	Broca	5.4 (moderate)
17	F	43	14	10	IMI	L	65	20	Broca	6.6 (mild)
18	M	57	13	1	I	L	67	18	Wernicke	not available
19	F	52	19	8	IMI	L	76	20	Broca	5.8 (moderate)
20	M	38	13	3	IMI	L	73	19	Broca	4.2 (moderate)
21	M	57	18	2	IMI	L	66	18	Broca	6 (mild)
22	F	67	16	5	IMI	L	76	20	Anomic	6.2 (mild)
23	M	74	15	7	IMI	L	67	20	Anomic	6.6 (mild)
24	M	78	15	6	IMI	L	not available	19	Wernicke	5.6 (moderate)

Note. IWA = individual with aphasia, F = female, M = male, P.O. = post onset, IMI = ischemic arteria cerebri media infarct, I = infarct, MA = arteria cerebri media aneurysm, VH = vertebrobasilar hemorrhage, L = left, R = right, ¹ LEMO 2.0 (Stadie et al., 2013) T3 = auditory lexical decision, T11 = auditory word-picture matching, ² Aachen Aphasia Test (Huber et al., 1983).

because the reduced capacity likely is processing speed (Caplan et al., 2015). Furthermore, increases in target fixations should be higher in correct versus incorrect trials.

Fixation patterns derived from prediction 2: Processing difficulty in complex vs. simple sentences, and a complexity-capacity interaction. Target fixations should diverge between simple and complex sentences, and the increase in target fixations should be higher in simple sentences. Furthermore, the RRH predicts a super-additive interaction between resource demands and resource capacity. Following Caplan et al. (2007) and Caplan et al. (2015), IWA with a lower overall accuracy are assumed to have a lower resource capacity, thus, they should show a more pronounced complexity effect. Consequently, if the overall accuracy of the IWA decreases, the difference in target fixations between simple and complex sentences should increase.

Fixation patterns derived from prediction 3: Unsystematic variability in the performance between test and retest. Following the RRH, fixation paths should vary randomly between the test and retest phase in IWA. That is, target fixations should not systematically increase faster over the course of a trial in the retest than in the test, as would be expected if practice effects were present (Mack et al., 2016).

2. Methods and Material

This visual world experiment investigated the processing of declarative sentences (henceforth declaratives), relative clauses and subject and object control structures (henceforth control structures) with an overt pronoun or a covert pronoun (henceforth PRO) in German in language-unimpaired control participants and IWA. In what follows, the specifics of the methods and materials are explained.

2.1. Participants

Overall, 43 participants, all native speakers of German completed the study: 21 IWA (9 females, mean age = 60 years, $SD = 11$, range = 38–78; mean education = 15 years, $SD = 3$, range = 8–22) and 22 age- and education-matched control participants (14 females, mean age = 58 years, $SD = 15$, range = 26–81; mean education = 16 years, $SD = 4$, range = 6–21). All participants had normal or corrected-to-normal

hearing and vision. Only control participants without known neurological disorders or language impairments were included. Inclusion criteria for IWA were the presence of chronic aphasia (>12 months post onset), no upper limb apraxia, and intact comprehension of nouns. Aphasia had to be apparent according to the Aachen Aphasia Test (Huber et al., 1983).⁴ Participants gave written consent in accordance with the ethics committee of the University of Potsdam and were paid for participation.

Control participants were recruited from the University of Potsdam and from a church parish. All control participants were right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). Control participants were screened for dementia using the Montreal Cognitive Assessment (MoCA, Nasreddine et al., 2005) and all participants were in the normal range, i.e., they scored at least 26/30 points (mean = 29 points, $SD = 1$, range = 26–30). Originally, data from 50 control participants were gathered. For age and education matching, 28 control participants were excluded prior to the analyses. Fig. A.1 in the appendix shows that the fixation paths of the 50 and the 22 control participants are qualitatively similar for all sentence types. Five additional control participants were excluded prior to the analyses because of neurological impairments (3 participants), or because they did not complete all tasks (2 participants).

IWA were recruited from a database of the University of Potsdam and from aphasia self-help groups in Potsdam and Berlin. Demographic and neurological information about the IWA is summarized in Table 1. All but one participant experienced a single stroke at least one year prior to the study. All but three participants were right-handed pre-morbidly as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). The Aachen Aphasia Test (Huber et al., 1983) was administered to determine

⁴ We did not exclude IWA with certain types of aphasia. It has been hypothesized that sentence complexity specifically affects people with Broca's aphasia (e.g., Drai and Grodzinsky, 1999). However, other authors (e.g., Caplan et al., 2015; Luzzatti et al., 2001) did not confirm a generalization of such a comprehension pattern to all people with Broca's aphasia and found a similar influence of sentence complexity on comprehension performance in individuals with different aphasia types. Therefore, we decided not to restrict our sample to people with a specific type of aphasia.

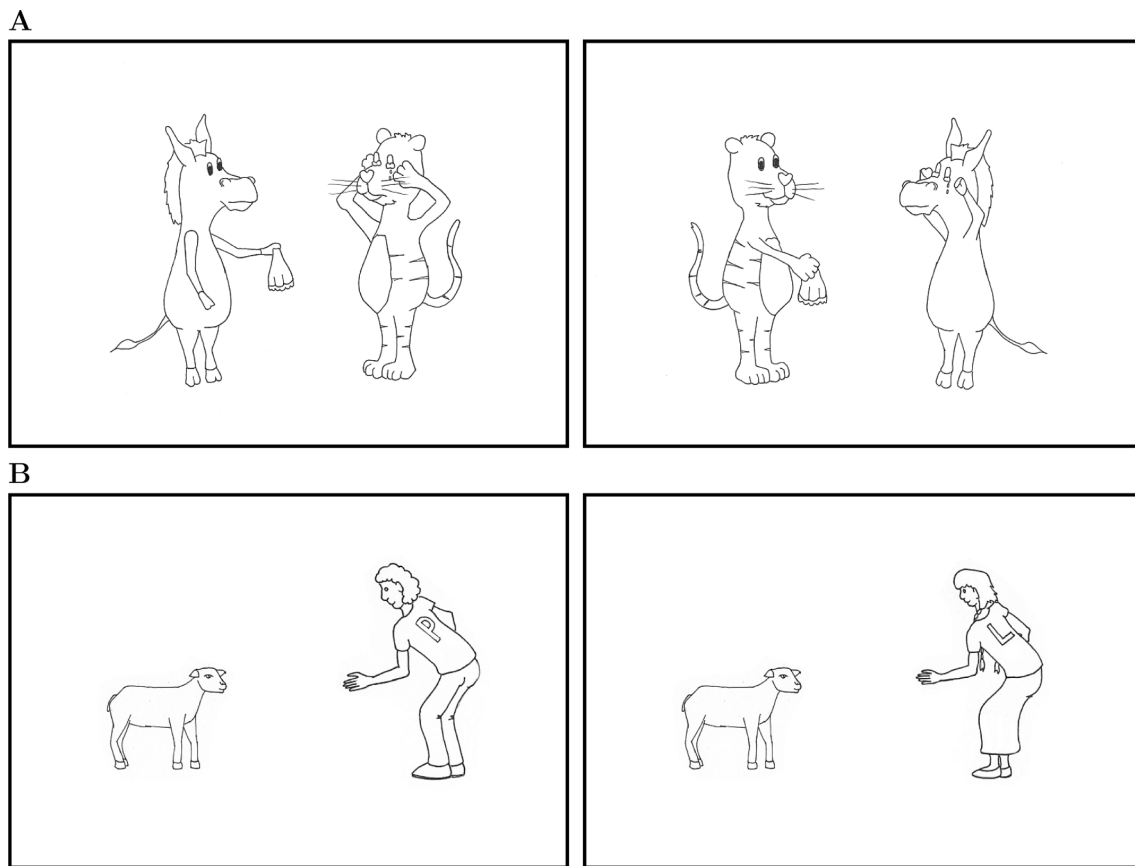


Fig. 2. Sample pictures of the part of the experiment with declaratives and relative clauses (A) and the part of the experiment with control structures (B). For the subject relative clause *Here is the tiger that_{nom} comforts the_{acc} donkey.*, the right picture A is the target and the left picture A the foil. For the object relative clause *Here is the tiger that_{acc} the_{nom} donkey comforts.*, the left picture A is the target and the right picture A the foil. For the object control sentence *Peter allows Lisa to pet the lamb.*, the right picture B is the target and the left picture B the foil. For the subject control sentence *Peter promises Lisa to pet the lamb.*, the left picture B is the target and the right picture B the foil.

the type and severity of aphasia (see Table 1). All IWA showed good auditory processing abilities for single nouns, assessed with an auditory word-picture matching task (all scores at least 90% correct) and a lexical decision task (all scores at least 88% correct) of the German psycholinguistic test battery LEMO 2.0 (Stadie et al., 2013). Although accuracy in the lexical decision task was lower in IWA compared to the control group, both participant groups were similarly influenced by psycholinguistic variables: Both groups gave faster responses for words than for non-words (lexicity effect), for high-frequency than for low-frequency words (frequency effect), and for concrete than for abstract words (effect of abstractness). Six additional IWA were excluded prior to data analysis due to no apparent aphasia in the Aachen Aphasia Test (3 participants), less than 90% accuracy in auditory word-picture matching (2 participants), or withdrawal (1 participant).

2.2. Procedure

This visual world experiment was part of a larger number of experiments that were carried out in a pseudo-randomized order with the same participants. All experiments were administered twice, i.e., in a test and retest phase spaced approximately two months apart. The specifics of the overall structure of the study are provided in Pregla et al. (2021).

The visual world experiment had two parts. The first part investigated the comprehension of control structures (see part *control structures* in the *materials* section), and the second part investigated the comprehension of declaratives and relative clauses (see part *declaratives and relative clauses* in the *materials* section). The two parts were presented to

participants in pseudo-randomized order. Both parts included five practice items for which feedback about response accuracy was provided, followed by the experimental items for which no feedback was provided. The part on control structures included one break after half of the items. The part on declaratives and relative clauses included breaks after each quarter of the items. Control participants and IWA completed the experiment in approximately 30 and 60 min respectively.

Prior to the experiment, participants were instructed that they were going to perform a sentence-picture matching task with two pictures and that their eye-movements would be recorded during the task. Items were presented in the following manner: 1) Preview of the pictures for 4000 ms and introduction of the displayed characters with a short sentence presented auditorily (e.g., *Hier sind Lisa und Peter.* 'Here are Lisa and Peter.' or *Hier sind Tiger und Esel.* 'Here are tigers and donkeys.'), 2) display of a central fixation cross for 500 ms, and 3), reappearance of the pictures and simultaneous auditory presentation of the experimental sentence. Pictures were shown until a picture was selected by the participant or for maximally 30 s. For picture selection, the lower left or right button on a Cedrus response pad (key layout RB-840) had to be pressed. In the experiment testing the comprehension of control structures, participants had to select the picture with the person (e.g., *Lisa*) that, according to the sentence, "does something with the animal". In the experiment testing the comprehension of declaratives and relative clauses, participants had to select the picture "that fits with the sentence" (see examples in Fig. 2). None of the participants had difficulties understanding the task or responding using the response pad.

A SensoMotoric Instruments (SMI RED250mobile) eye-tracker (binocular tracking, Experiment Center version 3.7, sampling rate

Table 2

Example of the control structures with PRO or an overt pronoun used in the experiment.

Condition	Sentence
s-ctrl	Peter _i verspricht nun Lisa _j PRO _i das kleine Lamm zu streicheln und zu kraulen.
(n = 10)	Peter _i promises now Lisa _j PRO _i to pet and to ruffle the little lamb.
o-ctrl	Peter _i erlaubt nun Lisa _j PRO _j das kleine Lamm zu streicheln und zu kraulen.
(n = 10)	Peter _i allows now Lisa _j PRO _j to pet and to ruffle the little lamb.
match	Peter _i verspricht nun Thomas _{MASC} dass er_i das kleine Lamm streichelt und krault.
(n = 10)	Peter _i promises now Thomas _{MASC} that he _i will pet and ruffle the little lamb.
mismatch	Peter _i verspricht nun Lisa _{FEM} dass er_i das kleine Lamm streichelt und krault.
(n = 10)	Peter _i promises now Lisa _{FEM} that he _i will pet and ruffle the little lamb.

Note. s-ctrl/o-ctrl = subject/object control, match/mismatch = gender match or mismatch of the main clause nouns. Critical region in bold.

Table 3

Example of the declaratives and relative clauses used in the experiment.

Sentence type	Condition	Sentence
Declaratives	SO	Hier tröstet der_{NOM} Tiger gerade den_{ACC} Esel
	(n = 10)	Here the _{NOM} tiger just comforts the _{ACC} donkey
	OS	Hier tröstet den_{ACC} Tiger gerade der_{NOM} Esel
Relative clause	(n = 10)	Here the _{ACC} tiger just comforts the _{NOM} donkey
	SRC	Hier ist der Tiger der_{NOM} den_{ACC} Esel gerade tröstet
	(n = 10)	Here is the tiger who _{NOM} comforts the _{ACC} donkey
	ORC	Hier ist der Tiger den_{ACC} der_{NOM} Esel gerade tröstet
	(n = 10)	Here is the tiger who _{ACC} the _{NOM} donkey comforts

Note. S = subject O = object, SRC/ORC = subject/object relative clause. Critical region in bold.

