Sentence comprehension and morphological cues in aphasia: What eye-tracking reveals about integration and prediction

Sandra Hanne

Department of Linguistics, University of Potsdam, Germany

Frank Burchert

Department of Linguistics, University of Potsdam, Germany

Ria De Bleser

Department of Linguistics, University of Potsdam, Germany

Shravan Vasishth

Department of Linguistics, University of Potsdam, Germany

Author Note

We would like to thank the individuals with aphasia who participated in this study. Without them and their families this work would not have been possible. We are grateful to David Caplan, Cynthia Thompson and the participants of the Academy of Aphasia 2012 meeting for helpful discussions of the work reported here. We thank the two anonymous reviewers for their constructive recommendations on an earlier version of this manuscript. We would also like to thank Tom Fritzsche, Felix Engelmann, Anne Adelt and Elisa Rath for assisting with data collection and programming, and Ruben Van de Vijver for advice on prosodic measures. We thank the Zentrum für Allgemeine Sprachwissenschaft (ZAS) in Berlin for providing space for participant testing. This research was supported by the Deutsche Forschungsgemeinschaft, grant VA-482 4-1.

Abstract

Comprehension of non-canonical sentences can be difficult for individuals with aphasia (IWA). It is still unclear to which extent morphological cues like case marking or verb inflection may influence IWA's performance or even help to override deficits in sentence comprehension. Until now, studies have mainly used offline methods to draw inferences about syntactic deficits and, so far, only a few studies have looked at online syntactic processing in aphasia. We investigated sentence processing in German IWA by combining an offline (sentence-picture matching) and an online (eye-tracking in the visual-world paradigm) method. Our goal was to determine whether IWA are capable of using inflectional morphology (number-agreement markers on verbs and case markers in noun phrases) as a cue to sentence interpretation. We report results of two visual-world experiments using German reversible SVO and OVS sentences. In each study, there were eight IWA and 20 age-matched controls. Experiment 1 targeted the role of unambiguous case morphology, while Experiment 2 looked at processing of number-agreement cues at the verb in case-ambiguous sentences. IWA showed deficits in using both types of morphological markers as a cue to non-canonical sentence interpretation and the results indicate that in aphasia, processing of case-marking cues is more vulnerable as compared to verb-agreement morphology. We ascribe this finding to the higher cue reliability of agreement cues, which renders them more resistant against impairments in aphasia. However, the online data revealed that IWA are in principle capable of successfully computing morphological cues, but the integration of morphological information is delayed as compared to age-matched controls. Furthermore, we found striking differences between controls and IWA regarding subject-before-object parsing predictions. While in case-unambiguous sentences IWA showed evidence for early subject-before-object parsing commitments, they exhibited no straightforward subject-first prediction in case-ambiguous sentences, although controls did so for ambiguous structures. IWA delayed their parsing decisions in case-ambiguous sentences until unambiguous morphological information, such as a subject-verb-number-agreement cue, was available. We attribute the results for IWA to deficits in predictive processes based on morpho-syntactic cues during sentence comprehension. The results indicate that IWA adopt a wait-and-see strategy and initiate prediction of upcoming syntactic structure only when unambiguous case or agreement cues are available.

Keywords: online sentence processing, online morphological processing, aphasia, sentence comprehension deficits, word order, case marking, subject-verb-agreement, prediction, eye-tracking, visual-world paradigm

Sentence comprehension and morphological cues in aphasia: What eye-tracking reveals about integration and prediction

Introduction

Many individuals with aphasia (IWA) show impairments in auditory sentence comprehension whenever reliance on syntactic structure is necessary in order to derive the correct sentence interpretation (Caramazza & Zurif, 1976; for an overview see Caplan, 2006). Often, severe comprehension deficits can be observed for semantically reversible non-canonical sentences such as object-verb-subject (OVS), passives, object clefts and object relative clauses, which are derived by movement operations, In contrast, IWA perform better with canonical structures like subject-verb-object (SVO), subject clefts and subject relative clauses (for example, Cho-Reyes & Thompson, 2012; Grodzinsky, 2000; Mitchum & Berndt, 2008). However, different syntactic structures may be affected to varying degrees across individual IWA and intra-individual patterns do not always yield significant canonicity effects (Berndt, Mitchum, & Haendiges, 1996; Caplan, Waters, DeDe, Michaud, & Reddy, 2007; Caramazza, Capitani, Rey, & Berndt, 2001; Luzzatti et al., 2011). Nevertheless, IWA's sentence comprehension abilities are significantly worse than controls' and this effect is frequently more pronounced for non-canonical structures.

Although traditionally sentence comprehension deficits have been associated with Broca's aphasia, there is overwhelming evidence that impairments in sentence comprehension, particularly, deficits in assigning thematic roles correctly, can occur across all aphasic syndromes (Caplan, Baker, & Dehaut, 1985; Caramazza & Miceli, 1991; Dronkers, Wilkins, Van Valin, Redfern, & Jaeger, 2004; Luzzatti et al., 2011). The answer to one important question remains unclear: to what extent do different morphological cues (for example, case marking or verb inflection) influence IWA's performance in non-canonical structures? Such cues might equally well hinder or help override sentence processing deficits (Burchert, De Bleser, & Sonntag, 2003).

As languages differ in the extent to which morphological cues are overtly realized and, thus, may constitute cues to sentence meaning, it is important to study the interplay of grammatical morphology and syntactic processing with reference to language-specific properties. In the case of studies involving English-speaking IWA, only limited conclusions can be drawn about the interplay of morphology and syntax from IWA's performance, because in English many morphological markers are not realized overtly. Therefore, morphological cues like case markers on nouns and person or number-agreement morphemes on verbs provide only limited information towards the meaning of a sentence. For most sentence structures, English heavily relies on a strict subject-verb-object word order principle.

In contrast, languages with rich grammatical morphology are less restricted in their word order and they provide overt morphological cues to sentence meaning. In German, for example, word order is less constrained than in English and sentences may deviate from canonical SVO order. Moreover, German has a rich case-marking system realized on nouns and determiners and, although some case syncretisms and ambiguities exist, the grammatical function and thematic role of many noun phrases (NPs) can be inferred from their case markings.[1] In addition, German verbs are inflected for person and number in agreement with the subject. Thus, it is possible to deduce theta-role assignments from the verb's inflectional cue even in sentences with case-ambiguous NPs. However, these overt morphological cues to theta-role-assignment, case and agreement, have different properties in terms of their reliability. Agreement cues on verbs are more reliable than case cues on NPs, because case markings can be ambiguous due to case syncretisms. Agreement cues, on the other hand, are not ambiguous when they mark third person singular or plural.

[1]One of the reviewers of this manuscript has pointed out that it is important to note that in some instances the mapping of theta-roles onto case-marked NPs is not necessarily isomorphic. However, in the case of the SVO and OVS sentences used in Experiment 1 of our study, the criterion of isomorphy is met.

In this study, we capitalize on these morpho-syntactic properties of the German language in order to investigate processing of case morphology and verb inflection cues in German IWA and age-matched neurologically unimpaired controls. So far, studies have mostly used offline methods to investigate sentence processing in aphasia and only a few studies have investigated this issue from the perspective of online processing (for example, Dickey, Choy, & Thompson, 2007; Meyer, Mack, & Thompson, 2012; Hanne, Sekerina, Vasishth, Burchert, & De Bleser, 2011). Therefore, this study focuses on online sentence processing in addition to offline comprehension.

Case morphology cues in sentence comprehension in aphasia

Studies looking at morpho-syntactic processing in aphasia in languages with overt morphology indicate that IWA show deficits in processing case markers as a cue to sentence comprehension (Friedmann & Shapiro, 2003; MacWhinney, Osmán-Sági, & Slobin, 1991; Kljajevic & Murasugi, 2010; Smith & Mimica, 1984; Smith & Bates, 1987; Yarbay Duman, Altinok, Özgirgin, & Bastiaanse, 2011).

Yet, some studies found preserved abilities in IWA (Lukatela, Crain, & Shankweiler, 1988; Milekić, Bosković, Crain, & Shankweiler, 1995; De Bleser, Dronsek, & Bayer, 1988). However, these studies included tasks like grammaticality judgement or constituent ordering, which do not explicitly target IWA's ability to use case markers as a cue for identifying theta-role relations in non-canonical sentences.

In contrast, across languages, IWA have been shown to be impaired in tasks targeting specifically this ability. For example, Kljajevic and Murasugi (2010) used a figurine-pointing task to study Croatian-speaking IWA's ability to rely on case morphology as a cue to thematic role identification in reversible subject- and object-extracted sentences. IWA's performance on this task was impaired, although the degree of impairment was different across aphasia syndromes. Using an enactment task, Smith and Mimica (1984) and Smith and Bates (1987) also observed impairments in Serbo-Croatian-speaking IWA's processing of case inflections. Impairments in correctly processing case markers in an enactment task have also been observed in Hungarian and Turkish-speaking IWA (MacWhinney et al., 1991). The results for Turkish-speaking IWA have been replicated in another study involving a sentence-picture matching task (Yarbay Duman et al., 2011). Sentence-picture matching has also been used by Friedmann, Reznick, Dolinski-Nuger, and Soboleva (2010) in order to investigate processing of case cues and theta-role assignment in Russian-speaking IWA. Consistent with the previous studies, the authors found no facilitatory effect of case-marking cues on IWA's sentence comprehension. Similar results were obtained for the use of case markers in Hebrew-speaking IWA (Friedmann & Shapiro, 2003).

Impaired abilities in using case marking as a cue for establishing thematic relations have also been found for German-speaking IWA (Heeschen, 1980; Burchert et al., 2003; Swoboda-Moll, Burchert, & De Bleser, 2002). In the study by Burchert et al. (2003), agrammatic IWA performed significantly worse in comprehending reversible non-canonical (OVS, object relatives) as compared to canonical sentences (SVO, subject relatives), even when unambiguous case markers provided explicit cues to theta-role assignments. Moreover, although individual patterns across IWA were heterogeneous, ranging between above chance (but still impaired) and chance performance, none of the IWA performed below chance on OVS sentences. Such a performance would be expected if case markers in object-first sentences were completely ignored. If IWA did not process any case morphology at all, they would be expected to apply a linear agent-first strategy in identifying thematic roles in OVS sentences. This way, a reversible OVS sentence would be interpreted as if it was an SVO structure, leading to an interpretation with reversed theta-roles. Thus, performance would be expected to be constantly below chance in tasks like enactment or sentence-picture matching with distractor pictures showing role reversals.

As mentioned earlier, most studies on sentence comprehension and morphological cue processing in aphasia have used offline methods and, so far, there is only one study which investigated the effect of case-marking cues on sentence comprehension in aphasia

online. Hanne et al. (2011) conducted an eye-tracking study in order to investigate theta-role assignment and sensitivity to case-marking cues in German IWA. Similar to the results of Burchert et al. (2003), IWA were impaired in using the case-marking cues to correctly interpret OVS sentences, for which performance was at chance level. The eye-tracking data, however, revealed that IWA's online sentence processing and integration of the morphological cue was not completely impaired but delayed, as correct offline responses were associated with normal-like fixation patterns. However, fixation patterns diverged from those of controls for IWA's incorrect responses.
Taken together, studies targeting the use of case markers as a cue to thematic-role assignment have found impairments in aphasic sentence comprehension irrespective of the presence or absence of morphological cues. However, although these studies indicate that, across different languages, IWA are severely impaired in using case cues, the frequent observation of chance and sometimes even above chance performance instead of below chance performance for non-canonical case-marked sentences indicates that some retained sensitivity to morphology could be present in aphasia (cf. Burchert et al., 2003; Friedmann & Shapiro, 2003). This view is also corroborated by the online results in Hanne et al. (2011), which point to a processing-based deficit in terms of a delay in integration of word order and case-marking cues rather than a complete insensitivity to morphological markers.