250 Hz) was used. Pictures were presented on a separate monitor (resolution: 1920 × 1080 pixels) on a grey screen with a distance of 60 pixels between the right border of the left picture and the left border of the right picture, which corresponded to a visual angle of 3°. Each picture subtended a visual angle of 37°. Participants were seated in front of the screen with a distance of approximately 60 cm. No chin-rest was used but participants were instructed to sit still. A 5-point calibration and validation were carried out before the practice phase, the test phase, and the second half of the test phase. If necessary, calibration could be manually initiated during the experiment. Both eyes were recorded and the fixation locations were determined based on the mean x and y coordinates of the eyes. Blinks, saccades, and fixations were detected with the velocity based algorithm of the SMI software BeGaze (version 3.7). Temporarily adjacent samples that did not exceed a velocity of 40°/s for at least 50 ms were treated as a fixation. Areas of interest (AoIs) consisted of the two pictures, and the number of fixations on the target picture (correct, counted as 1) in proportion to the fixations on the foil picture or no picture (counted as 0) was calculated. In the results section, the proportion of target fixations will be reported.

2.3. Materials

Below, the sentence structures, the auditory stimuli and the pictures will be presented.

2.3.1. Control structures

Examples for the sentences are given in Table 2 (for all items, see appendix). These sentences were used to test for the comprehension of subject and object control structures. In control structures, the subject of an embedded clause is identified with an argument of a matrix clause (Stiebels, 2007), i.e., the argument in the matrix clause controls the meaning of the subject in the embedded clause. Participants had to decide which of the arguments of the matrix clause, the subject or the

object, controls the subject of the embedded clause. This decision depended on the matrix clause verb that either led to a subject control interpretation (e.g., *versprechen*, ‘promise’) or an object control interpretation (e.g., *erlauben*, ‘allow’). The critical region of the sentence was the first phrase of the embedded clause. This phrase included the overt pronoun or PRO and thus was the point where the decision about the controlling argument should take place (highlighted in bold in Table 2).

A set of 50 control structures was used. In 20 sentences, the subject of the embedded clause was a pronoun controlled by the subject of the matrix clause (see Table 2, match and mismatch). In a further 20 sentences, the subject of the embedded clause was PRO, i.e., the pronoun was not pronounced overtly. PRO was controlled by the subject or the object in ten sentences respectively (see Table 2, s-ctrl and o-ctrl). Finally, ten filler sentences were included. Sentences were pseudo-randomized with at most three consecutive repetitions of the same sentence type.

To construct the sentences, 10 control verbs (5 subject control, 5 object control) with a mean lemma frequency of 4,713 ($SD = 2,146$) per million tokens in dlexDB (Heister et al., 2011) were used. In the sentences with PRO, control type was manipulated to vary the distance between the controlling argument and PRO. Based on earlier findings, subject control structures were regarded as complex because the distance between the controlling argument and PRO is longer than in object control structures (e.g., Kwon and Sturt, 2016; Caplan and Hildebrandt, 1988). In the sentences with a pronoun, only subject control verbs were used. The main clause nouns were common two-syllable German unambiguously male or female first names. In the sentences with PRO, nouns were always of different gender. In the sentences with a pronoun, the gender of the second noun of the matrix clause was manipulated such that it either matched or mismatched in gender with the first noun. This was done to manipulate the similarity of the nouns. Based on previous findings, sentences with gender-matching nouns were regarded as complex because the nouns are more similar than in sentences with gender-mismatching nouns (e.g., Stewart et al., 2000; Schroeder, 2007). Fillers included an object control verb and an overt pronoun (e.g. *Peter erlaubt nun Lisa, dass sie das kleine Lamm streichelt und krault.*, ‘Peter allows now Lisa that she pets and ruffles the little lamb.’).

2.3.2. Declaratives and relative clauses

Examples of the sentences are given in Table 3 (for all items, see appendix). In these sentences, the order of the nominative subject and the accusative object was varied. They were used to study the processing of canonical and non-canonical word order. In German, word order is canonical when the subject precedes the object, and it is non-canonical when the subject follows the object. Participants had to decide which of the arguments was the subject and the object. This decision depended on the case marking of the determiners and relative pronouns that were unambiguously marked for nominative case or accusative case. The critical region of the sentence was the phrase where the order of the arguments was disambiguated (highlighted in bold in Table 2). In the declaratives, this region consisted of the first noun phrase. In relative clauses, this region consisted of the relative pronoun.

A set of 80 sentences was used: 20 declaratives and 60 relative clauses.⁵ 10 declaratives had a canonical word order, i.e., the subject preceded the object (SO). The other 10 declaratives had a non-canonical word order, i.e., the subject followed the object (OS). Based on previous studies for German, SO declaratives will be regarded as simple and OS declaratives as complex (e.g., Hanne et al., 2011; Vogelzang et al., 2019). Relative clauses consisted of 30 subject and 30 object relative clauses. They were further divided into subject and object modifying relative clauses, and relative clauses with a plural noun (10 items respectively). In the present study, only the 20 subject modifying subject and object relative clauses with singular nouns were analyzed. The

⁵ This part of the experiment did not include filler sentences

analysis was restricted to these sentences because the study investigated the predictions of the RRH with respect to the processing of simple and complex sentences. The other conditions were included to test predictions with respect to changes of number and case between main clause and subclause which were not the focus of this paper.⁶ Based on previous findings for German, subject relative clauses will be regarded as simple and object relative clauses as complex (e.g., Adelt et al., 2017; Bader and Meng, 1999). Sentences were pseudo-randomized with a maximum of three consecutive repetitions of the same sentence type.

To construct the sentences, 10 transitive action verbs with two syllables and a mean lemma frequency of 85 ($SD = 211$) per million tokens in dlexDB (Heister et al., 2011) were used. The nouns referred to animals with masculine gender, and had a length of two-syllable and a mean lemma frequency of 356 ($SD = 400$) per million tokens in dlexDB (Heister et al., 2011). Twenty-three students rated the plausibility of the animals as agent or patient of the actions to ensure that all sentences were pragmatically reversible.

2.3.3. Auditory stimuli

Sentences were spoken with a neutral prosodic contour at a rate of 4.79 syllables per second in the experiment on declaratives and relative clauses and at a rate of 3.95 syllables per second in the experiment on control structures. These rates fall in the range of 3–6 syllables per second, which is considered a normal speech rate (Levett, 2001). Sentences were recorded in a sound-proof booth with a trained female native speaker of German. Recordings were post-processed with Praat (Boersma and Weenink, 2018). The same recordings were used for pairs of simple and complex sentences (e.g., subject and object relative clauses) by exchanging the manipulated region (e.g., *der* 'the.NOM' and *den* 'the.ACC') in the sound files.⁷

2.3.4. Pictures

Pictures consisted of pairs of black-and-white drawings. In the part of the experiment on declaratives and relative clauses, the target picture displayed the agent acting on the patient, and the foil picture displayed the referents with reversed thematic roles (e.g., Fig. 2, A). In the part of the experiment on control structures, target and foil picture displayed the target or distractor referent respectively interacting with the animal mentioned in the sentence (e.g., Fig. 2, B). Referents had the same size and adopted the same postures. Human referents were identifiable by their initials (e.g., *L* for *Lisa*). The action direction (from left to right or reversed) was balanced. Target and foil pictures were presented in the center of the screen adjacent to each other. The order of the pictures was counterbalanced so as to avoid any bias due to presentation order.

2.4. Data analysis

Data were analyzed separately for IWA and control participants, and for the four sentence structures (i.e., declaratives, relative clauses, control structures with an overt pronoun and control structures with PRO). The data of the test and retest were pooled, i.e., the statistical models included 20 observations per sentence type. The data of the two participant groups were not combined into one model to reduce computation time, which was up to a week for the models presented here. Blinks and saccades were excluded from the analyses. The data

⁶ Examples of these sentences are: Object modifying subject/ object relative clause: *Ich seh den Tiger, der den Esel gerade tröstet/ den der Esel gerade tröstet.*, 'I see the tiger who just comforts the donkey/ who the donkey just comforts.'. Subject modifying subject/ object relative clause with plural noun in the relative clause: *Hier ist der Tiger, der die Esel gerade tröstet/ den die Esel gerade trösten.*, 'I see the tiger who just comforts the donkeys/ who the donkeys just comfort.'.

⁷ It was checked in a pilot with four students and four elderly control participants that the spliced stimuli sounded natural.

were analyzed in two different ways: 1) Time bin analysis, in which the data were sliced in 50 ms time bins as done in a growth curve analysis (Mirman, 2014). This fine grained measure was used to determine where in the sentence a change in target fixations occurred at the group level. 2) Time window analysis, in which target fixations were averaged across three broad time windows and the two test phases. This broad measure was chosen to estimate the target fixations for each individual participant as recommended by McMurray (2020). It was not possible to determine the fine grained fixation path of each individual participant from the time bin analysis because the number of observations per participant was too low to get reliable participant-level estimates of the target fixations in each time bin. All data and code are available from <https://osf.io/mc2rn/>.

2.4.1. Time bin analysis

Analyses included all fixations from the onset of the critical region to a designated cutoff point after the end of the sentence. They were limited to this period to reduce computation time. In declaratives and relative clauses, the critical region was the first determiner or the relative pronoun (see region in bold in Table 3). In control structures, the critical region was the onset of the subclause (see region in bold in Table 2). The designated cut-off point after the end of the sentence was the mean reaction time of the respective participant group for the respective sentence type. Consequently, the analyses of the IWA include longer periods of silence than the analyses of the control participants because IWA had longer mean response times.

Fixations were averaged across 50 ms bins. About 99% of the obtained mean fixations were binary (i.e., 1 target fixated, 0 target not fixated), the remaining mean fixations were binarized (cf. Huang and Snedeker, 2020): If the mean proportion of target fixations in a particular bin was smaller or equal to 0.5, a 0 was inserted, otherwise, a 1 was inserted. The mean fixations were analyzed using R (Version 3.6.3; R Core Team, 2020) and the R-package *brms* (Version 2.17.0; Bürkner, 2017; Bürkner, 2018) with Bayesian hierarchical generalized linear mixed models with a logit link and full variance covariance matrices for the random effects of participants and items. Model estimates were back-transformed into proportions for ease of interpretation.

All models included the predictors COMPLEXITY, TEST PHASE and TIME and their interaction. The models of the IWA additionally included the predictor ACCURACY and interactions of all predictors. For COMPLEXITY, sum contrasts were used, where complex sentences were coded as -1 (i.e., OS declaratives, object relative clauses, subject control structures, and control structures with gender matching nouns) and simple sentences as $+1$ (i.e., SO declaratives, subject relative clauses, object control structures, and control structures with gender mismatching nouns). Similarly, a sum coding was used for TEST PHASE (-1 test, $+1$ retest) and ACCURACY ($+1$ correct, -1 incorrect). Following Mirman (2014), higher-order orthogonal polynomials were used for the predictor TIME to account for the fact that the change in proportion of target fixations over time is not linear. In all models, fourth order polynomials were used.

The prior distributions for the parameters in our models were specified as follows: The prior of the intercept was set to $Normal(0, 1.5)$, the priors of the slopes were set to $Normal(0, 1)$, and the prior standard deviations of the random effects to $Normal_+(0, 1)$ truncated at zero because standard deviations cannot be negative. The prior of the correlation between the random intercepts and slopes was set to $LKJ = 2$ (Lewandowski et al., 2009) to disfavor extreme correlations. The model output consisted of the posterior distributions of the parameters. The estimated 95% credible interval (CrI) of the posterior was extracted. The CrI is the range of plausible values of the parameters given the data and model.

The 95% CrIs were used to estimate the point in time of a divergence in proportion of target fixations between two conditions. These divergences were calculated to investigate the predictions of the RRH. More specifically, the following divergences were scrutinized based on the predictions: For prediction 1 (Normal-like processing in correct

Table 4

Mean and standard error of the accuracy (in %) and response times (in ms) in the sentence-picture matching task of the visual world experiment in individuals with aphasia and control participants.

		SO	OS	SRC	ORC	mis-match	match	o-ctrl	s-ctrl
Accuracy									
IWA	test	63.8 (3.3)	42.4 (3.5)	67.1 (3.2)	40.5 (3.4)	60.5 (3.5)	50.5 (3.6)	67.5 (3.3)	50.5 (3.6)
	retest	77.6 (2.8)	43.8 (3.5)	67.1 (3.2)	45.2 (3.5)	76.7 (2.9)	70.5 (3.2)	76.2 (2.9)	62.9 (3.3)
CP	test	99.1 (0.6)	89.1 (2.1)	96.4 (1.3)	92.3 (1.8)	100 (0)	96.8 (1.2)	99.5 (0.5)	96.8 (1.2)
	retest	98.2 (0.9)	93.2 (1.7)	96.8 (1.2)	95.9 (1.3)	100 (0)	98.6 (0.8)	100 (0)	99.1 (0.6)
Response time									
IWA	test	4116 (227)	4975 (297)	3935 (247)	4031 (233)	2553 (236)	3109 (300)	1922 (180)	2511 (223)
	retest	4502 (261)	5299 (294)	4252 (275)	4237 (231)	2061 (203)	2936 (279)	2135 (191)	2344 (233)
CP	test	976 (75)	1875 (130)	1421 (124)	1695 (116)	400 (42)	625 (74)	500 (88)	463 (56)
	retest	885 (60)	1495 (100)	1372 (123)	1445 (110)	398 (36)	499 (43)	371 (34)	439 (39)

Note. IWA = individuals with aphasia, CP = control participants, SO/ OS = declarative sentence with canonical/ non-canonical word order, SRC/ ORC = subject/ object relative clause, match/ mismatch = gender of the main clause nouns is the same/ different, s-ctrl/ o-ctrl = subject/ object control.

trials), it was checked whether there was a divergence in target fixations between control participants and IWA, a divergence in target fixations between the correct and incorrect trials of the IWA, and a divergence from 50% target fixations (i.e., the point in time where participants started to fixate the target picture more than the foil picture, cf. [Wendt et al., 2014](#)). For prediction 2 (Processing difficulty in complex vs. simple sentences, complexity-capacity interaction), it was checked whether there was a divergence in target fixations between simple and complex sentences, and an interaction between target fixations and response accuracy (the latter analysis is described in detail in the section *Time window analysis*). For prediction 3 (Unsystematic variability in the performance between test and retest), it was checked whether there was a divergence in target fixations between test and retest.