Subject-verb agreement cues in sentence comprehension in aphasia

When it comes to the processing of *verb inflection* as a cue to sentence comprehension, studies using tasks like grammaticality judgement or word-monitoring have reported preserved sensitivity to inflectional morphology in Italian-, English- and German-speaking IWA (Devescovi et al., 1997; Friederici, Wessels, Emmorey, & Bellugi, 1992; Wulfeck & Bates, 1991). However, as mentioned above, these tasks do not explicitly tap whether theta-roles have successfully been assigned for sentence comprehension.

Other studies have used tasks which allow us to draw more conclusions about the use of inflectional markers as a cue to assigning thematic roles (Bates, Friederici, & Wulfeck, 1987; Burchert et al., 2003; Nicol, Jakubowicz, & Goldblum, 1996; Smith & Bates, 1987). Using an enactment task, Bates et al. (1987) found that German-speaking IWA were impaired in processing agreement markers on verbs as a cue for theta-role assignment.[2] Similarly, in the study by Smith and Bates (1987), Serbo-Croatian-speaking IWA's sentence comprehension was impaired when verb agreement cues alone indicated the agent-theme relations of a sentence. Their ability to use the inflectional cue for assigning theta-roles depended on the presence of additional converging cues like case markers. In the study by Nicol et al. (1996), English- and French-speaking IWA failed to use the number-marking inflection on the verb in order to correctly perform a sentence-picture matching task.

Burchert et al. (2003) studied IWA's use of verb inflection as a cue to theta-role assignment in German case-ambiguous sentences using sentence-picture matching. Consistent with Smith and Bates (1987), they found no positive effect of number-agreement cues on IWA's comprehension performance, and, moreover, IWA appeared to be more impaired in using agreement as compared to case-marking cues (the use of which, nevertheless, was impaired, see above). However, similar to the findings for case-marking cues, analyses at an individual level revealed that IWA performed heterogeneously as two out of the seven IWA could benefit from the presence of verb agreement cues.

In sum, although some studies reported retained sensitivity to subject-verb-agreement cues in tasks like grammaticality judgement, there is no evidence for beneficial effects of verb morphology on sentence comprehension in aphasia in tasks that more explicitly tap

[2] We note that the results of Bates et al. should be considered with caution and are difficult to interpret as many ungrammatical sentences have been used in the enactment task. In addition, due to ambiguities in the German case-marking system, some sentences were globally ambiguous in their meaning and it is unclear to what extent this may have influenced the results.

theta-role assignment. Importantly, no study has, to our knowledge, investigated *online* processing of subject-verb agreement cues for sentence comprehension in aphasia.

Processing of case and verb morphology cues in unimpaired sentence comprehension

German word order is rather free and grammatical roles as well as thematic relations must often be derived from the case-marking cues contained in NPs. For example, in simple declarative sentences, the first NP may be the subject as well as the object of the sentence. In sentences containing NPs that are unambiguously marked for case, word order (SVO or OVS, respectively) is signaled by the nominative or accusative marker at NP1 (see (1)). However, due to case syncretisms, a sentence like (2) is globally ambiguous with respect to word order and the SVO as well as the OVS reading is possible. Yet, case-ambiguous sentences like (3) may be disambiguated at the verb by subject-verb-agreement cues.

(1) a. **SVO:**
Der Arzt schubst den Dieb.
the$_{NOM}$ doctor pushes the$_{ACC}$ thief
'The doctor pushes the thief.'

b. **OVS:**
Den Arzt schubst der Dieb.
the$_{ACC}$ doctor pushes the$_{NOM}$ thief
'The thief pushes the doctor.'

(2) **Ambiguous:**
Das Kind schubst die Tante.
the$_{NOM/ACC}$ child pushes the$_{NOM/ACC}$ aunt
'The child pushes the aunt. / The aunt pushes the child.'

(3) a. **Verb-agreement based disambiguation towards SVO:**

Das Kind schubst die Tanten.

the$_{NOM/ACC}$ child$_{singular}$ pushes$_{3rd-pers-singular}$ the$_{NOM/ACC}$ aunts$_{plural}$

'The child pushes the aunts.'

b. **Verb-agreement based disambiguation towards OVS:**

Das Kind schubsen die Tanten.

the$_{NOM/ACC}$ child$_{singular}$ push$_{3rd-pers-plural}$ the$_{NOM/ACC}$ aunts$_{plural}$

'The aunts push the child.'

Overall, as revealed by reading time studies, there is a strong subject-before-object preference for German sentences in which the first NP is ambiguous between nominative and accusative case (Gorrell, 2000; Hemforth & Konieczny, 2000) and processing of OVS sentences with unambiguous case marking like (1-b) is associated with increased reading times as compared to SVO sentences. It has been suggested that the subject-first bias cannot solely be explained in terms of the frequency of particular structures, but rather reflects the application of language-specific grammatical principles in order to predict upcoming structural positions (Bornkessel, Schlesewsky, & Friederici, 2002). However, although SVO is the preferred order, unambiguous morphological cues indicating a non-canonical ordering of arguments are processed incrementally, leading to rapid revision of the assumed SVO template in case-ambiguous OVS sentences. Yet, reanalysis is costly, as reflected by the increased end-of-sentence reading times (Bader & Meng, 1999; Schlesewsky, Fanselow, Kliegl, & Krems, 2000; Schriefers, Friederici, & Kuhn, 1995) and findings of an early negativity together with a P600-component in ERPs (Frisch, Schlesewsky, Saddy, & Alpermann, 2002; Matzke, Mai, Nager, Rüsseler, & Münte, 2002; Mecklinger, Schriefers, Steinhauer, & Friederici, 1995).

Given the presence of case ambiguities on the one hand and the incremental nature of processing unambiguous morphological cues on the other, Schlesewsky and Bornkessel (2004) have suggested a two-pathway processing system for theta-role assignment in German sentence comprehension. When sentences are case ambiguous, listeners rely on positional order information in terms of a subject-first bias. However, when there is unambiguous case-marking information, a morphological pathway is pursued and theta-role assignment is based solely on morphological information (Bornkessel, Schlesewsky, & Friederici, 2003; Bornkessel, McElree, Schlesewsky, & Friederici, 2004).

Besides reading time and ERP studies, processing of morphological cues in German has also been studied using the visual-world paradigm (Kamide, Scheepers, & Altmann, 2003; Knoeferle, 2007; Knoeferle, Crocker, Scheepers, & Pickering, 2005). The visual-world paradigm provides an excellent method to investigate language comprehension in real time (Boland, 2005). By recording the eye-movements on visually presented scenes or objects during spoken language processing, studies applying the paradigm reveal the rapid online processing mechanisms used during syntactic parsing and sentence comprehension. These processes may remain concealed in experiments looking at offline comprehension or end-of-sentence reaction times in isolation. Thus, although offline results in reading time studies suggest that processing of OVS sentences is associated with higher processing load, visual-world studies reveal that, in German declaratives, listeners incrementally integrate unambiguous case-marking information (Kamide et al., 2003). Moreover, visual-world data has provided evidence that in case of a sentence-initial NP unambiguously marked for accusative case, the preferred SVO template is rapidly revised and listeners immediately pursue an OVS parse after processing of the verb. Applying the visual-world paradigm, Knoeferle (2007) also found that identification of non-canonical structure in case-unambiguous sentences is directly time-locked to processing of the case-marking cue in NP1, even before the verb has been processed. For sentences with a case-ambiguous first NP, on the other hand, participants preferred an SVO interpretation. However, just as in unambiguous sentences, listeners rapidly reanalyzed their subject-first expectation immediately after morphological information later in the sentence (e.g., the case marker at NP2) indicated an OVS structure. Such disambiguation is directly time-locked to processing of the respective lexical or morphological information, as revealed by reading times and ERP patterns (Knoeferle, Habets, Crocker, & Münte, 2008; Matzke et al., 2002).

When it comes to processing of verb inflection cues, studies have looked at wh-questions with a case-ambiguous wh-NP (i.e. a case-ambiguous wh-question-word together with a case-ambiguous noun), which is disambiguated by a number-agreement cue at the verb (beim Graben, Saddy, Schlesewsky, & Kurths, 2000; Meng & Bader, 2000). These studies revealed that, in general, there is a subject-first bias for such sentences. However, the inflectional cue at the verb triggers immediate reanalysis towards an object-first when it is not in agreement with the number of the initial NP. According to Meng and Bader (2000), in wh-questions, this reanalysis is associated with higher processing costs as compared to reanalysis triggered by case cues. However, in German, object wh-questions occur more frequently than object-initial declaratives and such differences in input frequency may affect sentence processing. Therefore, the finding of different costs of reanalysis in questions may not be directly extended to declaratives. Furthermore, to our knowledge, so far no study has used the visual-world paradigm to investigate online processing of inflectional agreement markers on verbs as a cue to theta-role assignment in unimpaired sentence comprehension in German. Thus, it is still an open question whether in German declaratives, reanalysis triggered by verb-agreement cues is indeed more costly than reanalysis triggered by case markers.

In summary, for processing of case-ambiguous German main clauses, studies have found a strong bias for an SVO interpretation of the sentence. However, morphological cues available later in the sentence (for example, case markers at NP2) are incrementally integrated into the current parse and the structure is rapidly revised towards an OVS order. Moreover, studies investigating processing of wh-questions revealed that, when no unambiguous case cues are available, verb inflection cues trigger rapid revision towards a non-subject-initial word order. Nevertheless, the costs of reanalysis are reflected in increased processing time, as revealed by studies measuring offline data such as reading times or end-of-sentence response times in additional tasks like sentence-picture matching. For sentences containing unambiguous morphological cues at the initial NP1, listeners override their subject-first bias even more quickly, due to rapid cue integration. However, just as in case-ambiguous sentences, end-of-sentence response times are still increased for case-unambiguous OVS sentences.

In recent years, prediction and integration processes have gained a lot of attention in sentence comprehension research (e.g., Gibson, 2000; Levy, 2008; Lewis & Vasishth, 2005). It is commonly assumed that when hearing a sentence like (1-a), listeners treat the sentence-initial NP as the subject of the sentence and predict a verb in third person singular. Once the listener encounters the transitive verb, a post-verbal object-NP is predicted. By contrast, in a sentence like (1-b), the unambiguous accusative cue at NP1 leads to prediction of an OVS structure with a transitive verb. On encountering the verb, listeners predict a post-verbal subject-NP. For sentences like (3-a) and (3-b), in the absence of unambiguous case cues, listeners base their predictions on canonical assumptions and presume that NP1 is the subject of the sentence. Due to its number marking, they predict a verb in third person singular. In a sentence like (3-a), this prediction is confirmed at the verb and the verb's agreement cue can be integrated properly and a post-verbal object-NP is predicted. In contrast, in a sentence like (3-b), a garden-path arises because the predicted verb morphology does not confine with the agreement cue in plural. In order to integrate this cue, listeners revise their parsing decision towards an OVS structure and finally predict a post-verbal subject-NP. Little is known about such purely morpho-syntactic predictive abilities related to sentence processing in aphasia. There is evidence that IWA are impaired in generating predictions about upcoming arguments (Mack, Ji, & Thompson, 2013), however, this study focused on lexical-semantic prediction based on the meaning of the verb in a sentence. No study so far has investigated prediction and integration based on morpho-syntactic cues in aphasia. We will return to these issues in the general discussion and discuss the impact of our findings on assumptions about predictive abilities in aphasia.