To be counted as a divergence, the 95% CrIs of the respective two conditions were not allowed to overlap for at least 4 consecutive time bins (i.e., 200 ms). To determine a confidence interval (CI) for a divergence point, bootstrapping analyses were carried out ([Stone et al., 2020](#)). Different from [Stone et al. \(2020\)](#), we did not fit t-tests for each time bin to determine divergence between two conditions but we used the 95% CrIs of the models previously fit with *brms*. Thus, to determine the CIs for the divergence points we only had to fit one Bayesian model for each sentence type instead of multiple t-tests. The 95% CrIs were resampled for each participant in each time bin, and the divergence between CrIs was calculated for the resampled data. Resampling was done 2000 times to generate a distribution of divergence points.

2.4.2. Time window analysis

This analysis was carried out to test the prediction of the RRH that there is an interaction between resource capacity of an IWA and the complexity of the sentence structure. For this analysis, the data of the test and retest were pooled and trials were divided into three regions of interest: 1) the first half of the target sentence up to and including the critical region, 2) the second half of the target sentence, and 3) the silence region after the sentence until the response key was pressed. For each region, the sum of target fixations and the total number of fixations in each trial was calculated. The sum of target fixations and the total number of fixations were entered as the dependent variables of binomial models with a logit link which were fit in *brms*. Model estimates were back-transformed into proportions for ease of interpretation.

The models included the following predictors: COMPLEXITY, ACCURACY, OVERALL ACCURACY and their interactions. The predictors COMPLEXITY and ACCURACY were sum coded (+1 simple, -1 complex; +1 correct, -1 incorrect). For the predictor OVERALL ACCURACY, the overall response accuracy of each IWA for each of the four sentence types were calculated. The response accuracy was then centered per sentence type, i.e., per sentence type, the average response accuracy was subtracted from the response accuracy of each IWA ([Schad et al., 2020](#)). In an additional model, OVERALL ACCURACY was replaced by SEVERITY which was the centered severity of each IWA in the Aachen Aphasia Test (see [Table 1](#)). The same

priors as in the time bin analyses were used.

3. Results

First, the results of the time bin analyses for the two participant groups will be reported. Afterwards, the results of the time window analyses for each single IWA will be presented. Accuracy and response times of the sentence-picture matching task have been analyzed and reported in [Pregla et al. \(2021\)](#). We will give a summary of the offline results before turning to the target fixations.

3.1. Summary of the offline results

Accuracy and response times are summarized in [Table 4](#). Control participants responded faster and displayed more correct responses than IWA. Both participant groups responded faster and displayed more correct responses in simple versus complex sentences, and in the retest versus the test phase. As visible in [Table 4](#), the response accuracy of the IWA was at or below 50% in OS declaratives, object relative clauses, and in the complex control structures (i.e., match and subject control) in the test phase. This result is addressed in the discussion.

3.2. Results of the Time Bin Analyses

The fixation paths in correct trials of the two participant groups are shown in [Fig. 3](#). The fixation paths of correct versus incorrect trials of the IWA are shown in [Fig. 4](#). In the following, the results are presented according to the ordering of the three predictions of the RRH as outlined in the theoretical background.

3.2.1. Normal-like processing in correct trials

This prediction of the RRH was tested with the following comparisons: 1) comparisons of the fixation paths of IWA and the control participants for each sentence type and test phase in correct trials, 2) comparison of the fixation paths against a threshold of 50% target fixations (i.e., the threshold above which participants fixated the target picture more than the foil picture) for each sentence type, test phase and participant group in correct trials, and 3) comparisons of the fixation paths in correct and incorrect trials for each sentence type and test phase in the IWA.

1) *Divergence between the participant groups*: The increases in target fixations in correct trials were greater in control participants than in IWA. Control participants' target fixations exceeded the IWA's target fixations in all sentence structures except subject control structures, and subject relative clauses in the test phase. The divergence between the groups started less than two seconds after the critical region, which was before or at the sentence end (estimates of the divergence onsets see [Table A.1](#) in the appendix).

2) *Divergence from 50% target fixations*: In both participant groups,

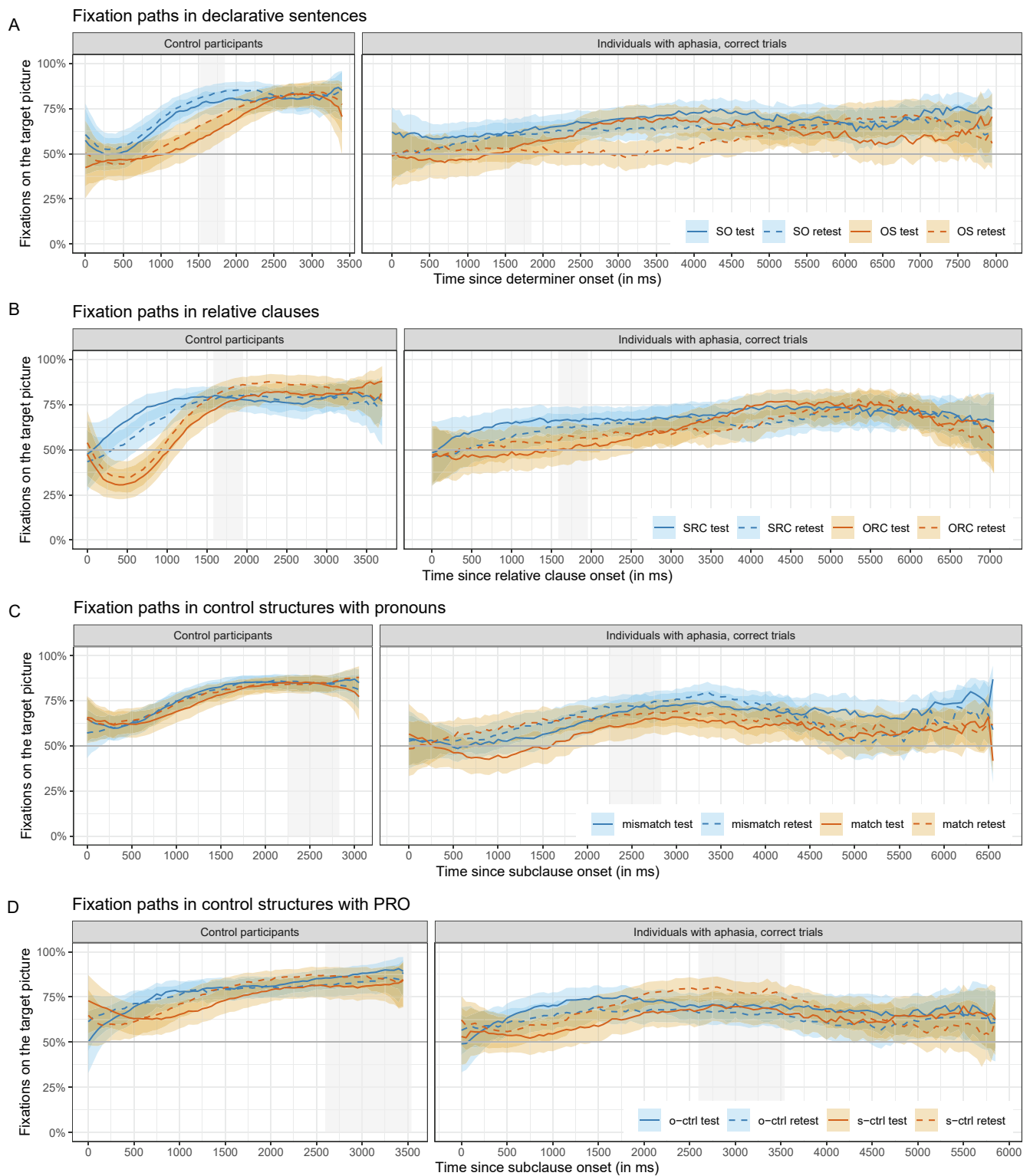


Fig. 3. Estimated fixation curves of the correct trials of the control participants and the individuals with aphasia within the time frame from the onset of the critical region until the response key was pressed. A: canonical (SO) and non-canonical (OS) declaratives; B: subject (SRC) and object (ORC) relative clauses; C: control structures with a pronoun with gender matching (match) and mismatching (mismatch) nouns; D: subject (s-ctrl) and object (o-ctrl) control structures with PRO. Solid and dashed lines represent the mean fixations in simple and complex sentences respectively and shaded areas represent the 95% credible intervals around the mean. Vertical bands shaded in grey mark the sentence end.

Fixation paths of the individuals with aphasia: correct versus incorrect trials

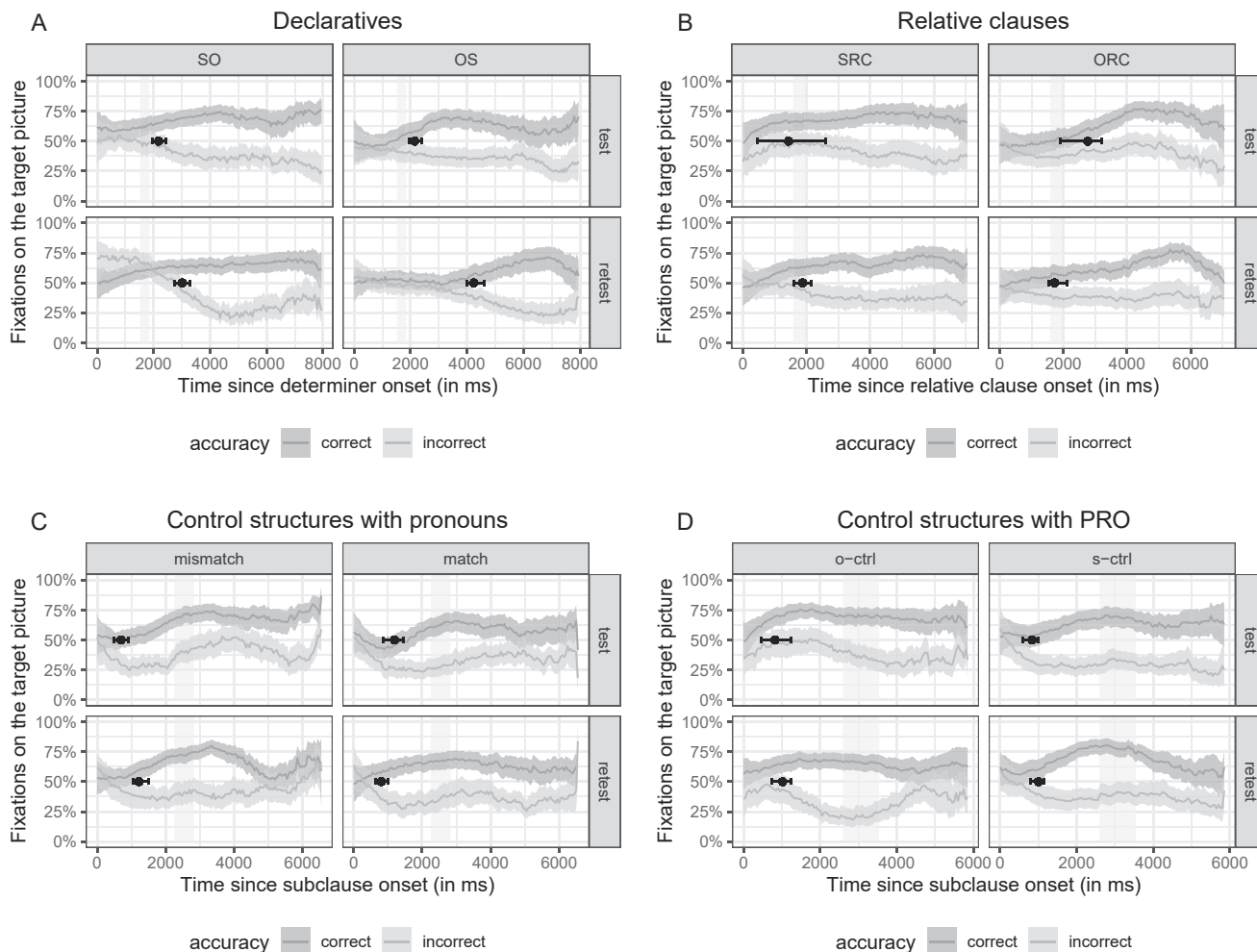


Fig. 4. Estimated fixation curves of the individuals with aphasia for the time frame from the onset of the critical region until the response key was pressed. A: canonical (SO) and non-canonical (OS) declaratives; B: subject (SRC) and object (ORC) relative clauses; C: control structures with a pronoun with gender matching (match) and mismatching (mismatch) nouns; D: subject (s-ctrl) and object (o-ctrl) control structures with PRO. Solid dark grey and light grey lines represent the mean fixations in correct and incorrect trials and shaded areas represent the 95% credible intervals around the mean. Dots represent the divergence onsets between correct and incorrect trials. Error bars represent bootstrapped confidence intervals. Vertical bands shaded in grey mark the sentence end. The width of these bands varies because audio files were not of equal length, and therefore, the sentence end can lie somewhere in between these bands. The minimum and maximum audio file length varies per sentence type, i.e., the minimum and maximum audio file length is different in declaratives, relative clauses, control structures with a pronoun and control structures with PRO. As such, the width of the band is different for each sentence type.

Bootstrapped onsets of the divergences from 50% target fixations in individuals with aphasia and control participants

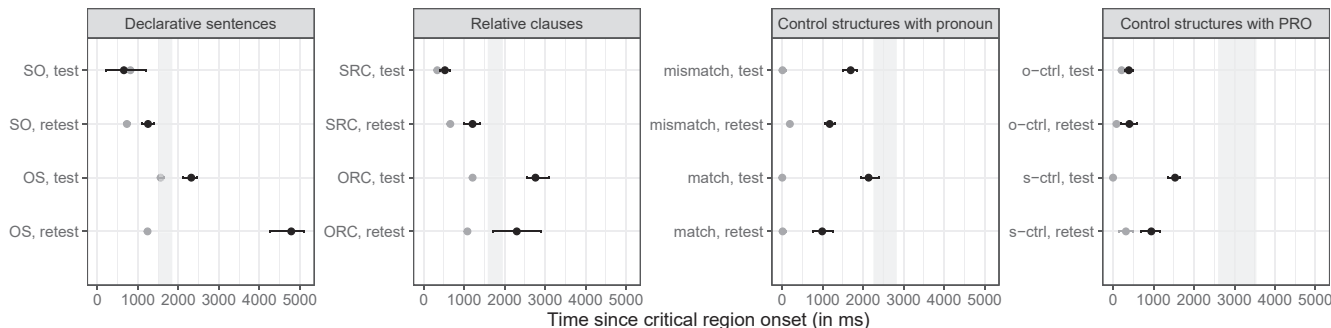


Fig. 5. Bootstrapped means (dots) and 95% confidence intervals (error bars) of the divergence onsets from 50% target fixations. Divergence onsets are shown separately for each test phase (test, retest) for canonical and non-canonical declaratives (SO and OS), subject and object relative clauses (SRC and ORC), control structures with gender matching and mismatching nouns with a pronoun (match and mismatch), and subject and object control structures with PRO (s-ctrl and o-ctrl). Vertical bands shaded in grey mark the sentence end.

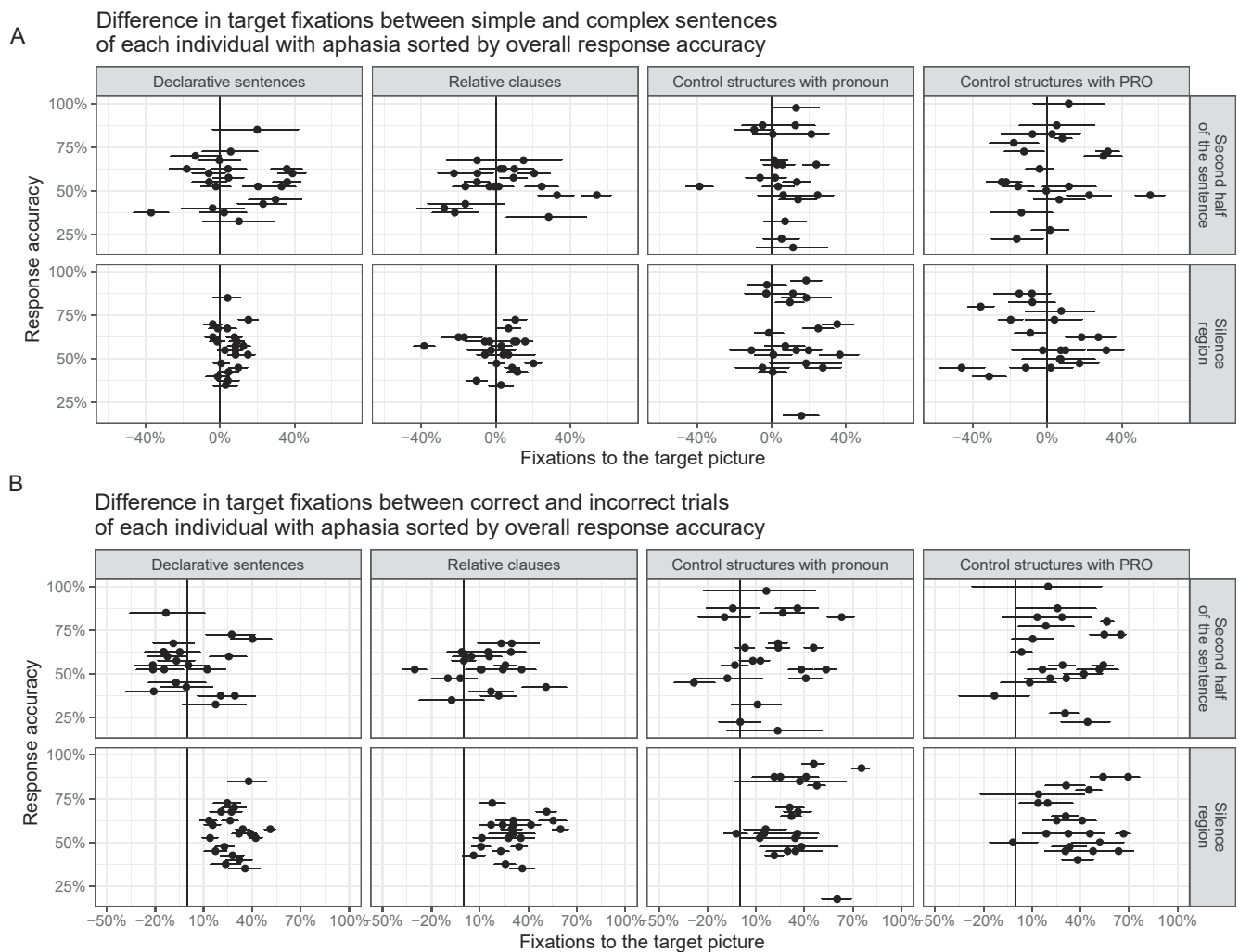


Fig. 6. Mean estimates (dots) and 95% credible intervals (horizontal lines) of (A) the difference in target fixations between simple and complex sentences and (B) the difference in target fixations between correct and incorrect trial in each individual with aphasia in the four investigated sentence types in the second half of the sentence after the critical word and in the silence region. Participants are displayed in descending order by their overall response accuracy in the respective sentence type. Distributions that are right-shifted denote higher proportions of target fixations in simpler sentences (A) or correct trials (B).

the fixation curves of the correct trials exceeded the 50% threshold in all sentence structures (estimates of the divergence onsets see Fig. 5 and Table A.1 in the appendix). With the exception of SO declaratives in the test phase, subject relative clauses in the test phase, and the subject control structures, the fixation paths of the control participants exceeded the 50% threshold earlier than the fixation paths of the IWA. In both participant groups, target fixations diverged from 50% earlier in the simple sentences than in the complex sentences. This was the case in the declaratives and relative clauses in the control participants and in all sentence types except for control structures with a pronoun in the IWA.

3) Divergence between correct and incorrect trials: In all sentence types and both test phases, IWA showed more target fixations in correct versus incorrect trials. Divergences occurred earlier in control structures with a pronoun or PRO than in declaratives and relative clauses (see Fig. 4, estimates of the divergence onsets see Table A.1 in the appendix). In all sentence types, the differences in target fixations between correct and incorrect trials were long lasting, extending over a period of at least two seconds.

3.2.2. Processing difficulty in complex vs. simple sentences, complexity-capacity interaction

This prediction of the RRH was tested by the juxtaposition of fixation paths in simple as opposed to complex sentences for each sentence type,

test phase and participant group in correct trials. Furthermore, the RRH predicts an interaction of sentence complexity and resource capacity of the IWA, which will be investigated in the section *Results of the Time Window Analyses*.

In the control participants, the fixation paths of the simple sentences exceeded the fixation paths of the complex sentences in declaratives and relative clauses in both test phases (divergence onsets: declaratives, test: 1008 ms CI: [950, 1100], retest: 1053 ms CI: [950, 1250], relative clauses, test: 223 ms CI: [200, 250], retest: 477 ms CI: [400, 600]). In the correct trials of the IWA, a divergence between simple and complex sentences occurred only in declaratives in the retest (divergence onset: 2847 ms CI: [2200, 3250]).

3.2.3. Unsystematic variability in the performance between test and retest

This prediction of the RRH was tested by comparing the fixation paths in the two test phases for each sentence type and participant group in correct trials.

In both participant groups, the fixation paths of the correct trials overlapped in test and retest. There was one exception: IWA showed earlier increases in target fixations in the test phase compared to the retest phase in OS declaratives (divergence onset: 2860 ms CI: [2550, 3150]).

Table 5

Means and 95% Credible Intervals (CrI) for the influence of the overall response accuracy in the sentence type on the target fixations in the second half of the sentence and in the silence region.

	Overall Accuracy	Overall Acc × Complexity	Overall Acc × Trial Acc
Region 2			
declaratives	0.3% CrI: [-1.5, 2.1]	0.1% CrI: [-1.9, 2.1]	-0.1% CrI: [-2, 1.9]
RC	-0.5% CrI: [-2.4, 1.3]	1.4% CrI: [-1.6, 4.3]	-1.8% CrI: [-4.4, 0.7]
pronoun	0.5% CrI: [-0.8, 1.8]	0.5% CrI: [-0.5, 1.5]	0.3% CrI: [-1.3, 2]
PRO	0.2% CrI: [-2.3, 2.6]	-0.3% CrI: [-2.2, 1.6]	1.3% CrI: [-1.4, 3.8]
Region 3			
declaratives	0.9% CrI: [-0.1, 1.8]	-0.4% CrI: [-0.9, 0.3]	0.5% CrI: [-0.7, 1.7]
RC	0.9% CrI: [-0.4, 2.3]	-0.6% CrI: [-2.4, 1.2]	0.1% CrI: [-2, 2.3]
pronoun	0.6% CrI: [-1.3, 2.4]	0.1% CrI: [-0.9, 1.1]	0.7% CrI: [-0.8, 2.2]
PRO	0.3% CrI: [-1.4, 1.8]	0.1% CrI: [-1.7, 1.9]	1.5% CrI: [-0.6, 3.5]

Note. Acc = Accuracy, Region 2 = second half of the sentence after the critical word, Region 3 = silence region, RC = relative clauses, pronoun = control structures with a pronoun, PRO = control structures with PRO.

3.3. Results of the Time Window Analyses

The time window analyses were carried out to investigate the relationship between overall comprehension accuracy of each IWA in a sentence structure and their target fixations. The analysis was based on the prediction of the RRH that there is a complexity-capacity interaction. The results are visualized in Fig. 6.

Fig. 6 A illustrates the relationship between overall response accuracy of each IWA and their differences in target fixations between simple and complex sentences in the second half of the sentence after the critical word and in the silence region. The interactions between overall response accuracy and sentence complexity were uninformative in all sentence types (for the estimates see Table 5). Fig. 6 B shows the relationship between overall response accuracy of each IWA and their differences in target fixations between correctly and incorrectly answered trials in the second half of the sentence after the critical word and in the silence region. As it can be seen, there was no indication that overall response accuracy systematically influenced the differences in target fixations in correct versus incorrect trials (for the estimates see Table 5). Rather, in the silence region, distributions were shifted to the right in all IWA and sentence types as visible in the lower part of Fig. 6 B. This means that all IWA fixated the target picture more in trials in which they answered correctly across sentence types.

3.3.1. Additional Time Window Analysis

In addition to the analyses above, an analysis that was not based on the predictions of the RRH was carried out in order to test whether there was a relationship between the severity grade measured with the Aachen Aphasia Test (see Table 1) and the individual target fixations in the second half of the sentence after the critical word or in the silence region. In our group of IWA with an aphasia severity grade ranging from mild to moderate there was no indication that the severity grade influenced the overall amount of target fixations or the differences in target fixations between simple and complex as well as between correct and incorrect trials (for the estimates see Table A.2 in the appendix). However, it cannot be ruled out that there is an influence of severity on the target fixations for a group of IWA with a wider range of severity levels.

4. Discussion

This study investigated predictions of the RRH (Caplan, 2012) regarding sentence processing in IWA. Sentence processing abilities were assessed with an auditory sentence-picture matching task by measuring the proportion of target fixations in the visual world paradigm. Fixation patterns were investigated across two test phases and four sentence types.

Before we discuss the fixation patterns with respect to the predictions of the RRH, it is important to check whether the response accuracies of the IWA in our study are representative for IWA. This validation check can be carried out by comparing our response data to previous visual world studies. We show below that our accuracy data exhibit very similar patterns to the patterns observed in 13 previously published visual world studies.

After presenting the validation check, we will discuss the fixation patterns with respect to the three investigated predictions of the RRH, namely: 1) Normal-like processing in correct trials, 2) Processing difficulty in complex vs. simple sentences, and complexity-capacity interaction, and 3) Variability in the performance between test and retest due to random noise.

4.1. Validation check: Comparison of the accuracy in this study to that of previous studies

As shown in Table 4, the accuracy of the IWA in the current study was at chance in the comprehension of complex sentences. To exclude the conclusion that this performance of the IWA was exceptionally low, we did a comparison with studies using similar tasks in the visual world paradigm with similar participants. The comparison included accuracy data in the comprehension of several sentences types from the following visual world studies: Adelt et al. (2017), Bos et al. (2014), Choy and Thompson (2005), Choy and Thompson (2010), Dickey et al. (2007), Dickey and Thompson (2009), Engel et al. (2018), Hanne et al. (2011), Hanne et al. (2015), Hanne et al. (2016), Mack et al. (2016), Mack and Thompson (2017), Meyer et al. (2012), Sheppard et al. (2015), Schumacher et al. (2015), Thompson et al. (2004). The extracted accuracies are provided in Table A.3 in the appendix. As shown in Fig. 7, the accuracies of the IWA in this study are within the range of accuracies of the IWA in previous studies. A linear model was fit with *lme4* (Bates et al., 2015) to the arcsine-transformed mean accuracy with study as random effect. According to the model, the mean accuracy of our study (65%, coded as 1) was not significantly different from the mean accuracy of earlier studies (60%, coded as -1, $\hat{\beta} = -0.04$, $SE = 0.05$, $t = -0.67$). This shows that there is no evidence that the accuracies in our study are atypical in any respect.