Aim of the study and predictions

We present two experiments which we conducted in order to obtain further insights into unimpaired and impaired processing of different syntactic and morphological cues for theta-role assignment. By investigating offline as well as online sentence processing using different kinds of syntactic structures and morphological parameters, we aimed to further identify the source of IWA's sentence comprehension deficits and to shed light on the processing strategies IWA rely on. We conducted two eye-tracking experiments in which we applied a version of the visual-world paradigm (Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995) that has been shown to be especially suitable for studying aphasic sentence comprehension (Dickey et al., 2007; Hanne et al., 2011; Meyer et al., 2012). In this version, eye-movements are being monitored while participants perform an auditory sentence-picture matching task frequently used to assess sentence comprehension in aphasia. The combination of both, the offline and the eye-tracking measures, enables us to characterize IWA's online processing patterns reflecting successful sentence comprehension (i.e., correct offline responses) as well as unsuccessful processing of syntactic and morphological cues (i.e., incorrect offline responses).

Specifically, in Experiment 1, we looked at the online processing of unambiguous case-marking cues in German SVO and OVS sentences in IWA and controls. Experiment 2 investigated online processing of a number-marking cue at the verb in case-ambiguous SVO and OVS sentences.

Regarding eye-movements, for controls, we expected to observe evidence for a subject-first bias for case-ambiguous sentences (Experiment 2). Moreover, we expected the gaze patterns to reflect rapid morphological cue integration and revision of the assumed SVO-template towards an OVS parse as soon as unambiguous morphological information comes in, i.e. after processing of the verb in Experiment 2 and after processing of NP1 in Experiment 1.

For IWA, we expected an effect of word order on IWA's offline accuracy in both experiments. That is, in the sentence-picture matching task, we expected lower accuracy in OVS as compared to SVO condition. Given the findings by Burchert et al. (2003), we hypothesized that the sentence comprehension deficit will be less pronounced for OVS sentences involving unambiguous case marking (Experiment 1) than OVS sentences involving a number marking cue (Experiment 2).

With respect to IWA's eye-movements, in Experiment 1, based on our findings in Hanne et al. (2011), we expected delayed cue integration in correct responses. If, similar to case cues, processing of inflectional cues is not impaired across the board in IWA, but delayed, we hypothesized that, in Experiment 2, IWA's gaze pattern in correct responses should reflect instances of successful, albeit late-emerging integration of the number-marking cue. Moreover, in both experiments, we were interested in determining whether IWA deploy the same subject-first preference and reanalyses processes as seen in controls.

Experiment 1

In this experiment, the target structures were German declarative sentences, in either SVO or OVS order. All sentences contained NPs unambiguously and overtly marked for nominative or accusative case. This allowed us to explore the effect of word order and its interaction with case morphology in controls and IWA.

The experiment is similar to the one reported by Hanne et al. (2011). However, in contrast to the previous study, the pictures and sentence stimuli in the matching task were not presented simultaneously, but with a fixed amount of preview time. This allowed participants enough visual processing in order to get the gist of the action pictures before the linguistic input comes into play. In addition, the inclusion of a preview time allows for a more straightforward comparison to previous experiments conducted with controls and IWA, because these also presented the pictures already before the sentences were played. Another improvement over the previous study is that it had included only eight control participants; in the present study, we have a larger group of controls.

Participants

Twenty-eight participants took part in this experiment: eight IWA and 20 control participants. The control group (labeled C in the figures below) included 20 adults (14 female, 6 male) ranging between 42 and 76 years of age ($M = 61$, $SD = 10$). They were recruited from the staff of the university and via announcement in the neighbourhood and they were paid. All control participants were right-handed and reported no history of neurological, psychiatric, learning or hearing impairment. The IWA (3 female, 5 male) were between 39 and 72 years old ($M = 55$, $SD = 13$) and post-onset time ranged between 5 and 18 years. Their aphasia was due to a single unilateral lesion in their dominant hemisphere. All but one IWA were pre-morbidly right-handed as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971). Demographic and neurological data of IWA are provided in Table 1. The group of IWA was age-matched to the group of controls ($t(11) = 1.3$, $p > .05$, unpaired t-test, two-sided). In addition, both groups were matched for education (in terms of years of schooling).

According to their medical report, hearing was unimpaired in the IWA. Syndrome classification and severity of aphasia was assessed using the Aachen Aphasia Test (Huber, Poeck, Weniger, & Willmes, 1983). In order to exclude IWA suffering from auditory comprehension difficulties at the word level (due to impairments in auditory analysis, the phonological input lexicon or semantics of single words), the IWA were tested on selected tasks (auditory discrimination, auditory word-picture matching) of the LeMo battery (De Bleser, Cholewa, Stadie, & Tabatabaie, 2004), a German psycholinguistic assessment of language processing similar to the English PALPA test (Kay, Coltheart, & Lesser, 1992). Details of IWA's results in these pre-tests are given in Table 2.

Syntactic comprehension abilities were assessed using a German sentence comprehension test (Saetze Verstehen; Burchert, Lorenz, Schröder, De Bleser, & Stadie, 2011), of which we conducted subtest one (irreversible sentences) and two (case-marked and number-marked SVO and OVS sentences). The test provides control data for two age groups (20-49 years, 50-80 years) and syntactic comprehension is considered impaired when performance is below 2 standard deviations from the respective control mean. All eight IWA performed below 2 SD from the control mean in at least the non-canonical conditions. Sentence comprehension scores are given in Table 3.

All participants were native speakers of German. Vision was normal or corrected-to-normal and participants wearing glasses or contact lenses were told to keep them on during the experiment and pre-tests.

Material

Sentence Stimuli. For 23 transitive German verbs, describing simple depictable events, we constructed a semantically reversible SVO as well as an OVS version, containing the same two animate nouns.

All nouns used were animate, denoting either human entities (family members, professions), animals or fairytale characters (like witch, dwarf, ghost). The two nouns in a sentence always belonged to the same of those three categories. No words of foreign origin were used and all nouns were simple and masculine. In all sentences, the accusative case of the object was unambiguously marked at the definite determiner of the NP. Lemma frequency for all nouns was obtained using the DLEXDB lexical database (Heister et al., 2011). No statistically significant difference in frequency was found between the set of nouns used as subjects and the set of nouns constituting the objects ($M_1 = 10028$, $SD_1 = 17995$; $M_2 = 6785$, $SD_2 = 9446$; $t = 0.32$, $p > .05$).

In order to make sure that the reversible sentences are controlled for their semantic plausibility, and that in each sentence both the subject and the object NP have equal probability of being the agent or theme of the action, we conducted a rating study. Details of this study are provided in the Appendix.

For 10 of the 23 verbs, additional irreversible sentences containing an animate subject and an inanimate object were designed in two conditions. Thus, the overall item set for the experiment comprised 66 German declarative sentences, of which 60 were test items and six served as practice items. Of the 60 test items, 40 sentences constituted the reversible targets in SVO and OVS condition. The 20 irreversible sentences were included as fillers. In order for the fillers to be distinct enough from the targets in terms of syntactic structure, they were assembled as canonical and non-canonical irreversible passive sentences.[3] The practice items comprised two SVO items, two OVS items, one canonical passive and one non-canonical passive. Examples for each sentence type are given in (4).

(4) a. **Case SVO:**

Der Arzt schubst den Dieb.

the$_{NOM}$ doctor is pushing the$_{ACC}$ thief

'The doctor is pushing the thief.'

b. **Case OVS:**

Den Arzt schubst der Dieb.

the$_{ACC}$ doctor is pushing the$_{NOM}$ thief

'The doctor is pushing the thief.'

c. **Case non-canonical passive:**

Der Sack wird vom Schmied gezogen.

the sack is by the smith pulled

'The sack is being pulled by the smith.'

d. **Case canonical passive:**

Vom Schmied wird der Sack gezogen.

by the smith is the sack pulled

'The sack is being pulled by the smith.'

[3] Note that besides the usual passive sentence structure, German allows passives in which the prepositional phrase (the by-phrase) is topicalized to sentence beginning, resulting in a non-canonical order of subject and object, but canonical ordering of thematic roles (agent before theme).

Sentence Recordings and Post-processing. Sentences were digitally recorded in a designated recording room at the University. They were spoken by a trained female native speaker of German keeping the prosodic contour of the sentences as neutral and constant as possible. Each experimental sentence was recorded several times keeping recording settings stable so that various recording versions were available.

Sentences in the two target conditions should only differ in their word order features and no prosodic cues to sentence meaning should be interfering. As processing of linguistic prosody can be impaired in aphasia (Baum & Pell, 1999; Kimelman, 1999), IWA cannot rely on prosodic cues to sentence interpretation as controls would do. Thus, controls could profit from prosodic cues whereas IWA could not, leading to uncontrolled negative effects on sentence comprehension in participants with aphasia.

Therefore, we post-processed the recorded sound files using Praat software (Boersma & Weenink, 2009) and matched the SVO and OVS version of an item with respect to three prosodic parameters: (1) F0 contour, (2) duration of sentence constituents, and (3) F0 rise time.

In order to assess the F0 contour, we determined, for each sound file, the overall amount of F0 rise within NP1 by calculating the difference between the minimum and maximum F0 within this NP. Following this, we paired this F0-rise value of each recorded version of an SVO item with the F0-rise value of each recorded version of the respective OVS item. For each pair, we calculated the ratio between the two values and determined, for each item, the pair for which this ratio was nearest to 1. Such a ratio of 1 for a given pair would denote that the two NP1s are comparable in their F0 contour. Ratios for the selected pairs ranged between 0.87 and 1.15 (M = 0.998, SD = 0.066).

In the next step, we compared the constituent durations of the selected recordings for SVO and OVS sentences. This revealed that, in the mean, NP1 had a slightly shorter duration in the SVO condition. This was probably due to the fact that the determiner *der* is articulated faster than the determiner *den*, which appeared in NP1 of the OVS condition. We therefore identified those pairs of items for which the difference in NP1 durations was most marked and manipulated the shorter NP1 by elongating it artificially. The factors used for this elongation procedure were very small (mean: 1.13, range: 1.09 - 1.2). Overall, 8 out of the 20 NPs in the SVO condition and 2 out of the 20 NPs in the OVS condition were manipulated. In order to rule out any effects of the NP1 manipulation on processing of the auditory sentences during the experiment, a rating study was carried out. The manipulations were not recognized by the participants. Details on the rating are provided in the online supplement. Final mean constituent durations per target condition are provided in Table 4. There were no more significant differences across conditions after applying the elongation procedure (NP1: $t = 0.26$, $p > .05$; verb: $t = 1.91$, $p > .05$; NP2: $t = 0.35$, $p > .05$).

The *F0 rise time* for each NP1 was measured by calculating the amount of time between the point of F0 minimum and the point of F0 maximum. This was done after the duration manipulations were applied to the sounds. If significant differences existed in F0 rise times between conditions, this could be indicative of a different weighting of prosodic cues across conditions. However, statistical comparisons showed that this was not the case (SVO condition: $M = 332$ms, $SD = 156$; OVS condition: $M = 321$ms, $SD = 109$; $t = 0.27$, $p > .05$, two-sided).