4.2. Processing in correct trials

According to the RRH, processing in correct trials should be normal-like. Although the accuracy of the IWA lies within chance range, the observation of an increase in target fixations above 50% speaks against guessing and in favor of normal-like processing. This assumption is further supported by the fact that the increase occurred early (on average 2 s after onset of the critical region, estimates see Table A.1) during the trial and not shortly before response selection in all sentence types of both test phases (Hanne et al., 2011; Burchert et al., 2013). Furthermore the early and stable difference in target fixations between correct and incorrect trials corroborates the notion of normal-like target decision (Hanne et al., 2011; Burchert et al., 2013). As in Hanne et al. (2012), the observation that each individual IWA displayed these differences in correct and incorrect trials (irrespective of the overall response accuracy) further advocates the assumption of normal-like processing (see Fig. 6 B).

In addition to normal-like processing, it was predicted that

Response accuracy of the individuals with aphasia in this study and previous visual world studies on sentence comprehension in aphasia

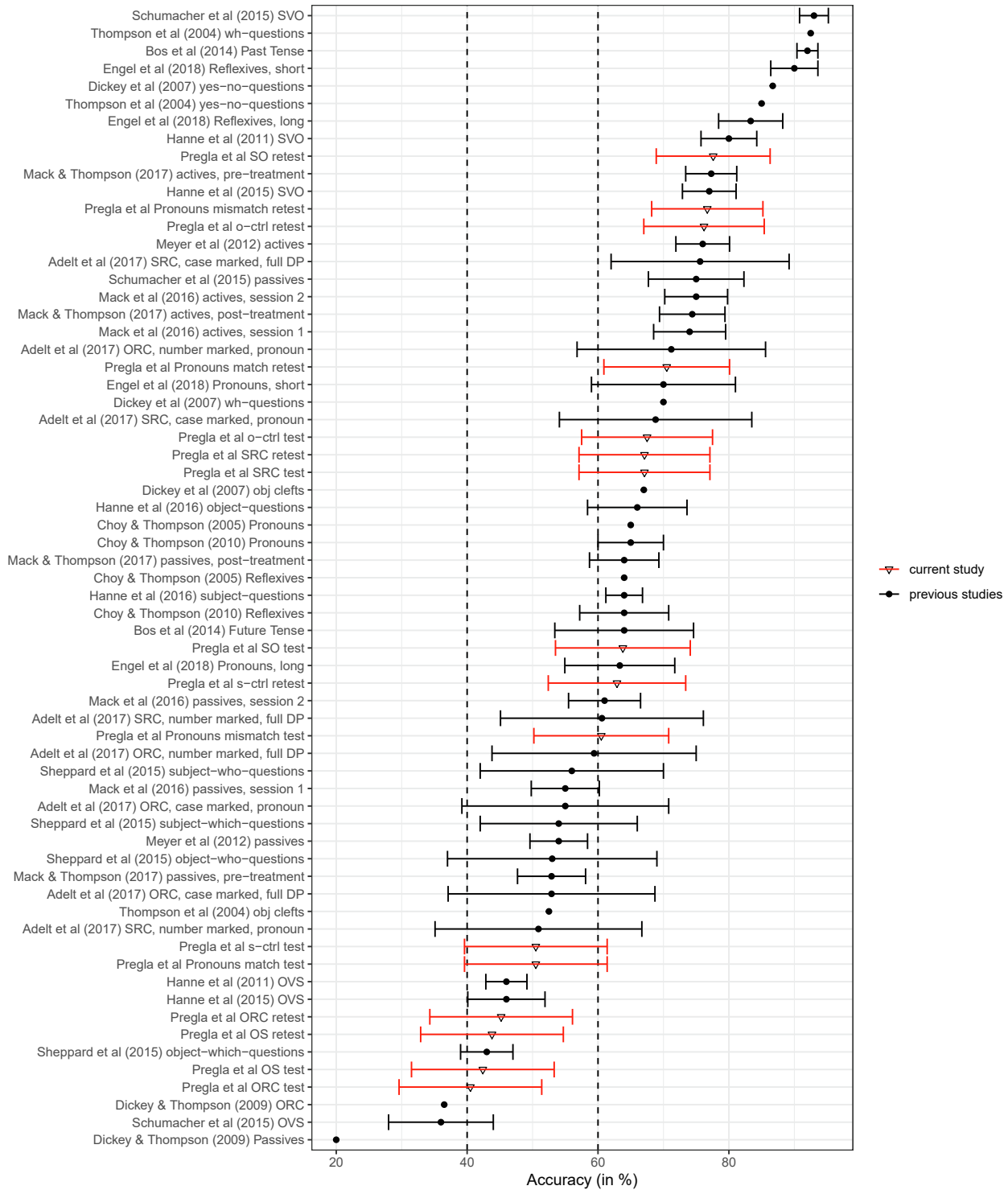


Fig. 7. Response accuracy of the individuals with aphasia in the current and previous visual world studies on sentence comprehension in aphasia sorted increasingly by mean accuracy. Dots and triangles represent the mean accuracy and error bars represent standard errors (if error bars are missing, standard errors could not be derived from the information provided in the study). Dashed lines mark an area of 40–60% accuracy which would be the chance area for a task in which the probability of getting a correct response by guessing is 50% and the number of items is 100.

processing speed is slowed down in IWA. To evaluate this prediction, the fixation paths of the IWA were compared to those of the control participants. Similar to previous studies (e.g., Mack et al., 2016; Meyer et al., 2012), IWA showed later increases in target fixations than control participants (i.e., the lower bound of the 95% CrI estimated for the

fixation paths of the control participants exceeded the upper bound of the 95% CrI estimated for the fixation paths of the IWA). This suggests that IWA do not process morpho-syntactic information as a cue to sentence processing as rapidly as control participants. Furthermore, the delayed increase in target fixations was visible across sentence

Table 6
Predictions of the resource reduction hypothesis, their expected expression in the visual world experiment and actual findings.

Predictions of the resource reduction hypothesis, expected fixation pattern for individuals with aphasia	Findings consistent with predictions?
1) Normal-like processing in correct trials	
– increases over 50% in target fixations in correct trials	yes, but reduced magnitude in target fixations
– higher increases in target fixations in correct vs. incorrect trials	yes
– slower increase in target fixations compared to control participants	yes
2) Processing difficulty in complex vs. simple sentences, complexity-capacity interaction	
– higher increases in target fixations in simple vs. complex sentences	no*, similar fixation paths in correct trials
– interaction complexity effect and overall response accuracy	inconclusive
3) Unsystematic variability in the performance between test and retest	
– unsystematic changes in fixation paths between test and retest	yes

Note. *The predicted pattern was only observed in declaratives in the retest phase.

structures with different types of morpho-syntactic information, namely case information (declaratives, relative clauses), gender information (control structures with a pronoun), and information about the verb's control type (control structures with a pronoun or PRO). Thus, it does not seem to be one specific type of morpho-syntactic information that leads to sentence processing difficulty in IWA. Rather, morpho-syntactic processing in IWA seems to be slowed down in general. This finding is in line with the RRH under the assumption that reduced resources in IWA are reflected by a reduction of processing speed, an assumption that was also put forward by Caplan et al. (2015).

Finally, the increase of target fixations in IWA was not as pronounced as in language unimpaired control participants. This is not in line with Nozari et al. (2016) who found similar increases in both in language-impaired and-unimpaired participants, however the increase was delayed in IWA. Following the reasoning of Nozari et al. (2016), if IWA would trade speed for accuracy target fixations should increase more slowly but to the same maximum as in control participants. Our finding of a less noticeable increase in target fixations in addition to a delay might

Table A.1

Bootstrapped onsets of the divergence of the fixation path (1) between participant groups, (2) from 50% target fixations, and (3) between correctly and incorrectly answered trials. Divergence onsets are represented as mean in milliseconds with a bootstrapped 95% confidence interval [CI].

	Declarative sentences	Relative clauses	Control structures with pronoun	Control structures with PRO
(1) Group				
test	s: 1535 [1400, 1800] c: 2243 [2100, 2500]	s: NA c: 1414 [1300, 1550]	s: 677 [500, 800] c: 626 [500, 750]	s: 2456 [2300, 2750] c: NA
retest	s: 1138 [1050, 1250] c: 1705 [1600, 1800]	s: 1367 [1100, 1650] c: 1333 [1250, 1450]	s: 760 [650, 850] c: 1101 [950, 1250]	s: 2046 [1100, 2750] c: NA
(2) 50%				
IWA: test	s: 653 [200, 1200] c: 2317 [2100, 2450]	s: 524 [400, 650] c: 2760 [2550, 3100]	s: 1690 [1500, 1850] c: 2133 [1950, 2400]	s: 389 [300, 500] c: 1532 [1350, 1650]
IWA: retest	s: 1248 [1100, 1400] c: 4787 [4250, 5100]	s: 1206 [1000, 1400] c: 2299 [1700, 2900]	s: 1172 [1050, 1300] c: 987 [750, 1250]	s: 404 [200, 600] c: 945 [700, 1150]
CP: test	s: 814 [750, 850] c: 1560 [1500, 1650]	s: 334 [300, 350] c: 1207 [1150, 1250]	s: 7 [0, 100] c: 1 [0, 0]	s: 211 [150, 250] c: 0 [0, 0]
CP: retest	s: 726 [650, 800] c: 1237 [1200, 1300]	s: 656 [600, 700] c: 1080 [1050, 1100]	s: 189 [150, 250] c: 10 [0, 100]	s: 86 [50, 150] c: 318 [150, 500]
(3) Accuracy				
IWA: test	s: 2176 [1950, 2450] c: 2143 [1950, 2400]	s: 1435 [450, 2600] c: 2762 [1900, 3200]	s: 699 [500, 900] c: 1189 [850, 1450]	s: 819 [450, 1250] c: 834 [600, 1000]
IWA: retest	s: 3013 [2750, 3300] c: 4245 [4000, 4600]	s: 1872 [1600, 2150] c: 1725 [1550, 2100]	s: 1225 [1050, 1500] c: 808 [650, 1000]	s: 1022 [750, 1250] c: 1001 [800, 1150]

Note. IWA = individuals with aphasia, CP = control participants, s = simple canonical or low interference conditions, c = complex non-canonical or high interference conditions, NA = no divergence.

suggest that in IWA the decision for the target picture was taken with less certainty. We will elaborate on what might lead to the reduced certainty in picture selection in the summary section below.

Overall, the data are consistent with the general conclusion of visual world studies in aphasia, namely that IWA do not guess and deliberately decide on the target picture in correct trials. This conclusion was confirmed at the group level and the individual participant level. This result is in line with the prediction of the RRH that processing in correct trials is normal-like.

4.3. Processing of complex sentences

According to the RRH, participants should have more processing difficulty in complex vs. simple sentences, and there should be an interaction between sentence complexity and resource capacity. To test this prediction, sentence complexity (i.e., canonicity, similarity of noun phrases, dependency length) in different sentence types (i.e., declaratives, relative clauses, control structures with pronoun or PRO) was varied, and the fixation paths of the simple and complex sentences in each sentence structure were compared.

In the control structures, neither the control participants nor the IWA showed differences in target fixations between the simple and complex sentences. Control structures with pronouns were regarded as simple when the gender of the pronoun's antecedent and a distractor noun mismatched and as complex when the gender of the two nouns matched. Control structures with PRO were regarded as simple when the antecedent directly preceded PRO (object control) and as complex when a noun intervened between the antecedent and PRO (subject control). Irrespective of the pronoun type, target fixations overlapped between the simple and complex sentences. Similar results have been obtained for language-unimpaired participants for reflexive pronouns, in which the distractor noun also did not influence pronoun resolution (Dillon et al., 2013; Schroeder, 2007; Sturt, 2003). A possible explanation for the lack of influence of the distractor might be that only antecedents accessible for binding are considered during pronoun resolution (Sturt, 2003).