For all target recordings finally entering the experiment, the overall speech rate in terms of syllables per second was determined. Again, we found no statistically significant differences between conditions ($M_{SVO} = 3.13$, $SD_{SVO} = 0.45$; $M_{OVS} = 3.12$, $SD_{OVS} = 0.43$; $t = 0.41$, $p > .05$). The different recordings for the filler items in the experiment (canonical and non-canonical passives) were judged by two native speakers of German who were not involved in the experimental design. They were told to choose for each filler the one recording version that sounded best in terms of articulation and recording quality. Finally, all chosen target and filler recordings were normalized using Audacity 2.0.2 software (a free, open source software for recording and editing sounds, version 2.0.2, web.audacityteam.org).

Visual Stimuli. The visual stimuli consisted of black-and-white line drawings depicting the semantically reversible events of the material set and for each event, two pictures were drawn. The two pictures for the target sentences differed with respect to the mapping of thematic roles on the two characters. Whereas in one picture, one of the characters was the agent of the action, the same character constituted the theme in the other picture. Therefore, for each sentence, one target picture (depicting the action mentioned in the sentence) and one foil picture (depicting a theta-role reversal) existed. For the irreversible fillers, the foil pictures depicted either a non-matching subject or a non-matching object. Sample pictures are provided in Figures 1 and 2.

According to the terminology of Henderson and Ferreira (2004), the pictures can be classified as scene sketches depicting a limited number of objects which interact in a semantically coherent and meaningful way. They all had comparable scene content and the depicted persons and objects were of approximately same size. Across all pictures, the action direction (i.e., the spatial position of the theme in relation to the agent) was balanced and there were as many pictures having their action directed to the left (with the theme being positioned to the left of the agent) as pictures with the action directed to the right (i.e., the theme was depicted to the right of the agent).

All pictures were controlled for their comprehension agreement in a norming study involving 30 participants (1 male, 29 female, mean age: 24 years, range: 19-39 years). In this study, the picture pairs for the 60 experimental items were shown in a pseudo-randomized order and participants performed an auditory sentence-picture matching task. Mean correct picture selection across all participants and items was at 99%.[4]

[4]After the first round of comprehension agreement, correct picture selection for one picture pair was below two standard deviations from the mean score. Hence, these two pictures were digitally post-processed to improve the clarity of the characters. Afterwards, comprehension agreement for this picture pair (and for three more picture pairs used as fillers) was assessed again with another group of 30 students (all female, mean age: 25 years, range: 19-49 years). This time, mean correct picture selection for the critical item was at 100%.

Procedure

The experiment consisted of a practice phase during which the practice items appeared in a fixed order, and a testing phase. For the presentation of the test items, four different randomization lists were designed, each starting with a filler sentence. Within each list, two targets were always paired with one filler item. Within these pairs, four different orders of conditions were possible (filler-SVO-OVS, filler-SVO-SVO, filler-OVS-SVO, filler-OVS-OVS) and these orders were equally distributed across the whole presentation list. The lists were constrained so that sentences containing the same verb were never adjacent but separated by at least three different sentences. Moreover, no more than three SVO or OVS sentences followed one after another.
The position of the target picture was balanced across all experimental trials, so that for half of the sentences the target appeared on the left hand side of the screen, and on the right hand side for the other half of sentences. In addition, the target picture position was identical for no more than three consecutive items.
Participants' eye-movements were monitored with a remote Tobii T120 eye-tracker (binocular tracking, Tobii Studio software version 1.7.2, accuracy: 0.5 degrees, head-move-tolerance: 30 x 22 x 30 cm) at a sampling rate of 60 Hz. The screen of the eye-tracker served as presentation screen for the experimental videos (screen size: 17 inch, resolution: 1280 x 1024 pixels). Participants were seated in a comfortable position at approximately 60 cm in front of the screen.
The task was to identify the picture correctly matching the sentence as accurately and fast as possible. Two buttons of a keyboard placed to the left of the sitting chair served as response buttons. Participants used two fingers (the index and middle or ring finger) of the left hand to press the button. One participant with aphasia who was pre-morbidly left-handed and had his lesion in the right hemisphere used the index and middle finger of the right hand.
For each experimental item, a movie file was prepared using Adobe Flash CS 4 Professional (version 10.0.2, Adobe Systems Incorporation, 2008). Within each movie,

at first, the two pictures were shown for 3000 ms, so that for each experimental item there was a fixed preview time for the visual material. Eye-tracking studies involving IWA often use rather long preview times, up to 10 seconds. In contrast, visual-world studies involving language-unimpaired participants usually apply a much shorter preview (for example, 1000 ms). However, in order to adjust for the different participant groups and to comply with the conventions in the field of aphasia research, we decided on 3000 ms. Both pictures were presented next to each other on black background with a black bar (100 pixel wide) between them. The picture size was 550 x 380 pixels, so that the video with both pictures separated by the black bar was 1200 x 380 pixels. Following the preview, both pictures disappeared and a black-and-white smiley was presented in the middle of the screen for 600 ms. This smiley served as visual attractor to centre participants' eye gaze before the critical regions of the test phase started for the item. In this phase, both pictures re-appeared on the screen together with the auditory sentence. The video presentation ended with the participant's button press. The maximum response time was set to 15 s and reaction times exceeding 15 s were considered time-outs. Before the presentation of the next experimental movie, a smiley was shown in the centre of the screen for 600 ms to prepare participants for the following item.

The eye-tracking experiment always started with a 9-point calibration procedure displaying red dots on a black screen. Following this, the practice phase started. After the practice items were finished participants were given time to ask questions. None of the participants needed to repeat the practice phase. Calibration was repeated before the beginning of the test phase. For IWA, there was a pause (lasting 5-10 minutes) after the first half of the items. Calibration was repeated for the second half of the experiment.

Prior to testing, every participant received a written description of the experiment including a specification of the eye-tracking method. For IWA, the form was additionally read out by the experimenter. After all possible questions had been answered, participants signed informed consent. In addition, the IWA received a paper-and-pencil version of the practice items in order to make sure that they were able to carry out the task (sentence-picture matching). Testing took place in a designated eye-tracking laboratory at the University of Potsdam or in a cooperating institution in Berlin. Overall, for participants with aphasia the experiment lasted approximately 45-60 minutes with a test phase of 25-35 minutes. For control participants, the whole experiment took about 30 minutes.

Data Analysis

Null responses (i.e., no reaction within 15 s after sentence onset) were treated as errors. For controls, only correct responses entered the analyses of reaction times and eye-movements. For IWA, in addition to the respective group comparisons, reaction times and eye-movement data were analyzed separately for correct and incorrect responses in the sentence-picture matching task.

For the analysis of eye-movements, visual *areas of interest* (AoIs) consisted of the two pictures presented on the screen and we measured gaze proportions to the target (correct) and foil (incorrect) picture, respectively. Gaze position was defined based on the combined eye-tracking data from both eyes. Moreover, we applied the criterion that the combined gaze position be stable in a radius of 35 pixels for at least 100 ms in order to be treated as a gaze. The gaze proportions were subsequently collapsed within each participant and item for each of the auditory *regions of interest* (RoIs). These were defined according to the sentence constituent structure (NP1, verb, NP2) individually for each item. The final RoI was the period of silence after sentence offset until the participant's button press.

For statistical analyses, we used linear mixed models with participants and items as random effects. For data on accuracy, a generalized linear mixed model with a binomial link function was fit. For reaction time and eye-movement data, we applied a linear mixed model with a Gaussian link function. Depending on the outcome measure, condition, group and RoI were treated as fixed effects. Some analyses also included accuracy as a predictor. The group of controls was coded as baseline category against which the group of IWA was compared. In order to statistically compare the gaze proportions across the different RoIs, we applied a successive differences contrast coding. Hence, from the second RoI onwards (NP1), gaze proportions during each RoI were compared to those of the previous RoI. For model parameter estimation, we used maximum likelihood estimation. For determining statistical significance of any predictor variable, we carried out the generalized likelihood ratio test: a mixed model with and without the predictor was compared. We report the chi-square statistic for those model comparisons in the Appendix. Residuals in the linear mixed models were checked for their distributional properties. For the coded contrasts, unless stated otherwise, coefficient estimates (b), their standard errors, and t- or z-scores (depending on the dependent measure) are provided in the Appendix. An absolute t-score of 2 or greater indicates significance at the alpha level of 0.05. For the generalized mixed models, p-values are provided in addition to z-scores.

Results

Sentence-picture matching accuracy and reaction times. Controls performed at ceiling in the two conditions (SVO: M = 0.96, SE = 0.01, OVS: M = 0.93, SE = 0.02). Patients performed worse than controls (SVO: M = 0.77, SE = 0.03, OVS: M = 0.46, SE = 0.04). There were no null responses. We fit a generalized linear mixed model with condition, group and the interaction between condition and group as predictors. As revealed by model comparisons, there were significant main effects of both predictors and an interaction. The (marginal) main effect of condition on accuracy was due to IWA performing worse with OVS compared to SVO items.

In order to investigate whether this effect of condition was true for all of the IWA, we additionally conducted a single-case analysis using the chi-square statistic. The results are provided in Table 5. The effect of word order was significant in six out of the eight

IWA (although for A06 it was only marginal).

Analyses of variance comparisons of linear mixed models with condition, group and the interaction of condition and group as predictors revealed main effects of condition and group, but no interaction. Reaction times (ms) were significantly higher in OVS compared to SVO condition (Controls: SVO: M = 2168, SE = 44.19, OVS: M = 2565, SE = 58.16; IWA: SVO: M = 4014, SE = 136.01, OVS: M = 4429, SE = 153.06), and, overall, IWA's RTs were significantly higher as compared to controls.

In order to look for an effect of response accuracy on IWA's RTs, we fit a mixed model with accuracy and condition as predictors. This revealed a significant main effect of accuracy, but no effect of condition and no accuracy:condition interaction. Thus, RTs for IWA's correct responses were generally faster as compared to their incorrect responses, irrespective of condition (M_{cor} = 3957, SE_{cor} = 121.88, M_{incor} = 4640, SE_{incor} = 176.95).

Eye-movements. Figure 3 shows gaze proportions to the target picture for controls and for IWA's correct responses. Gaze proportions for IWA's incorrect responses in the sentence-picture-matching task are provided in Figure 3.

As for the accuracy and RT data, we fit linear mixed models with participants and items as random effects. The fixed effects included group, condition and RoI, as well as their interaction. Patients' gaze data for correct and incorrect responses were analyzed separately.

For *correct* responses, we found main effects of group, condition, and RoI and there was a significant interaction of group, condition, and RoI. The three-way interaction coefficient was significant at NP1, indicating that for this RoI contrast the increase in gaze proportions to the target was lower in IWA as compared to controls in OVS, but not in SVO condition. Moreover, for controls, no significant differences were found in gaze proportions in SVO and OVS condition during NP1. At the verb, both groups showed an increase in gaze proportions to the target, but this effect did not interact with group or condition. At NP2, the overall increase in gaze proportions was also significant, but it was reduced in IWA as compared to controls. Furthermore, a significant interaction of group, condition, and RoI indicates that, for IWA, the increase in gaze proportions at NP2 was steeper in the OVS condition as compared to the SVO condition. For the silence region, both groups showed increased gaze proportions as compared to the previous RoI, however, this increase was more marked in IWA as compared to controls, irrespective of condition.