In the declaratives and relative clauses, control participants showed differences between sentences with a canonical and non-canonical word order in both test phases. That is, irrespective of sentence type, there were more target fixations in canonical sentences than in non-canonical sentences. As in previous studies (e.g., Hanne et al., 2015; Mack et al., 2016; Meyer et al., 2012), these differences in fixations between

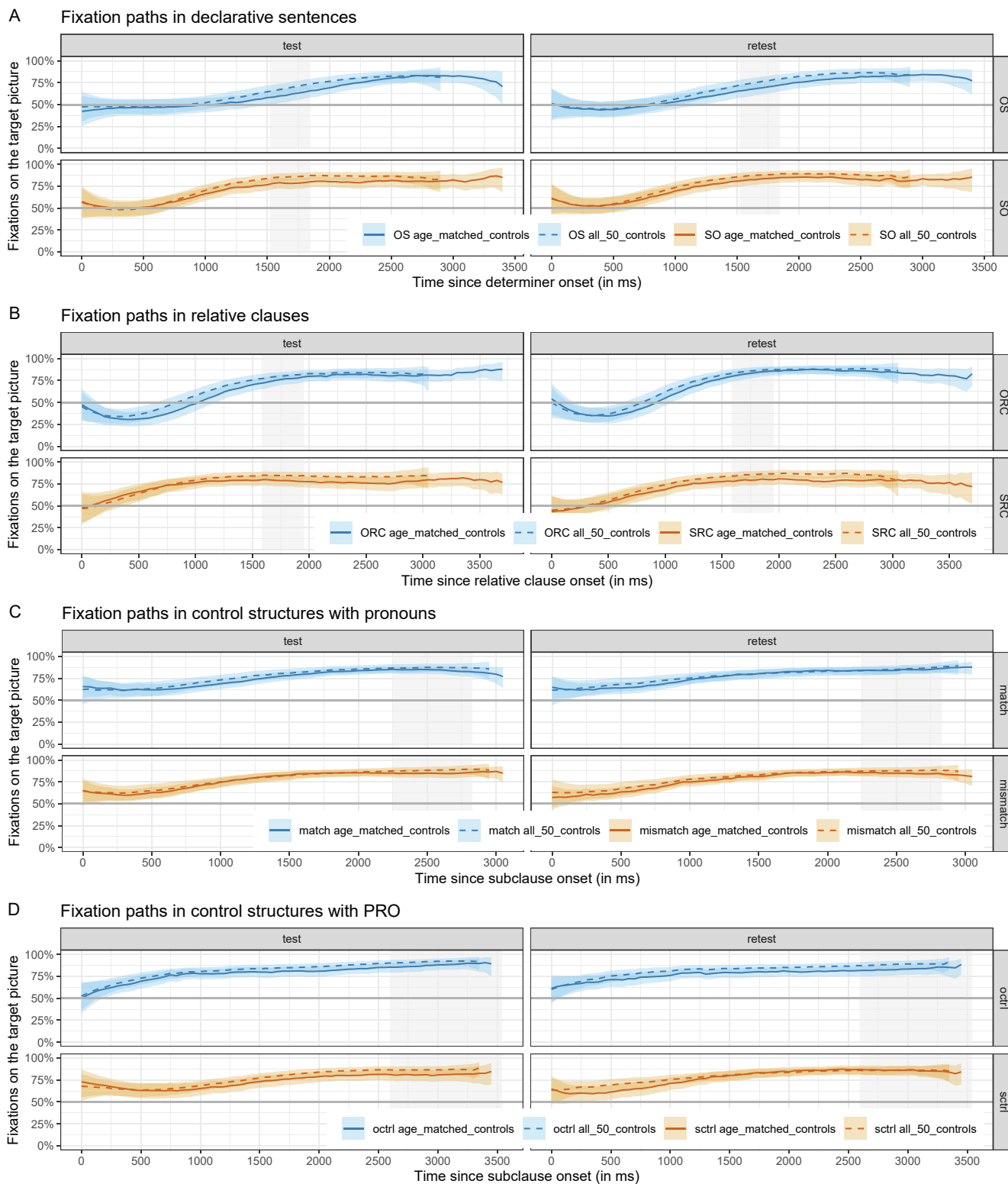


Fig. A.1. Estimated fixation paths of the whole control group ($n = 50$ participants, mean age = 48 years, $SD = 20$, range = 19–83; mean education = 18 years, $SD = 4$, range = 6–26, dashed lines) and the age and education matched control group ($n = 22$ participants, mean age = 58 years, $SD = 15$, range = 26–81; mean education = 16 years, $SD = 4$, range = 6–21, solid lines). A: canonical (SO) and non-canonical (OS) declaratives; B: subject (SRC) and object (ORC) relative clauses; C: control structures with a pronoun with gender matching (match) and mismatching (mismatch) nouns; D: subject (s-ctrl) and object (o-ctrl) control structures with PRO. Solid and dashed lines represent the mean fixations in simple and complex sentences respectively and shaded areas represent the 95% credible intervals around the mean. Vertical bands shaded in grey mark the sentence end.

Table A.2

Means and 95% Credible Intervals (CrI) for the influence of the aphasia severity grade in the Aachen Aphasia Test (AAT) on the target fixations in the second half of the sentence and in the silence region.

	Severity	Severity × Complexity	Severity × Trial Accuracy
Region 2			
declaratives	0.2% CrI: [-9, 9.6]	4.3% CrI: [-6, 15.1]	-5.1% CrI: [-15.3, 5.2]
RC	-3.6% CrI: [-10.4, 3]	6.8% CrI: [-4.1, 17.4]	-3% CrI: [-13.3, 7.2]
pronoun	-1.3% CrI: [-12.5, 9.9]	2.2% CrI: [-6.5, 10.7]	-8% CrI: [-21.3, 5.7]
PRO	-3.7% CrI: [-17.9, 10.8]	4.3% CrI: [-7.2, 15.7]	6% CrI: [-9.8, 21.6]
Region 3			
declaratives	1.4% CrI: [-4, 6.9]	-1.4% CrI: [-4.7, 2.1]	6% CrI: [0.4, 11.6]
RC	-1.8% CrI: [-7.5, 4]	-4% CrI: [-11.4, 3.2]	3.2% CrI: [-5, 11.5]
pronoun	-4.7% CrI: [-19.7, 9.9]	3.7% CrI: [-4.9, 12.4]	3% CrI: [-9.4, 15.7]
PRO	0.6% CrI: [-8.8, 10.2]	6.9% CrI: [-3.8, 17.4]	-2.7% CrI: [-16.2, 10.6]

Note. Region 2 = second half of the sentence after the critical word, Region 3 = silence region, RC = relative clauses, pronoun = control structures with a pronoun, PRO = control structures with PRO.

sentences with a canonical and non-canonical word order can be regarded as agent-first processing pattern. That is, control participants expected the canonical word order, which is more frequent in German than the non-canonical order in non-experimental settings (Bader and Häussler, 2010). Control participants rapidly revised this expectation in non-canonical sentences after they encountered the disambiguating information in the input. These results are consistent with the established findings regarding processing in language-unimpaired control participants (e.g., Hanne et al., 2015; Mack et al., 2016; Meyer et al., 2012).

In contrast to the control participants, IWA displayed no differences in target fixations between sentences with a canonical and non-canonical word order, with the exception of declaratives in the retest. At first glance, the absence of differences in online processing (i.e., the overlap in target fixations between canonical and non-canonical sentences) is surprising given the fact that we observed differences in offline processing (i.e., lower response accuracy in non-canonical versus canonical sentences, Pregla et al., 2021). This contradiction between the offline and online data can be explained by the fact that only the fixations of the correct trials entered the analyses. The result can therefore be interpreted as follows: Non-canonical sentences induced a higher number of incorrect responses as compared to canonical sentences. However, if a correct response was given, processing (as indicated by fixation patterns) was similar for both non-canonical and canonical sentences. Thus, the overlapping fixation paths in correct trials suggest that IWA were able to process the sentences correctly regardless of complexity. In principle, this conclusion is consistent with the RRH, according to which both canonical and non-canonical sentences are processed normal-like provided the randomly fluctuating resources of the IWA are high enough. However, the conclusion that processing in IWA is normal-like may be premature as the comparison of the participant groups in the next section shows.

The results of the current study and previous studies suggest an agent-first fixation pattern for control participants (Hanne et al., 2015; Mack and Thompson, 2017). This pattern consists of increasing fixations to the distractor picture in non-canonical trials followed by increasing fixations to the target picture reflecting a revision of the prediction. In contrast, the results of the current study and previous studies suggest no agent-first fixation pattern for IWA (Hanne et al., 2015; Mack et al., 2016; Meyer et al., 2012). How can the absence of the agent-first fixation pattern in IWA be explained? The RRH assumes that IWA have a reduced and fluctuating resource, and this reduced resource likely manifests itself in a reduction of processing speed. Importantly, a slowdown in processing speed should entail a slow emergence of agent-first predictions. Moreover, resource fluctuation should lead to variation with respect to the emergence of agent first predictions during sentence processing. Due to this fluctuation, we assume that IWA may or may not create an agent-first prediction before the unambiguous case cue occurs in the input. If they do not create an agent-first prediction before the unambiguous case cue occurs, there is no mismatch between the agent-

first prediction and the information that is provided by the cue. As a result a correct response is given. This processing pattern matches with the fixation paths in correct trials: Due to the absence of an agent-first prediction, no revision of the prediction is needed in non-canonical sentences. Therefore, fixation paths overlap in canonical and non-canonical sentences. In contrast, if IWA do engage in an agent-first interpretation before the cue information is given, a mismatch arises between this prediction and the cue. We assume that this conflict cannot be solved because IWA are unable to revise a previously made prediction, thus resulting in an incorrect response. This could explain IWA's high number of incorrect responses in non-canonical trials. This interpretation is supported by the fixation patterns in incorrect non-canonical trials: Due to the agent-first prediction, IWA show increasing distractor fixations, and as they are not able to revise their prediction, these fixation patterns do not change, i.e. IWA continue to fixate the distractor picture. The conclusion that IWA might be impaired in revising initial sentence interpretations is consistent with the results of Lissón et al. (2021). Using computational modeling, these authors found that IWA have a much lower probability of *backtracking* (i.e., revision of an incorrect sentence interpretation to the correct one) than control participants. That is, incorrect initial sentence interpretations, e.g., agent-first predictions in non-canonical sentences, might result in incorrect responses, as incorrect interpretations cannot be revised. Put differently, the results do not suggest that IWA are in general unable to make agent-first predictions, but that IWA have difficulties *revising* their agent-first predictions based on the morpho-syntactic information of the input. Overall, the fixation patterns of IWA for non-canonical sentences hint at a processing pattern that is not only slower but also different from normal-like processing in that the revision of agent-first predictions is impaired.

With respect to individual participants, the RRH predicted an interaction between sentence complexity and resource capacity. Applied to fixation data, it was assumed that the differences in target fixations between simple and complex sentences should be larger in IWA with lower resource capacity. Resource capacity was operationalized as the overall response accuracy (i.e., low accuracy = low capacity and high accuracy = high capacity). As shown in Fig. 6, the patterns were not consistent with the predicted interaction. Possibly, the interaction could not be detected since the IWA as a group also did not show clear differences between complex and simple sentences, as discussed in the previous paragraphs. Furthermore, the number of 20 observations per sentence type might have been too low to find differences between individuals.

In sum, the data discussed in this section are consistent with the RRH's prediction that processing difficulty is higher in complex versus simple sentences. More specifically, although the number of target fixations in IWA was not clearly different in simple and complex sentences, the number of incorrect trials was higher in complex versus simple sentences. One unexpected finding is that IWA did not show an agent-

Table A.3
Response accuracy in visual world studies on sentence comprehension in aphasia.

Study and sentence type	Mean accuracy	Reported uncertainty measure	Reported uncertainty values	N IWA	N items	Calculated standard error
Adelt et al. (2017) ORC, case marked, full DP	0.53	sd	0.501	10	16	0.16
Adelt et al. (2017) ORC, case marked, pronoun	0.55	sd	0.499	10	16	0.16
Adelt et al. (2017) ORC, number marked, full DP	0.59	sd	0.493	10	16	0.16
Adelt et al. (2017) ORC, number marked, pronoun	0.71	sd	0.454	10	16	0.14
Adelt et al. (2017) SRC, case marked, full DP	0.76	sd	0.431	10	16	0.14
Adelt et al. (2017) SRC, case marked, pronoun	0.69	sd	0.465	10	16	0.15
Adelt et al. (2017) SRC, number marked, full DP	0.61	sd	0.49	10	16	0.16
Adelt et al. (2017) SRC, number marked, pronoun	0.51	sd	0.501	10	16	0.16
Bos et al. (2014) Future Tense	0.64	sd	0.26	6	20	0.11
Bos et al. (2014) Past Tense	0.92	sd	0.04	6	20	0.02
Choy & Thompson (2005) Pronouns	0.65	NA	NA	8	20	NA
Choy & Thompson (2005) Reflexives	0.64	NA	NA	8	20	NA
Choy & Thompson (2010) Pronouns	0.65	sd	0.141	8	20	0.05
Choy & Thompson (2010) Reflexives	0.64	sd	0.192	8	20	0.07
Dickey & Thompson (2009) ORC	0.36	range	0.16 to 0.75	8	12	NA
Dickey & Thompson (2009) Passives	0.20	range	0 to 0.67	8	12	NA
Dickey et al. (2007) obj clefts	0.67	range	0.2 to 1	12	10	NA
Dickey et al. (2007) wh-questions	0.70	range	0 to 1	12	10	NA
Dickey et al. (2007) yes-no-questions	0.87	range	0.6 to 1	12	10	NA
Engel et al. (2018) Pronouns, long	0.63	sd	0.2066	6	10	0.08
Engel et al. (2018) Pronouns, short	0.70	sd	0.2683	6	10	0.11
Engel et al. (2018) Reflexives, long	0.83	sd	0.1211	6	10	0.05
Engel et al. (2018) Reflexives, short	0.90	sd	0.0894	6	10	0.04
Hanne et al. (2011) OVS	0.46	se	0.0314	7	20	0.03
Hanne et al. (2011) SVO	0.80	se	0.0426	7	20	0.04
Hanne et al. (2015) OVS	0.46	sd	0.166	8	20	0.06
Hanne et al. (2015) SVO	0.77	sd	0.116	8	20	0.04
Hanne et al. (2016) object-questions	0.66	sd	0.186	6	20	0.08
Hanne et al. (2016) subject-questions	0.64	sd	0.068	6	20	0.03
Mack & Thompson (2017) actives, post-treatment	0.74	sd	0.159	10	24	0.05
Mack & Thompson (2017) actives, pre-treatment	0.77	sd	0.123	10	24	0.04
Mack & Thompson (2017) passives, post-treatment	0.64	sd	0.169	10	24	0.05
Mack & Thompson (2017) passives, pre-treatment	0.53	sd	0.166	10	24	0.05
Mack et al. (2016) actives, session 1	0.74	sd	0.189	12	24	0.06
Mack et al. (2016) actives, session 2	0.75	sd	0.167	12	24	0.05
Mack et al. (2016) passives, session 1	0.55	sd	0.18	12	24	0.05
Mack et al. (2016) passives, session 2	0.61	sd	0.192	12	24	0.06
Meyer et al. (2012) actives	0.76	sd	0.13	10	20	0.04
Meyer et al. (2012) passives	0.54	sd	0.14	10	20	0.04
Pregla et al. o-ctrl retest	0.76	sd	0.42	21	10	0.09
Pregla et al. o-ctrl test	0.68	sd	0.46	21	10	0.10
Pregla et al. ORC retest	0.45	sd	0.5	21	10	0.11
Pregla et al. ORC test	0.40	sd	0.5	21	10	0.11
Pregla et al. OS retest	0.44	sd	0.5	21	10	0.11
Pregla et al. OS test	0.42	sd	0.5	21	10	0.11
Pregla et al. Pronouns match retest	0.70	sd	0.44	21	10	0.10
Pregla et al. Pronouns match test	0.50	sd	0.5	21	10	0.11
Pregla et al. Pronouns mismatch retest	0.77	sd	0.39	21	10	0.08
Pregla et al. Pronouns mismatch test	0.60	sd	0.47	21	10	0.10
Pregla et al. s-ctrl retest	0.63	sd	0.48	21	10	0.10
Pregla et al. s-ctrl test	0.50	sd	0.5	21	10	0.11
Pregla et al. SO retest	0.78	sd	0.4	21	10	0.09
Pregla et al. SO test	0.64	sd	0.47	21	10	0.10
Pregla et al. SRC retest	0.67	sd	0.46	21	10	0.10
Pregla et al. SRC test	0.67	sd	0.46	21	10	0.10
Schumacher et al. (2015) OVS	0.36	se	0.08	12	16	0.08
Schumacher et al. (2015) passives	0.75	se	0.073	12	16	0.07
Schumacher et al. (2015) SVO	0.93	se	0.022	12	16	0.02
Sheppard et al. (2015) object-which-questions	0.43	se	0.04	8	46	0.04
Sheppard et al. (2015) object-who-questions	0.53	se	0.16	8	46	0.16
Sheppard et al. (2015) subject-which-questions	0.54	se	0.12	8	46	0.12
Sheppard et al. (2015) subject-who-questions	0.56	se	0.14	8	46	0.14
Thompson et al. (2004) obj clefts	0.52	NA	NA	4	10	NA
Thompson et al. (2004) wh-questions	0.92	NA	NA	4	10	NA
Thompson et al. (2004) yes-no-questions	0.85	NA	NA	4	10	NA

Note. IWA = individuals with aphasia, SRC/ORC = subject/ object relative clause, DP = determiner phrase, obj = object, SVO/SO = declarative sentence with subject (verb) object order, OVS/OS = declarative sentence with object (verb) subject order, o-ctrl/ s-ctrl = object/ subject control, se = standard error, sd = standard deviation, NA = not reported.

first fixation pattern. The absence of this pattern might indicate that agent-first predictions emerge slow in IWA and that IWA have difficulties successfully revising agent-first predictions once they emerged.

4.4. Processing variability between test phases

The RRH predicts variability in the performance of IWA caused by random fluctuations in resources. To test this prediction, the target fixations of the test phase and the retest phase were compared.

The control participants did not exhibit notable changes (i.e., increases or decreases) in target fixations in the retest phase. This result is inconsistent with Mack et al. (2016), where control participants showed systematic increases in target fixations in the retest. A reason for the diverging results could be that the interval between the test phase and the retest phase was different between this study and Mack et al. (2016). In this study, the gap between test phases was two months, whereas, in Mack et al. (2016), it was only one week. The short gap in Mack et al. (2016) could have enabled participants to remember the task better than in this study, which could explain the differences in practice effects between the two studies.

With respect to IWA, Mack et al. (2016) observed a systematic increase in target fixations between test and retest which they interpreted as a practice effect. In our study, we did not observe such a systematic increase in target fixations in the retest. Furthermore, clear changes in fixation paths between the test phases occurred only in one sentence structure, namely the OS declaratives. In OS declaratives, target fixations increased more slowly in the retest than in the test phase. This result is unexpected when assuming a practice effect, because a practice effect should have led to a faster increase in target fixations. A similar slowdown in sentence processing has been observed by Warren et al. (2016). In their study, the IWA became slower in reading low predictable sentences over the course of the experiment, while the control group became faster. The authors speculated that IWA do not adjust to experimental sentences in the same way as control participants (Warren et al., 2016). The slower increase in the retest in our study might therefore suggest that IWA had difficulties adapting to sentences with a non-canonical word order. However, the results of both studies are not fully comparable with each other since Warren et al. (2016) studied changes in the behaviour within a single session and not between different sessions. Furthermore, the slowdown in target fixations is only present in the OS declaratives and not in the other sentence structures. Therefore, the difference in target fixations in OS declaratives could be an accidental finding.

In sum, the RRH predicted variability in the performance because of random fluctuations in processing resources and the pattern of target fixations in test and retest is overall in line with this prediction.

5. Summary and conclusion

Table 6 provides an overview of the predictions of the RRH, the expected fixation patterns, and our results.

Four findings were consistent with the RRH. First, there were stable increases over 50% target fixations in correct trials, and early and stable divergences between correct and incorrect trials. These fixation patterns occurred in simple and complex sentences, across all sentence structures and test phases, both at the group level and at the individual participant level. The latter results indicate that IWA do not choose a picture at random but settle on a picture in correct trials in the sentence-picture matching task. This finding is consistent with the prediction of the RRH that the processing of IWA in correct trials is normal-like. Second, IWA showed a slower-than-normal increase in target fixations. This slowdown is compatible with the RRH because processing speed might be the resource that is reduced in IWA. Third, while the expected divergence in target fixation between simple and complex sentences was not confirmed, the number of incorrect trials was higher for complex sentences. Taking response accuracy into account, this finding is in line

with the RRH according to which processing should be successful irrespective of sentence complexity once the resource demands are met. Fourth, IWA did not show systematic increases in target fixations in the retest. This finding is consistent with the prediction of the RRH that sentence processing should be variable in IWA.

Three findings diverged from the predictions of the RRH. First, the magnitude of target fixations was lower in IWA than in the control participants in correct trials, which could reflect a reduced certainty in picture selection. Second, IWA did not exhibit an agent-first fixation pattern, which points towards an impairment in the revision of structural predictions. Third, IWA showed increased canonicity effects in the retest phase, which might indicate that IWA have difficulties adjusting to the input. Under the RRH, these findings for correct trials are unexpected given that processing in correct trials should be normal-like. Caplan et al. (2015) also found differences in processing between IWA and control participants for correct trials. They concluded that the impairment has graded effects, "...at times slowing incremental processing without leading to an error" (Caplan et al., 2015, p. 305). While our results also indicated a slowdown, it is questionable whether slowed processing alone can explain the difficulties in correct trials. Possibly, an additional source might cause these difficulties. For example, IWA might struggle matching their expectations about the sentence structure with the actual linguistic input, which requires correct perception of the input, detection of the mismatch between input and expectation, and updating the expectations (Cope et al., 2017). Difficulty in matching expectation and input might cause IWA to be less certain in picture selection than control participants and impaired in revising incorrect expectations. The impairment in revising expectations might eventually lead to difficulties adjusting to complex non-canonical sentences. Overall, it seems that processing difficulties may not only underlie incorrect trials but also correct trials. Thus, our results confirm the RRH in part, but not completely, since in some respects processing of the IWA in correct trials is not normal-like.

To conclude, our findings were consistent with the predictions of the RRH that processing difficulty is more frequent in complex versus simple trials, and that processing varies unsystematically between test phases. Also the observed processing slowdown in IWA is compatible with the RRH. However, our results were mixed with respect to the prediction that processing in correct trials is normal-like. On the one hand, IWA showed a deliberate decision for the target picture in correct trials in all sentence structures and both test phases, which speaks for normal-like processing. On the other hand, IWA showed a reduced certainty in picture selection, difficulty in revising sentence interpretations, and difficulty in adjusting to complex sentences, which speaks for processing difficulties in correct trials. Further research is needed to investigate whether these performance patterns can be attributed to the effects of slowed processing.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank the individuals with aphasia and the language-unimpaired volunteers who participated in this study. We are grateful to Andreas Schmidt for helpful discussions of the work. We would also like to thank Silke Böttger, Sarah Düring and Therese Mayr for assisting with data collection. This research was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Project number 317633480 – SFB 1287, project B02 (PIs: Shravan Vasishth, Frank Burchert, and Nicole Stadie).

Appendix A

A.1. Sentence stimuli

Declarative sentences

The manipulated determiners (*nominative/ accusative*) are presented in italics.

- Hier badet *der/ den* Esel gerade *den/ der* Tiger.
Here *the (nom/acc)* donkey just bathes *the (acc/nom)* tiger.
- Hier zeichnet *der/ den* Büffel gerade *den/ der* Panther.
Here *the (nom/acc)* buffalo just draws *the (acc/nom)* panther.
- Hier kitzelt *der/ den* Hamster gerade *den/ der* Igel.
Here *the (nom/acc)* hamster just tickles *the (acc/nom)* hedgehog.
- Hier rettet *der/ den* Pudel gerade *den/ der* Kater.
Here *the (nom/acc)* poodle just rescues *the (acc/nom)* tomcat.
- Hier bürstet *der/ den* Kater gerade *den/ der* Pudel.
Here *the (nom/acc)* tomcat just brushes *the (acc/nom)* poodle.
- Hier tröstet *der/ den* Tiger gerade *den/ der* Esel.
Here *the (nom/acc)* donkey just comforts *the (acc/nom)* tiger.
- Hier leitet *der/ den* Panther gerade *den/ der* Büffel.
Here *the (nom/acc)* panther just guides *the (acc/nom)* buffalo.
- Hier füttert *der/ den* Igel gerade *den/ der* Hamster.
Here *the (nom/acc)* hedgehog just feeds *the (acc/nom)* hamster.
- Hier findet *der/ den* Eber gerade *den/ der* Otter.
Here *the (nom/acc)* boar just finds *the (acc/nom)* otter.
- Hier streichelt *der/ den* Otter gerade *den/ der* Eber.
Here *the (nom/acc)* otter just pets *the (acc/nom)* boar.

Relative clauses

The manipulated determiners to get subject and object relative clauses (*nominative/ accusative*) are presented in italics.

- Hier ist der Esel, *der/ den den/ der* Tiger gerade badet.
Here is the donkey *who (nom/ acc) the (nom/ acc)* tiger just bathes.
- Hier ist der Büffel, *der/ den den/ der* Panther gerade zeichnet.
Here is the buffalo *who (nom/ acc) the (nom/ acc)* panther just draws.
- Hier ist der Hamster, *der/ den den/ der* Igel gerade kitzelt.
Here is the hamster *who (nom/ acc) the (nom/ acc)* hedgehog just tickles.
- Hier ist der Pudel, *der/ den den/ der* Kater gerade rettet.
Here is the poodle *who (nom/ acc) the (nom/ acc)* tomcat just rescues.
- Hier ist der Kater, *der/ den den/ der* Pudel gerade bürstet.
Here is the tomcat *who (nom/ acc) the (nom/ acc)* poodle just brushes.
- Hier ist der Tiger, *der/ den den/ der* Esel gerade tröstet.
Here is the tiger *who (nom/ acc) the (nom/ acc)* donkey just comforts.
- Hier ist der Panther, *der/ den den/ der* Büffel leitet.
Here is the panther *who (nom/ acc) the (nom/ acc)* buffalo just guides.
- Hier ist der Igel, *der/ den den/ der* Hamster gerade füttert.
Here is the hedgehog *who (nom/ acc) the (nom/ acc)* hamster just feeds.
- Hier ist der Eber, *der/ den den/ der* Otter gerade findet.
Here is the boar *who (nom/ acc) the (nom/ acc)* otter just finds.
- Hier ist der Otter, *der/ den den/ der* Eber gerade streichelt.
Here is the otter *who (nom/ acc) the (nom/ acc)* boar just pets.

Sentences with PRO

The manipulated verb (*subject control/ object control*) is presented in italics.

- Peter *verspricht/ erlaubt* nun Lisa, das kleine Lamm zu streicheln und zu kraulen.
Peter now *promises/ allows* Lisa to pet and to ruffle the little lamb.
- Thomas *versichert/ gestattet* nun Anna, das dicke Rind zu melken und zu hüten.
Thomas now *assures/ allows* Anna to milk and to tend the thick cattle.
- Thomas *droht/ befiehlt* nun Lisa, das schnelle Huhn zu jagen und zu fangen.
Thomas now *threatens/ commands* Lisa to chase and to catch the fast chicken.
- Peter *garantiert/ empfiehlt* nun Anna, das stolze Ross zu bürsten und zu striegeln.
Peter *guarantees/ recommends* now Anna to brush and to comb the proud steed.
- Thomas *schwört/ rät* nun Anna, das süße Ferkel zu waschen und zu säubern.
Thomas now *swears/ advises* Anna to wash and to clean the sweet piglet.
- Lisa *verspricht/ erlaubt* nun Peter, das alte Schaf zu impfen und zu pflegen.
Lisa now *promises/ allows* Peter to vaccinate and to nurse the old sheep.
- Anna *versichert/ gestattet* nun Thomas, das junge Kalb zu malen und zu zeichnen.
Anna now *assures/ allows* Thomas to paint and to draw the young calf.
- Anna *droht/ befiehlt* nun Peter, das kluge Schwein zu füttern und zu mästen.
Anna now *threatens/ commands* Peter to feed and to fatten the clever pig.
- Lisa *garantiert/ empfiehlt* nun Thomas, das scheue Reh zu locken und zu suchen.
Lisa now *guarantees/ recommends* Thomas to lure and to search the shy deer.
- Lisa *schwört/ rät* nun Peter, das schöne Pferd zu satteln und zu zäumen.
Lisa now *swears/ advises* Peter to saddle and to bridle the nice horse.