For IWA's *incorrect* responses, model comparisons revealed a main effect of group and RoI. The effect of condition was not significant and did not interact with any of the other predictors. A significant group-RoI interaction was seen at NP2 and silence. This indicates that, although in controls, gaze proportions to the target increased from verb to NP2, and from NP2 to silence, the pattern was different in IWA. Thus, we fit a separate model for IWA's gaze data including condition, RoI and the condition-RoI interaction as fixed effects. This revealed that, for IWA's incorrect responses, gaze proportions in a given RoI were significantly different from the previous RoI only for the silence region. Moreover, in contrast to controls, gaze proportions decreased, leading to a reversed effect. This effect did not interact with condition.

Discussion

This experiment investigated online processing of case-marking cues to theta-role assignment in German SVO and OVS sentences. We will first review the results for control participants, before discussing the findings for the IWA.

Eye-movements of controls show that the accusative case marker at NP1 in OVS sentences was integrated immediately into the current parse tree. Moreover, gaze proportions to the target significantly increased immediately after processing of NP1 in both conditions, indicating that, regarding controls' gaze data, OVS sentences did not show a disadvantage over subject-initial sentences. In line with previous eye-tracking studies (Kamide et al., 2003; Knoeferle, 2007) this constitutes further evidence that, in German, case markers indicating a non-subject-first structure are processed incrementally. In addition, this is evidence that the preferred SVO template is overridden very quickly.

However, in our experiment, end-of-sentence reaction times were higher in the OVS as compared to the SVO condition, replicating the results for controls in Hanne et al. (2011). This suggests that, although morphological cue integration is incremental, the overall processing demands are higher in OVS as compared to SVO structures. In order to disentangle this seeming contradiction, it is important to keep in mind the differences in methodology between our and previous studies within the visual-world paradigm involving language-unimpaired participants. Usually, visual-world studies employ visual displays of four to five objects and, in most cases, there is no explicit task. In contrast, we used rather complex images depicting action scenes and, in addition, participants were to perform a sentence-picture matching task. It is possible that the higher reaction times we observed in the selection task in OVS condition were caused by an interaction of linguistic structure and task demands or complexity of the visual material. As suggested by Hanne et al. (2011), it is possible that determining the theta-role relations of an OVS sentence and matching them to the correct picture takes longer as compared to SVO structures. Under this view, our finding of higher processing demands for unambiguously case-marked OVS as compared to SVO sentences in German is in line with previous findings looking at offline measures such as reading times or end-of-sentence responses (Bader & Meng, 1999; Bornkessel, Zysset, Friederici, von Cramon, & Schlesewsky, 2005; Schlesewsky et al., 2000; Schriefers et al., 1995).

Turning towards the results for IWA, as expected, they were less accurate than controls and we found an effect of word order on IWA's performance in the sentence-picture matching task, i.e., they were less accurate with OVS as compared to SVO sentences.

Analyzing each IWA separately, the word order effect was replicable in six out of the eight IWA. Two IWA were equally impaired with both sentence types. However, it should be noted that this may be related to the fact that, overall, IWA showed some increased difficulties with the SVO sentences in this experiment. Performing only at

77% correct in the mean, their performance on SVO sentences was lower as compared to their performance in the pre-test on syntactic comprehension abilities. In this pre-test, the IWA had performed either within or only slightly below the normal range on the case-marked SVO sentences. There are various reasons which could account for the higher error rate in the eye-tracking experiment. Firstly, although the syntactic structures were the same, the lexical material used in the experimental sentences was not matched to that of the pre-test and, therefore, different from it. This could have influenced the results. Secondly, performing a computerized sentence-picture matching task on the eye-tracker may be considered more difficult as compared to the paper-and-pencil task in the pre-test. Furthermore, in the eye-tracking experiment, participants were asked to find the matching picture as fast as possible while in the pre-test, they were given as much time as needed. Finally, during the experiment, the interstimulus interval (i.e., the time between a participant's response and the presentation of the following item) was set at 600 ms, whereas no prescribed interstimulus interval was given in the pretest (and moving towards the next item could potentially proceed slower in a paper-and-pencil task as compared to computerized presentation). In any case, although IWA's performance was overall worse than in the pretest, they yet performed less accurately on OVS as compared to SVO structures, complying to the hypothesized effect of word order. In contrast, as expected, word order had no effect on accuracy in any of the two controls groups.

We now turn to the discussion of IWA's eye-movements in cases where the case cue was processed correctly. When appropriate, we compare their performance with that of controls. For correct responses in both conditions, IWA showed delays in their gaze pattern. They were not as fast as the age-matched controls in directing their gaze towards the target picture. This is consistent with the finding of delayed syntactic processing in Hanne et al. (2011) and constitutes further evidence for the slowed processing account of sentence comprehension in IWA (Burkhardt, Avrutin, Piñango, & Ruigendijk, 2008; Dickey et al., 2007).

As indicated by the three-way-interaction, during processing of the unambiguous NP1, IWA exhibited a higher proportion of gazes towards the picture promoting a subject-first interpretation (i.e., the target picture for SVO sentences and the foil picture for OVS sentences). This suggests that, for processing of sentences with unambiguous case markers, IWA initially commit to an SVO prediction. However, after processing of the verb in the SVO condition, their prediction process is different from controls: while controls' gazes to the target start increasing immediately after the verb (indicating successful prediction of the post-verbal object-NP), this is not the case for the IWA. By contrast, IWA's gazes to the target significantly increase immediately after NP2 was processed. This observation is indicative of impairments in predictive processes in aphasic sentence comprehension. For processing of NP1 in the OVS condition, controls' data is not indicative of an initial subject-first prediction as the accusative case marker was integrated immediately. IWA, on the other hand, initially pursued an SVO prediction, suggesting that upon processing the unambiguous case marker, they initiated prediction of upcoming syntactic structure. However, IWA were not using the case-marker as quickly as controls do in order to build the OVS parse. During processing of the verb, both groups showed an increase in gaze proportions to the target, but there were no statistically significant differences between the two groups. At NP2, IWA's gaze proportions to the target increased in the OVS condition. This is evidence that, in correct responses, IWA did actually process the accusative case marker at NP1 in OVS sentences, although for them, the integration of the cue and the following revision of the presumed SVO template took longer as compared to controls.

Thus, we conclude that, in aphasia, two distinct mechanisms of processing unambiguous sentences are affected by a slow-down: (1) the integration of the case-marking cue and (2) in case of OVS sentences, the revision of the preferred SVO-template. In addition, our results suggest deficits in predictive processes during parsing in aphasia. Although the unambiguous case cue of NP1 lead to prediction of upcoming syntactic structure in both groups, on encountering an accusative cue IWA's predictive processes were initially restricted to subject-first assumptions whereas controls immediately engaged in OVS predictions.

The results for IWA's correct responses in the OVS condition are slightly different from the earlier study (Hanne et al., 2011), in which no early gazes to the picture promoting the SVO interpretation had been observed. However, when comparing both studies, differences in procedure and participants need to be taken into account. Specifically, in the earlier study, the pictures and sentences were presented simultaneously without providing a preview time of the visual material and this may have led to different results. In fact, in the previous study, a similar subject-first bias, which was seen in IWA at NP1 in the current study, was evident only slightly later, namely during processing of the verb. Moreover, similar to the current results, there were indications of delayed attempts to revise the SVO-template, although these were often unsuccessful. We therefore believe that the current study may not be inconsistent with the earlier results. In addition, applying the modified method, which included a preview, contributed to a more detailed understanding of the effects of slow-down on morpho-syntactic processing in aphasia.

In contrast to controls and to IWA's correct responses, the gaze pattern in IWA's incorrect responses in the OVS condition was marked by non-significant differences in looks to target and foil picture until most of the sentence had been processed. Similar differences between controls and IWA during early regions in the sentence have also been observed for IWA's erroneous responses in our earlier study. While incorrect responses in that study were marked by an early-emerging preference for an SVO interpretation, the current results point to a late-emerging gaze preference for the foil picture in incorrect responses. However, as mentioned above, the earlier study did not include a picture preview and, thus, the gazes during early stages of the sentence presentation may have been influenced by cognitive processes other than syntactic parsing (for example, visual processing). Yet, in the current study the procedure was designed in a manner that allows us to tap syntactic processing more explicitly. We therefore argue, that the rather late gaze preference for the picture promoting the SVO interpretation in incorrect OVS responses reflects a delay in arriving at a decision to build an SVO or OVS structure, and is indicative of non-successful processing of case-marking cues.

Experiment 2

In this experiment, we investigated processing of case-ambiguous SVO and OVS sentences in German. Both NPs in the sentences were ambiguous between nominative and accusative case. The only morphological cue towards an SVO or OVS interpretation was given through the verb morphology in terms of a subject-verb-number-agreement marker. Thus, this experiment investigated whether and how controls and IWA make use of verb inflection as a morpho-syntactic cue for theta-role assignment.

Participants

This experiment comprised 28 participants: the same eight IWA who had participated in Experiment 1 and 20 controls (14 female, 6 male, M = 62.5, SD = 9.5, range 44 -76) without any history of neurological, psychiatric, learning or hearing impairment. They were recruited via announcement and were paid for participating. Five of the control participants had taken part in the other experiment, however, there was a gap of at least 6 months between both experiments. Moreover, after the first experiment, there was no debriefing with respect to the specific goal of the experiment. All controls were right-handed. For IWA, demographic and neurological data are provided in Tables 1 and 2. Table 3 gives results of language assessments. We made sure that, for each IWA, there was a lag of at least 2 weeks between participating in each of the two experiments. In addition, the order of both experiments was randomized across all IWA. All participants were native speakers of German.

Material

Sentence Stimuli. As in Experiment 1, the item set comprised 66 sentences, of which 6 served as practice items. The reversible target sentences appeared in SVO and

OVS conditions (n = 20, each). Details of a rating study, conducted to control for semantic plausibility of the targets, are provided in the Appendix.

All nouns were simple and animate, and they denoted either family members, animals or fairytale characters. One of the two nouns in a sentence had neuter gender and appeared in singular, hence taking the German determiner *das*. The other noun was feminine and appeared in plural. In German, plural NPs always take the determiner *die*. Both determiners are ambiguous between nominative and accusative case. In SVO sentences, the verb was in present tense third person singular and, hence, agreed with the singular noun at sentence beginning. Verbs in OVS sentences had the present tense third person plural marking and, thus, were in agreement with the post-verbal plural NP. All but one of the plural nouns had the same plural morpheme (-*en*, for example: *die Frau* - *die Frauen*; the woman - the women).

In addition to the targets, there were 20 irreversible fillers (n = 10 canonical passives and n = 10 non-canonical passives). Examples for each sentence type are given in (5). For the target sentences, the set of nouns serving as subjects and the set of nouns serving as objects were frequency-matched ($M_1 = 22231$, $SD_1 = 26571$; $M_2 = 23303$, $SD_2 = 31948$; $t = 0.12$, $p > .05$).

(5) a. **Number SVO:**

Das Kind faengt die Frauen.

the$_{NOM/ACC}$ girl catch$_{3rdpers,sg}$ the$_{NOM/ACC}$ women

'The girl is catching the women.'

b. **Number OVS:**

Das Kind fangen die Frauen.

the$_{NOM/ACC}$ girl catch$_{3rdpers,pl}$ the$_{NOM/ACC}$ women

'The women are catching the girl.'

c. **Number non-canonical passive:**

Das Paket wird von den Frauen geschoben.

the bundle is by the women pushed

'The bundle is being pushed by the women.'

d. **Number canonical passive:**

Von den Frauen wird das Paket geschoben.

by the women is the bundle pushed

'The bundle is being pulled by the women.'

Sentence Recordings and Post-processing. Several tokens of each sentence, spoken by a trained female native speaker of German, were digitally recorded in a recording room at the University of Potsdam. For post-processing of the audio files, the same procedure was applied as in Experiment 1.

First, we calculated the overall F0 rise value within NP1 of each recording and compiled the list of paired F0-rise times of SVO and OVS recordings for each item. Following this, we determined the ratio between the two F0-rise values and identified for each item the pair for which the ratio was nearest to 1. Ratios for the selected pairs ranged between 0.91 and 1.16 ($M = 0.99$, $SD = 0.05$).

Next, we compared the duration of the sentence constituents in the selected SVO and OVS sentence recordings. This revealed that, in the mean, the verb had a longer duration in OVS recordings. This is because in German, the third person singular inflection at the verb is realized through the morpheme *-t* added to the verb stem (although for some verbs, a linking element is required, for example *e* as in *badet* - *bathes*). On the other hand, the third person plural inflection morpheme is *-en*. Hence, plural forms have two syllables whereas most third person singular verb forms are monosyllabic. This explains the longer verb durations in OVS sentences. In order to control for this difference, the shorter verbs were elongated artificially with a mean factor of 1.27. This manipulation was not perceptible, as revealed by a rating study, details of which are provided in the online supplement. Final mean constituent durations per target condition are provided in Table 6. Statistical comparisons showed that the durations did not differ significantly across conditions (NP1: $t = 1.65$, $p >$ .05; verb: $t = 1.94$, $p > .05$; NP2: $t = 1.07$, $p > .05$).

With respect to the F0 rise times of NP1s, there were no statistically significant differences between the conditions (SVO condition: $M = 504$, $SD = 141$; OVS condition: $M = 660$, $SD = 115$; $t = 1$, $p > .05$, two-sided, data in ms).

The overall speech rate in terms of syllables per second was 3.34 ($SD = 0.45$) for SVO sentences and 3.45 ($SD = 0.43$) for OVS sentences. This difference was not significant ($t = 1.72$, $p > .05$). All recordings were finally normalized using Audacity 2.0.2 software (version 2.0.2, web.audacityteam.org).

Visual Stimuli. Black-and-white pictures, complying with the same criteria set up for the pictures in Experiment 1, were drawn. For each target sentence, a target and a foil picture, depicting the event with reversed theta-roles, existed. For the irreversible fillers, foil pictures contained either a non-matching object or subject. In a norming study including 35 participants (mean age: 22 years, range: 18-40 years), mean correct picture selection was 99%. Sample pictures are provided in Figures 4 and 5.

Procedure and Data analysis

The procedure and data analysis was the same as in Experiment 1.

Results

Sentence-picture matching accuracy and reaction times. Controls performed at ceiling in both conditions (SVO: $M = 0.98$, $SE = 0.007$, OVS: $M = 0.91$, $SE = 0.01$). The group of IWA performed worse than the controls (SVO: $M = 0.78$, $SE = 0.03$, OVS: $M = 0.64$, $SE = 0.04$). There were no null responses. We fit a generalized linear mixed model including condition, group and the respective interaction as predictors. Model comparisons revealed overall significant main effects of both predictors and a significant interaction. The effect of condition on accuracy was due to IWA performing worse with OVS compared to SVO items.

As in Experiment 1, in addition to the group analysis of accuracy data, we conducted a single-case analysis in order to investigate the effect of condition in individual IWA. Five out of the eight IWA did not show a significant dissociation between the two conditions. However, for the other three IWA, the effect was significant. The detailed results are provided in Table 7.

For controls, RTs were higher in OVS compared to SVO condition (SVO: $M = 2441$, $SE = 39.5$, OVS: $M = 2999$, $SE = 54.43$). Patients were overall slower (SVO: $M = 4148$, $SE = 122.3$, OVS: $M = 4281$, $SE = 124.65$). Analyses of variance comparisons of linear mixed models with condition, group and the interaction of condition and group as predictors revealed significant main effects and an interaction. Overall, IWA's RTs were significantly higher as compared to controls, and the significant interaction reveals that the effect of condition was different in IWA as compared to controls. While condition had an effect on RTs in controls, this was not the case for IWA's RT data.

The mixed model for the effect of response accuracy on IWA's RTs revealed a significant main effect of accuracy. This effect did not interact with condition. Thus, as in Experiment 1, RTs for IWA's correct responses were generally faster as compared to their incorrect responses, irrespective of condition ($M_{cor} = 4095$, $SE_{cor} = 98.02$, $M_{incor} = 4507$, $SE_{incor} = 178.67$).

Eye-movements. Gaze proportions to the target picture for controls and IWA's correct responses are provided in Figure 6. As in Experiment 1, IWA's gaze data for correct and incorrect responses were analyzed separately and gaze data of incorrect responses in the sentence-picture-matching task are provided in a separate plot. The linear mixed model for statistical analyses included group, condition, RoI and the 3-way interaction as fixed effects. Participants and items were included as random effects.

For both groups' *correct* responses, the model comparisons revealed main effects of group, condition and RoI. Overall, IWA had less gazes to the target picture compared to controls and, overall, there were fewer looks to the target in OVS condition at NP1 and the verb. Moreover, there was an interaction of group, condition, and RoI, which became significant at the verb and NP2. Note that the direction of the contrast changed between the two RoIs from positive to negative. For the other two RoI contrasts, the 3-way interaction was not significant. Thus, whereas the effect of condition was similar in IWA and controls at NP1 and during silence, it was different at the verb and NP2. For controls, in OVS condition, gazes to the target significantly decreased during NP1 and the verb, i.e. there were less gazes to the target picture, but more gazes to the foil picture showing the SVO interpretation. In contrast to controls, IWA showed no decrease in gaze proportions to the target at the verb in OVS condition. Furthermore, the subsequent increase in gaze proportions, which we found for both conditions at NP2 in controls, was not seen for IWA in OVS condition (although it was evident for SVO). However, IWA's gaze proportions to the target did significantly increase in OVS after processing of NP2.

For the comparison of IWA's *incorrect* against controls' correct responses, we found main effects of group, condition, RoI and a three-way-interaction. As for correct responses, overall, IWA had less gazes to the target in both conditions. The three-way-interaction was significant at the verb and at silence. Separate models for the two conditions revealed that the interaction coefficient at the verb was negative for SVO and positive in OVS condition. Thus, at the verb in the SVO condition, controls showed increasing gazes to the target, whereas IWA for their incorrect responses did not (see left panels of Figures 6a and 6b). In contrast, in the OVS condition, controls showed a marked decrease in gazes to the target which was not present in IWA (see right panels of Figures 6a and 6b). Moreover, for the OVS model, there was a group:RoI interaction at NP2 and silence, indicating that gaze proportions to the target decreased in IWA for these two RoIs whereas they increased in controls. Although there was no such interaction at NP2 and silence in SVO condition, IWA nevertheless behaved differently from controls because they had significantly less gazes to the target, as revealed by the overall main effect of group in the SVO model.

Discussion

In this experiment, we investigated whether and how controls as well as IWA make use of verb inflection as a cue for theta-role assignment in case-ambiguous SVO and OVS sentences.

Regarding controls' eye-movements, we expected to find evidence for an initial subject-first assumption which would be rapidly revised towards an OVS structure due to incremental integration of the number-agreement cue at the verb. The gaze pattern of controls confirmed this prediction. For SVO sentences, we observed significantly more gazes to the target picture (i.e., the picture showing the SVO interpretation) already at the case-ambiguous NP1, indicating participants' expectation of a subject-initial sentence structure. This kind of subject-first prediction was also reflected in their high gaze proportions to the foil picture during the case-ambiguous NP1 region in OVS sentences. However, after hearing the verb's number inflection cue, which is not in agreement with NP1 and thus signals an OVS structure, participants' gazes immediately switched towards the target picture. This is evidence for successful agreement cue processing and rapid revision of the predicted subject-first structure towards an OVS interpretation.

Although both groups showed evidence for rapid cue integration and successful reanalysis, end-of-sentence RTs were higher for OVS as compared to SVO sentences. Similar to Experiment 1, this corroborates previous findings on increased offline processing demands for case-ambiguous OVS sentences in end-of-sentence tasks (Meng & Bader, 2000).

Turning to the accuracy and RT results for the IWA, in line with our prediction, they performed less accurately than controls and exhibited an effect of word order with fewer correct responses in the OVS as compared to the SVO condition. However, analyses at an individual level revealed a significant effect of word order only for three out of the eight IWA. This stands in contrast to Experiment 1, in which accuracy was even lower in the OVS condition (46% as compared to 64%) and six out of the eight IWA could be shown to be affected by the word order manipulation. However, the three IWA who had shown the effect in Experiment 2 (A05, A07, A08) also did so in Experiment 1.

We now turn to the discussion of IWA's eye-movements for trials with correct offline responses. In contrast to the findings for controls, during processing of the verb in both conditions, IWA did not show a preference for the picture promoting the SVO interpretation. We therefore conclude, that, contrary to the age-matched controls, IWA were not strongly biased to interpret the case-ambiguous NP1 as the subject of the sentence. Thus, IWA show less capabilities of predicting a subject-initial syntactic structure when sentence-initial cues are ambiguous. The absence of a subject-first processing bias and the issue of impairments in predictive processing is in line with results for online processing of passive sentences and selectional information of verbs in aphasia (Mack et al., 2013; Meyer et al., 2012). However, for SVO sentences, IWA's gaze patterns reflected a subject-first parse immediately after processing of the verb at NP2 and during silence. Thus, as soon as unambiguous morphological information was available (in this case the number-agreement cue at the verb disambiguating the sentence towards an SVO structure), IWA initiated prediction of syntactic structure and posited an SVO template.

Although, in the absence of unambiguous morphological cues, IWA show impairments in early prediction of syntactic structure, the fact that they do commit towards an SVO parsing once unambiguous morphological information is available constitutes evidence for retained abilities to engage in prediction of syntactic structure. This further constitutes evidence that IWA are indeed capable of processing the unambiguous morphological cue at the verb in correct responses, also when this cue finally forces them to pursue an OVS interpretation. Otherwise, they would not start directing their gaze on the picture promoting the OVS interpretation after processing the agreement cue in OVS sentences. Crucially, although the increase in gazes to the target picture mostly arises after NP2, this NP is not carrying any unambiguous morphological information promoting an OVS interpretation. Instead, the only cue signaling OVS structure is the verb inflection.

However, similar to the findings for case cues in Experiment 1, the integration of the inflectional cue was not as fast as in controls. This is indicated by the interaction we found at NP2, which reveals that, after processing of the verb, IWA's gaze proportions towards the target picture did not increase as much as in controls. However, the marked increase in IWA's gaze proportions in the following RoI suggested successful cue integration and their commitment to an OVS parse.

The fact that, for case-ambiguous sentences, the SVO prediction is not instantiated as fast as in controls indicates that IWA delay predicting upcoming sentence structure until unambiguous morphological information–either complying with a subject-first structure or defeating it–has been processed. This is indicative of non-deterministic parsing principles being used by IWA and points towards application of a top-down wait-and-see strategy which gets overridden once unambiguous cues are available.