Sentences with a pronoun

The manipulated noun (*same gender/ different gender*) is presented in italics.

- Peter verspricht nun *Thomas/ Lisa*, dass er das kleine Lamm streichelt und krault.
Peter now promises *Thomas/ Lisa* that he will pet and ruffle the little lamb.
- Thomas versichert nun *Peter/ Anna*, dass er das dicke Rind melkt und hütet.
Thomas now assures *Peter/ Anna* that he will milk and tend the thick cattle.
- Thomas droht nun *Peter/ Lisa*, dass er das schnelle Huhn jagt und fängt
Thomas now threatens *Peter/ Lisa* that he will chase and catch the fast chicken.
- Peter garantiert nun *Thomas/ Anna*, dass er das stolze Ross bürstet und striegelt.
Peter guarantees now *Thomas/ Anna* that he will brush and comb the proud steed.

5. Thomas schwört nun *Peter/ Anna*, dass er das süße Ferkel wäscht und säubert.
Thomas now swears *Peter/ Anna* that he will wash and clean the sweet piglet.
6. Lisa verspricht nun *Anna/ Peter*, dass sie das alte Schaf impft und pflegt.
Lisa now promises *Anna/ Peter* that she will vaccinate and nurse the old sheep.
7. Anna versichert nun *Lisa/ Thomas*, dass sie das junge Kalb malt und zeichnet.
Anna now assures *Lisa/ Thomas* that she will paint and draw the young calf.
8. Anna droht nun *Lisa/ Peter*, dass sie das kluge Schwein füttert und mästet.
Anna now threatens *Lisa/ Peter* that she will feed and fatten the clever pig.
9. Lisa garantiert nun *Anna/ Thomas*, dass sie das scheue Reh lockt und sucht.
Lisa now guarantees *Anna/ Thomas* that she will lure and search the shy deer.
10. Lisa schwört nun *Anna/ Peter*, dass sie das schöne Pferd sattelt und zäumt.
Lisa now swears *Anna/ Peter* that she will saddle and bridle the nice horse.

A.2. Estimates of the bootstrapped divergence onsets

Table A.1

A.3. Comparison of the fixation paths of 50 and 22 control participants

Fig. A.1

A.4. Estimates of the time window analysis comparing target fixations and aphasia severity

Table A.2

A.5. Comparison of the response accuracy in this study and previous visual world studies on sentence comprehension in aphasia

Table A.3

References

- Adelt, A., Stadie, N., Lassotta, R., Adani, F., & Burchert, F. (2017). Feature dissimilarities in the processing of German relative clauses in aphasia. *Journal of Neurolinguistics*, *44*, 17–37.
- Arantzeta, M., Bastiaanse, R., Burchert, F., Wieling, M., Martinez-Zabaleta, M., & Laka, I. (2017). Eye-tracking the effect of word order in sentence comprehension in aphasia: Evidence from Basque, a free word order ergative language. *Language, Cognition and Neuroscience*, *32*, 1320–1343.
- Bader, M., & Häussler, J. (2010). Word order in German: A corpus study. *Lingua*, *120*, 717–762.
- Bader, M., & Meng, M. (1999). Subject-object ambiguities in German embedded clauses: An across-the-board comparison. *Journal of Psycholinguistic Research*, *28*, 121–143.
- Barbieri, E., Mack, J. E., Chiappetta, B., Europa, E., & Thompson, C. K. (2019). Recovery of offline and online sentence processing in aphasia: Language and domain-general network neuroplasticity. *Cortex*, *120*, 394–418.
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.
- Boersma, P., & Weenink, D. (2018). Praat: Doing phonetics by computer [computer program]. version 6.0.37. Retrieved February 13, 2018, from <http://www.praat.org/>
- Bos, L. S., Hanne, S., Wartenburger, I., & Bastiaanse, R. (2014). Losing track of time? Processing of time reference inflection in agrammatic and healthy speakers of German. *Neuropsychologia*, *65*, 180–190.
- Burchert, F., Hanne, S., & Vasishth, S. (2013). Sentence comprehension disorders in aphasia: The concept of chance performance revisited. *Aphasiology*, *27*, 112–125.
- Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*, 1–28.
- Bürkner, P.-C. (2018). Advanced Bayesian multilevel modeling with the R package brms. *The R Journal*, *10*, 395–411.
- Caplan, D. (2012). Resource reduction accounts of syntactically based comprehension disorders. In R. Bastiaanse, & C. K. Thompson (Eds.), *Perspectives on agrammatism* (pp. 48–62). Hove: Psychology Press.
- Caplan, D., DeDe, G., & Michaud, J. (2006). Task-independent and task-specific syntactic deficits in aphasic comprehension. *Aphasiology*, *20*, 893–920.
- Caplan, D., & Hildebrandt, N. (1988). *Disorders of syntactic comprehension*. Cambridge, MA: MIT Press.
- 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*, 21–33.
- Caplan, D., Michaud, J., & Hufford, R. (2015). Mechanisms underlying syntactic comprehension deficits in vascular aphasia: New evidence from self-paced listening. *Cognitive Neuropsychology*, *32*, 283–313.
- 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*, 103–150.
- Choy, J. J., & Thompson, C. K. (2005). Online comprehension of anaphor and pronoun constructions in Broca's aphasia: Evidence from eyetracking. *Brain and Language*, *95*, 119–120.
- Choy, J. J., & Thompson, C. K. (2010). Binding in agrammatic aphasia: Processing to comprehension. *Aphasiology*, *24*, 551–579.
- Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, *6*, 84–107.
- Cope, T. E., Sohoglu, E., Sedley, W., Patterson, K., Jones, P. S., Wiggins, J., & Rowe, J. B. (2017). Evidence for causal top-down frontal contributions to predictive processes in speech perception. *Nature Communications*, *8*, 1–16.
- 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*, 1–22.
- Dickey, M. W., & Thompson, C. K. (2009). Automatic processing of wh-and NP-movement in agrammatic aphasia: Evidence from eyetracking. *Journal of Neurolinguistics*, *22*, 563–583.
- 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*, 85–103.
- Drai, D., & Grodzinsky, Y. (1999). Comprehension regularity in Broca's aphasia? There's more of it than you ever imagined. *Brain and Language*, *70*, 139–143.
- Engel, S., Shapiro, L. P., & Love, T. (2018). Proform-antecedent linking in individuals with agrammatic aphasia: A test of the intervener hypothesis. *Journal of Neurolinguistics*, *45*, 79–94.
- Hanne, S., Burchert, F., De Bleser, R., & Vasishth, S. (2012). Offline chance performance in sentence comprehension: What single-case analyses reveal about online sentence processing in aphasia. Poster presented at the Thirtieth European Workshop on Cognitive Neuropsychology. Bressanone, Italy.
- Hanne, S., Burchert, F., De Bleser, R., & Vasishth, S. (2015). Sentence comprehension and morphological cues in aphasia: What eye-tracking reveals about integration and prediction. *Journal of Neurolinguistics*, *34*, 83–111.
- Hanne, S., Burchert, F., & Vasishth, S. (2016). On the nature of the subject-object asymmetry in wh-question comprehension in aphasia: Evidence from eye tracking. *Aphasiology*, *30*, 435–462.
- 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*, 221–244.
- Heister, J., Würzner, K.-M., Bubbenzer, J., Pohl, E., Hanneforth, T., Geyken, A., & Kliegl, R. (2011). DlexDB—eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychologische Rundschau*, *62*, 10–20.
- Huang, Y., & Snedeker, J. (2020). Evidence from the visual world paradigm raises questions about unaccusativity and growth curve analyses. *Cognition*, *200*, 104251.
- Huber, W., Poeck, K., Weniger, D., & Willmes, K. (1983). *AAT-Aachener Aphasie Test*. Göttingen: Hogrefe.
- Just, M. A., & Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, *99*, 122.
- Kwon, N., & Sturt, P. (2016). Processing control information in a nominal control construction: An eye-tracking study. *Journal of Psycholinguistic Research*, *45*, 779–793.
- Levitt, W. J. (2001). Spoken word production: A theory of lexical access. *Proceedings of the National Academy of Sciences*, *98*, 13464–13471.
- Lewandowski, D., Kurowiczka, D., & Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, *100*, 1989–2001.
- Lissón, P., Pregla, D., Nicenboim, B., Paape, D., van het Nederend, M. L., Burchert, F., & Vasishth, S. (2021). A computational evaluation of two models of retrieval processes in sentence processing in aphasia. *Cognitive Science*, *45*, e12956.
- Luzzatti, C., Toraldo, A., Guasti, M. T., Ghirardi, G., Lorenzi, L., & Guarnaschelli, C. (2001). Comprehension of reversible active and passive sentences in agrammatism. *Aphasiology*, *15*, 419–441.
- Mack, J. E., & Thompson, C. K. (2017). Recovery of online sentence processing in aphasia: Eye movement changes resulting from treatment of underlying forms. *Journal of Speech, Language, and Hearing Research*, *60*, 1299–1315.
- Mack, J. E., Wei, A. Z.-S., Gutierrez, S., & Thompson, C. K. (2016). Tracking sentence comprehension: Test-retest reliability in people with aphasia and unimpaired adults. *Journal of Neurolinguistics*, *40*, 98–111.

- McMurray, B. (2020). I'm not sure that curve means what you think it means: Toward a [more] realistic understanding of the role of eye-movement generation in the visual world paradigm. Retrieved October 19, 2020, from <https://psyarxiv.com/pb2c6/>.
- Meyer, A. M., Mack, J. E., & Thompson, C. K. (2012). Tracking passive sentence comprehension in agrammatic aphasia. *Journal of Neurolinguistics*, 25, 31–43.
- Mirman, D. (2014). *Growth curve analysis and visualization using R* (1st ed.). Boca Raton: CRC Press.
- 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, 671–717.
- Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, 53, 695–699.
- Nozari, N., Mirman, D., & Thompson-Schill, S. L. (2016). The ventrolateral prefrontal cortex facilitates processing of sentential context to locate referents. *Brain and Language*, 157, 1–13.
- Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9, 97–113.
- Pregla, D., Lissón, P., Vasishth, S., Burchert, F., & Stadie, N. (2021). Variability in sentence comprehension in aphasia in German. *Brain and Language*, 222, 105008.
- R Core Team. (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. Retrieved from <https://www.R-project.org/>.
- Schad, D. J., Vasishth, S., Hohenstein, S., & Kliegl, R. (2020). How to capitalize on a priori contrasts in linear (mixed) models: A tutorial. *Journal of Memory and Language*, 110, 104038.
- Schroeder, S. (2007). *Interaktion gedächtnis- und erklärungs-basierter Verarbeitungsprozesse bei der pronominalen Auflösung. Analyse der Effekte von Impliziten Kausalitäts- und Gender-Informationen durch die Model lierung von Reaktionszeitverteilungen*. Universität zu Köln. Doctoral dissertation.
- Schumacher, R., Cazzoli, D., Eggenberger, N., Preisig, B., Nef, T., Nyffeler, T., & Müri, R. M. (2015). Cue recognition and integration – Eye tracking evidence of processing differences in sentence comprehension in aphasia. *PLoS ONE*, 10, e0142853.
- Sharma, S., Kim, H., Harris, H., Haberstroh, A., Wright, H. H., & Rothermich, K. (2021). Eye tracking measures for studying language comprehension deficits in aphasia: A systematic search and scoping review. *Journal of Speech, Language, and Hearing Research*, 64, 1008–1022.
- Sheppard, S. M., Walenski, M., Love, T., & Shapiro, L. P. (2015). The auditory comprehension of wh-questions in aphasia: Support for the intervener hypothesis. *Journal of Speech, Language, and Hearing Research*, 58, 781–797.
- Stadie, N., Cholewa, J., & De Bleser, R. (2013). *LEMO 2.0: Lexikon model lorientiert: Diagnostik für Aphasie, Dyslexie und Dysgraphie*. Hofheim: NAT-Verlag.
- Stewart, A. J., Pickering, M. J., & Sanford, A. J. (2000). The time course of the influence of implicit causality information: Focusing versus integration accounts. *Journal of Memory and Language*, 42, 423–443.
- Stiebels, B. (2007). Towards a typology of complement control. *ZAS Papers in Linguistics*, 47, 1–58.
- Stone, K., Lago, S., & Schad, D. (2020). Divergence point analyses of visual world data: Applications to bilingual research. *Bilingualism: Language and Cognition*, 1–9.
- Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, 48, 542–562.
- Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995). Integration of visual and linguistic information in spoken language comprehension. *Science*, 268, 1632–1634.
- 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, 124–125.
- Vogelzang, M., Thiel, C.M., Rosemann, S., Rieger, J., & Ruigendijk, E. (2019). Cognitive abilities to explain individual variation in the interpretation of complex sentences by older adults. In Proceedings of the 41th Annual Conference of the Cognitive Science Society (pp. 3036–3042). Cognitive Science Society.
- Warren, T., Dickey, M. W., & Lei, C.-M. (2016). Structural prediction in aphasia: Evidence from either. *Journal of Neurolinguistics*, 39, 38–48.
- Wendt, D., Brand, T., & Kollmeier, B. (2014). An eye-tracking paradigm for analyzing the processing time of sentences with different linguistic complexities. *PLoS ONE*, 9, e100186.