Taken together, the data suggest that in the absence of unambiguous morphological cues prediction of syntactic structure is impaired in aphasia, whereas integration of morpho-syntactic cues is still preserved, although delayed. The idea of less capacities in prediction is also supported by parallel findings in studies investigating bilingual sentence comprehension in adult L2 comprehenders (DeLong, Urbach, & Kutas, 2005; Martin et al., 2013).[5]

Looking at IWA's eye-movements in cases in which the agreement cue was not processed correctly, the findings are similar to those for the incorrect responses in Experiment 1. The gaze preference for the foil picture, which was (incorrectly) chosen as the matching one, emerged late, after the overall sentence had been processed. This corroborates the idea of late parsing decisions and failure to successfully predict upcoming sentence structure and to integrate the inflectional cue correctly. Moreover, in incorrect responses, IWA show instances of unsuccessful attempts to overcome their

[5]We thank an anonymous reviewer for raising this issue and pointing out the similarities between our results for IWA and findings involving bilingual speakers.

non-deterministic parsing of case-ambiguous sentences, i.e., not only do they show deficits in predicting syntactic structure when morphological cues are ambiguous, but they also tend to avoid parsing decisions when inflectional cues indicate an OVS structure. We speculate that the combination of these two aspects is responsible for their incorrect offline processing of case-ambiguous OVS sentences.

General Discussion

Two visual-world experiments involving IWA and age-matched controls investigated the online processing of unambiguous case-marking cues (Experiment 1) and inflectional number-agreement cues in case-ambiguous German SVO and OVS sentences (Experiment 2). This is the first online study investigating processing of subject-verb agreement cues for sentence comprehension in real time in aphasia. In both experiments, participants' eye-movements were being monitored while they performed an auditory sentence-picture matching task. This allowed us to gain insights into IWA's online sentence processing in cases of successful (correct offline responses) as well as unsuccessful (incorrect offline responses) cue integration. The results provide new insights regarding the source of IWA's sentence comprehension deficits and the morpho-syntactic processing strategies they rely on. Moreover, the findings corroborate previous results on processing of morpho-syntactic cues to theta-role assignment in unimpaired listeners.

For processing of German main clauses in unimpaired listeners, our results are in line with previous studies providing evidence for a subject-first bias, when the sentence-initial NP is case ambiguous. The eye-tracking data further confirm findings of rapid incremental integration of verb-agreement cues signaling non-canonical OVS struture and successful immediate revision of the predicted SVO template. Moreover, in line with previous eye-tracking studies, we found evidence that, in German, unambiguous case markers indicating an OVS structure are processed incrementally for assigning thematic relations in case-unambiguous sentences. However, for both case-ambiguous and unambiguous OVS sentences, we found increased end-of sentence RTs in the offline sentence-picture matching task, supporting the view that OVS sentences are more demanding in terms of overall processing time, despite the incremental nature of morphological cue integration.

Our findings also shed new light on morpho-syntactic processing and the source of sentence comprehension deficits in aphasia. Comparing IWA's accuracy in both experiments reveals that, for both types of morphological cues (case and number-agreement), comprehension is more impaired in OVS as compared to SVO sentences. The finding of such a word order effect on offline responses is in line with previous findings in the field. However, the eye-tracking method combined with the sentence-picture matching task allowed us to gain insights into IWA's *online* processing of case and verb agreement cues. By analyzing IWA's trials with correct and those with incorrect end-of-sentence responses seperately, we were able to characterize aphasic processing mechanisms reflecting correct as well as erroneous morpho-syntactic cue processing.

Although the offline results indicate impaired, i.e., less accurate, processing of these morphological cues, the gaze pattern of IWA's correct offline responses revealed that they are in fact sensitive to agreement markers in ambiguous sentences and to case markers in unambiguous sentences. Yet, they cannot integrate them as incrementally as age-matched controls do in the service of sentence comprehension. Similar delays have been found for the integration of lexical information in sentence processing (Choy & Thompson, 2010; Thompson & Choy, 2009).

When we look at the gaze patterns in the correct responses in both experiments, a striking difference emerges: when an unambiguous case marker appeared in Experiment 1, IWA initially adopted a subject-first prediction, and only later revised their parse. By contrast, in Experiment 2, when IWA were confronted with a case-ambiguous marker (e.g., Das$_{NOM/ACC}$ Kind), in contrast to controls, they did not predict an SVO structure, but rather seemed to rely on a wait-and-see strategy until unambiguous morphological information (the verb inflection) comes in. This absence of predictive processes when only ambiguous morphological information is available is similar to the observation of divergences in parsing predictions based on lexical-semantic cues provided by the meaning of a verb (Mack et al., 2013; Meyer et al., 2012) and converges with findings of less capacities in prediction in L2 comprehenders (DeLong et al., 2005).

We speculate that one possible explanation for the differences we observed between case-ambiguous and unambiguous sentences may be as follows. During sentence processing, IWA rely on a top-down wait-and-see strategy which gets overridden as soon as unambiguous cues are available. Once unambiguous cue information is available, IWA initiate prediction of upcoming syntactic structure. Yet, initially they commit to a subject-first prediction. In the case of cues signaling OVS structure, the predicted structure then needs to be revised towards an object-first parse in order to integrate the cue properly. This account explains the delay we observed in IWA's cue processing.

Under this view, the wait-and-see strategy (as a kind of adaptive process) is responsible for the missing subject-first preference at NP1 in sentences with an ambiguous case cue (Experiment 2). In contrast, in Experiment 1, the presence of the unambiguous case cue at NP1 makes prediction possible to occur early and the wait-and-see strategy is overridden very early. However, in case of an unambiguous accusative cue at NP1 in OVS sentences, IWA suffer from interfering competitive co-activation of the nominative marker because it occurs more frequently in sentence-initial position. Evidence supporting the claim of interfering co-activated competitors comes from the study by Dickey et al. (2007) who found interfering effects of lexical competitors during processing of filler-gap constructions. Our results provide no evidence for such competition in controls. In cases of correct offline comprehension, despite interfering competition, IWA succeed in revising their parse towards an OVS structure and finally integrate the accusative case-marker correctly (but this process is delayed). Future research is needed to investigate in more detail how deficient prediction and interfering co-activation of more frequent cues contribute to sentence comprehension deficits in aphasia.

By contrast, in the case-ambiguous sentences in Experiment 2, in the absence of unambiguous case information, the wait-and-see strategy operates until the subject-verb agreement cue is being processed. As this is unambiguous in nature, it drives prediction of an SVO structure. For initially ambiguous sentences with verb-based disambiguation towards an SVO structure, the predicted structure corresponds with the incoming sentence material and, thus, the increase in gazes to the target is observed already during NP2 and further during silence. However, for OVS sentences (in which the inflectional cue is not in agreement with NP1), IWA's gazes to the target increased slightly later (mostly after NP2) as compared to the SVO condition. This finding could be associated with their commitment to a subject-first prediction evoked by the unambiguous agreement cue (although this cue signals OVS structure). However, slightly later, they directed their gaze towards the picture promoting the OVS interpretation, indicating that the SVO prediction was revised towards an OVS parse and the inflectional cue is integrated correctly–albeit delayed.

The gaze patterns in incorrect OVS responses in both experiments suggest a slow build-up of the incorrect parsing decision (to build an SVO structure) and failure to integrate the case or agreement cue. This suggests that the process which leads IWA to building erroneous parses is similar in both experiments.

With respect to the question of why in the IWA processing of case markers in unambiguous OVS sentences was more impaired than processing of agreement cues in case-ambiguous OVS sentences, we suggest the following explanation. As argued earlier, subject-verb-number-agreement marking is a more reliable cue for assignment of theta-role relations in a sentence as compared to case cues. Thus, it is possible that processing of agreement cues is less affected in aphasia because it is the more reliable cue, and therefore, less prone to impairments. This view is also consistent with data on language acquisition showing that children can process agreement cues earlier than case cues, which indicates that more reliable cues are acquired earlier. However, there are other reasons why processing of agreement cues may be more robust against language breakdown as compared to processing of case marking. For example, it is possible that the subject-verb agreement morpheme is more salient than rather unstressed case morphology cues on determiners and, therefore, it might be easier to perceive. Processing of case cues could also be more affected because, according to Fodor and Inoue (2000), case markers are less robust cues during sentence comprehension as revising an incorrect case assignment has been shown to be easier than revising other parsing decisions. A more principled test of the relative strength of agreement vs. case marking cues would need an experiment involving sentences in which both types of cues are present simultaneously. Such a direct comparison of these cues warrants future research.

In summary, there are four main findings regarding morpho-syntactic processing in IWA. First, there is evidence for a subject-first bias for processing of sentences containing *unambiguous* morphological cues, which is not the case in controls. Second, while controls pursue an early subject-first prediction for processing of *ambiguous* sentences, IWA do not, but the build-up of the subject-first parse is delayed as compared to controls. Thus, in the absence of unambiguous morphological cues, IWA show deficits in predicting upcoming sentence structure. Third, for both case-ambiguous and case-unambiguous sentences, we identified marked deficits in integrating case cues on time. Finally, overall, processing of agreement cues in case-ambiguous OVS sentences is less affected than processing of case markers in case-unambiguous sentences. We attribute this finding to the higher cue reliability of agreement cues, which renders them more resistant against impairments in aphasia.

References

Adobe Systems Incorporation. (2008). *Adobe Flash CS4 Professional [Computer program].*

Bader, M., & Meng, M. (1999). Subject-Object Ambiguities in German Embedded Clauses: An Across-the-Board Comparison. *Journal of Psycholinguistic Research*, *28*(2), 121–143.

Bates, E., Friederici, A., & Wulfeck, B. (1987). Comprehension in aphasia: A cross-linguistic study. *Brain and language*, *32*(1), 19–67.

Baum, S. R., & Pell, M. D. (1999). The neural bases of prosody: Insights from lesion studies and neuroimaging. *Aphasiology*, *13*(8), 581–608.

beim Graben, P., Saddy, J. D., Schlesewsky, M., & Kurths, J. (2000). Symbolic dynamics of event-related brain potentials. *Physical review. E, Statistical physics, plasmas, fluids, and related interdisciplinary topics*, *62*(4 Pt B), 5518–41.

Berndt, R. S., Mitchum, C. C., & Haendiges, A. N. (1996). Comprehension of reversible sentences in "agrammatism": A meta-analysis. *Cognition*, *58*(3), 289–308.

Boersma, P., & Weenink, D. (2009). *Praat: Doing phonetics by computer [Computer program].*

Boland, J. E. (2005, April). Visual arguments. *Cognition*, *95*(3), 237–274.

Bornkessel, I., McElree, B., Schlesewsky, M., & Friederici, A. D. (2004). Multi-dimensional contributions to garden path strength: Dissociating phrase structure from case marking. *Journal of Memory and Language*, *51*(4), 495–522.

Bornkessel, I., Schlesewsky, M., & Friederici, A. D. (2002). Grammar overrides frequency: Evidence from the online processing of flexible word order. *Cognition*, *85*(2), B21–B30.

Bornkessel, I., Schlesewsky, M., & Friederici, A. D. (2003). Eliciting thematic reanalysis effects: The role of syntax-independent information during parsing. *Language and Cognitive Processes*, *18*(3), 269–298.

Bornkessel, I., Zysset, S., Friederici, A. D., von Cramon, D. Y., & Schlesewsky, M.

(2005). Who did what to whom? The neural basis of argument hierarchies during language comprehension. *NeuroImage*, *26*(1), 221–33.

Burchert, F., De Bleser, R., & Sonntag, K. (2003). Does morphology make the difference? Agrammatic sentence comprehension in German. *Brain and Language*, *87*(2), 323–342.

Burchert, F., Lorenz, A., Schröder, A., De Bleser, R., & Stadie, N. (2011). *Sätze verstehen - Neurolinguistische Materialien für die Untersuchung von syntaktischen Störungen beim Satzverständnis.* Hofheim: NAT-Verlag.

Burkhardt, P., Avrutin, S., Piñango, M. M., & Ruigendijk, E. (2008). Slower-than-normal syntactic processing in agrammatic Broca's aphasia: Evidence from Dutch. *Journal of Neurolinguistics*, *21*, 120–137.

Caplan, D. (2006). Aphasic deficits in syntactic processing. *Cortex*, *42*(6), 797–804.

Caplan, D., Baker, C., & Dehaut, F. (1985). Syntactic determinants of sentence comprehension in aphasia. *Cognition*, *21*(2), 117–175.

Caplan, D., Waters, G. S., DeDe, G., Michaud, J., & Reddy, A. (2007). A study of syntactic processing in aphasia I: Behavioral (psycholinguistic) aspects. *Brain and Language*, *101*(2), 103–150.

Caramazza, A., Capitani, E., Rey, A., & Berndt, R. S. (2001). Agrammatic Broca's aphasia is not associated with a single pattern of comprehension performance. *Brain and Language*, *76*(2), 158–184.

Caramazza, A., & Miceli, G. (1991). Selective Impairment of Thematic Role Assignment in Sentence Processing. *Brain and Language*, *41*(3), 402–436.

Caramazza, A., & Zurif, E. B. (1976). Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and Language*, *3*(4), 572–582.

Cho-Reyes, S., & Thompson, C. K. (2012). Verb and sentence production and comprehension in aphasia: Northwestern Assessment of Verbs and Sentences (NAVS). *Aphasiology*, *26*(10), 1250–1277.

Choy, J. J., & Thompson, C. K. (2010). Binding in agrammatic aphasia: Processing to comprehension. *Aphasiology*, *24*(5), 551–579.

Cooper, R. M. (1974). Control of of Eye Spoken Fixation by the Meaning Language: A New Methodology for the Real-Time Investigation of Speech Perception, Memory, and Language Processing. *Cognitive Psychology*, *6*, 84–107.

De Bleser, R., Cholewa, J., Stadie, N., & Tabatabaie, S. (2004). *LeMo - Lexikon modellorientiert: Einzelfalldiagnostik bei Aphasie, Dyslexie und Dysgraphie.* München: Urban & Fischer.

De Bleser, R., Dronsek, C., & Bayer, J. (1988). Morphosyntactic processing in German agrammatism: A replication and revision of von Stockert/Bader (1976). *Cortex*, *24*, 53–76.

DeLong, K. a., Urbach, T. P., & Kutas, M. (2005, August). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature neuroscience*, *8*(8), 1117–21.

Devescovi, A., Bates, E., D'amico, S., Hernandez, A., Marangolo, P., Pizzamiglio, L., et al. (1997). An on-line study of grammaticality judgements in normal and aphasic speakers of Italian. *Aphasiology*, *11*(6), 543–579.

Dickey, M. W., Choy, J. J., & Thompson, C. K. (2007). Real-time comprehension of wh- movement in aphasia: Evidence from eyetracking while listening. *Brain and Language*, *100*(1), 1–22.

Dronkers, N. F., Wilkins, D. P., Van Valin, R. D., Redfern, B. B., & Jaeger, J. J. (2004). Lesion analysis of the brain areas involved in language comprehension. *Cognition*, *92*(1-2), 145–177.

Fodor, J. D., & Inoue, A. (2000). Garden Path Repair: Diagnosis and Triage. *Language and Speech*, *43*(3), 261–271.

Friederici, A. D., Wessels, J., Emmorey, K., & Bellugi, U. (1992). Sensitivity to Inflectional Morphology in Aphasia: A Real-Time Processing Perspective. *Brain and Language*, *43*, 747–763.

Friedmann, N., Reznick, J., Dolinski-Nuger, D., & Soboleva, K. (2010). Comprehension and production of movement-derived sentences by Russian speakers with agrammatic aphasia. *Journal of Neurolinguistics*, *23*, 44–65.

Friedmann, N., & Shapiro, L. P. (2003). Agrammatic Comprehension of Simple Active Sentences With Moved Constituents: Hebrew OSV and OVS Structures. *Journal of Speech, Language, and Hearing Research*, *46*(2), 288–297.

Frisch, S., Schlesewsky, M., Saddy, D., & Alpermann, A. (2002). The P600 as an indicator of syntactic ambiguity. *Cognition*, *85*, B83—B92.

Gibson, E. (2000). Dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita, & W. O'Neil (Eds.), *Image, language, brain: Papers from the first mind articulation project symposium.* Cambridge, MA: MIT Press.

Gorrell, P. (2000). The Subject-before-Object Preference in German Clauses. In B. Hemforth & L. Konieczny (Eds.), *German sentence processing* (p. 25–63). Dordrecht: Kluwer Academics Publishers.

Grodzinsky, Y. (2000). The neurology of syntax: Language use without Broca's area. *Behavioral and Brain Sciences*, *23*(1), 1–71.

Hanne, S., Sekerina, I. A., Vasishth, S., Burchert, F., & De Bleser, R. (2011). Chance in agrammatic sentence comprehension: What does it really mean? Evidence from eye movements of German agrammatic aphasic patients. *Aphasiology*, *25*(2), 221–244.

Heeschen, C. (1980). Strategies of decoding actor-object-relations by aphasic patients. *Cortex*, *16*(1), 5–19.

Heister, J., Würzner, K.-M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., et al. (2011). dlexDB – eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychologische Rundschau*, *62*(1), 10–20.

Hemforth, B., & Konieczny, L. (2000). Cognitive Parsing in German - An introduction. In B. Hemforth & L. Konieczny (Eds.), *German sentence processing* (pp. 1–24).

Dordrecht: Kluwer Academics Publishers.

Henderson, J. M., & Ferreira, F. (2004). Scene perception for psycholinguists. In J. M. Henderson & F. Ferreira (Eds.), *The interface of language, vision, and actions: Eye movements and the visual world* (pp. 1–58). Hove, UK: Psychology Press.

Huber, W., Poeck, K., Weniger, D., & Willmes, K. (1983). *Aachener Aphasie Test (AAT)*. Göttingen: Hogrefe.

Kamide, Y., Scheepers, C., & Altmann, G. T. (2003). Integration of syntactic and semantic information in predictive processing: Cross-linguistic evidence from German and English. *Journal of Psycholinguistic Research*, *32*(1), 37–55.

Kay, J., Coltheart, M., & Lesser, R. (1992). *PALPA - Psycholinguistic Assessments of Language Processing in Aphasia*. Hove: Psychology Press.

Kimelman, M. D. (1999). Prosody, linguistic demands, and auditory comprehension in aphasia. *Brain and Language*, *69*(2), 212–221.

Kljajevic, V., & Murasugi, K. (2010). The role of morphology in the comprehension of wh -dependencies in Croatian aphasic speakers. *Aphasiology*, *24*(11), 1354–1376.

Knoeferle, P. (2007). Comparing the time course of processing initially ambiguous and unambiguous German SVO / OVS sentences in depicted events. In R. van Gompel, M. Fischer, W. Murray, & R. Hill (Eds.), *Eye-movement research* (pp. 517–536). Oxford: Elsevier.

Knoeferle, P., Crocker, M. W., Scheepers, C., & Pickering, M. J. (2005). The influence of the immediate visual context on incremental thematic role-assignment: Evidence from eye-movements in depicted events. *Cognition*, *95*(1), 95–127.

Knoeferle, P., Habets, B., Crocker, M. W., & Münte, T. F. (2008). Visual scenes trigger immediate syntactic reanalysis: Evidence from ERPs during situated spoken comprehension. *Cerebral Cortex*, *18*(4), 789–795.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*, 1126–1177.

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

Lukatela, K., Crain, S., & Shankweiler, D. (1988). Sensitivity to inflectional morphology in agrammatism: Investigation of a highly inflected language. *Brain and Language*, *33*(1), 1–15.

Luzzatti, C., Toraldo, A., Guasti, M. T., Ghirardi, G., Lorenzi, L., & Guarnaschelli, C. (2011). Comprehension of reversible active and passive sentences in agrammatism. *Aphasiology*, *15*(5), 419–441.

Mack, J. E., Ji, W., & Thompson, C. K. (2013). Effects of verb meaning on lexical integration in agrammatic aphasia: Evidence from eyetracking. *Journal of Neurolinguistics*, *26*(6), 619–636.

MacWhinney, B., Osmán-Sági, J., & Slobin, D. I. (1991). Sentence comprehension in aphasia in two clear case-marking languages. *Brain and Language*, *41*(2), 234–49.

Martin, C. D., Thierry, G., Kuipers, J.-R., Boutonnet, B., Foucart, A., & Costa, A. (2013). Bilinguals reading in their second language do not predict upcoming words as native readers do. *Journal of Memory and Language*, *69*(4), 574–588.

Matzke, M., Mai, H., Nager, W., Rüsseler, J., & Münte, T. (2002). The costs of freedom: An ERP – study of non-canonical sentences. *Clinical Neurophysiology*, *113*(6), 844–52.

Mecklinger, A., Schriefers, H., Steinhauer, K., & Friederici, A. D. (1995). Processing relative clauses varying on syntactic and semantic dimensions: An analysis with event-related potentials. *Memory & Cognition*, *23*(4), 477–94.

Meng, M., & Bader, M. (2000). Mode of Disambiguation and Garden-Path Strength: An Investigation of Subject-Object Ambiguities in German. *Language and Speech*, *43*(1), 43–74.

Meyer, A. M., Mack, J. E., & Thompson, C. K. (2012). Tracking passive sentence comprehension in agrammatic aphasia. *Journal of Neurolinguistics*, *25*(1), 31–43.

Milekić, S., Bosković, Z., Crain, S., & Shankweiler, D. (1995). Comprehension of nonlexical categories in agrammatism. *Journal of Psycholinguistic Research*, *24*(4), 299–311.

Mitchum, C. C., & Berndt, R. S. (2008). Comprehension and Production of Sentences. In R. Chapey (Ed.), *Language intervention strategies in aphasia and related neurogenic communication disorders* (5th ed., pp. 632–653). Philadelphia: Wolters Kluwer.

Nicol, J. L., Jakubowicz, C., & Goldblum, M.-C. (1996). Sensitivity to grammatical marking in English-speaking and French-speaking non-fluent aphasics. *Aphasiology*, *10*(6), 593–622.

Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113.

Schlesewsky, M., & Bornkessel, I. (2004). On incremental interpretation: Degrees of meaning accessed during sentence comprehension. *Lingua*, *114*(9-10), 1213–1234.

Schlesewsky, M., Fanselow, G., Kliegl, R., & Krems, J. (2000). The subject preference in processing of locally ambiguous wh-questions in German. In B. Hemforth & L. Konieczny (Eds.), *German sentence processing* (pp. 65–94). Dordrecht: Kluwer Academics Publishers.

Schriefers, H., Friederici, A., & Kuhn, K. (1995). The Processing of Locally Ambiguous Relative Clauses in German. *Journal of Memory and Language*, *34*(4), 499–520.

Smith, S. D., & Bates, E. (1987). Accessibility of Case and Gender Contrasts for Agent-Object Assignment in Broca ' s Aphasics and Fluent Anomies. *Brain and Language*, *30*, 8–32.

Smith, S. D., & Mimica, I. (1984). Agrammatism in a case-inflected language: Comprehension of agent-object relations. *Brain and Language*, *21*(2), 274–90.

Swoboda-Moll, M., Burchert, F., & De Bleser, R. (2002). Agrammatic Comprehension of agent-object-relaions: A German replication of Smith and Mimica (1984). *Cortex*, *38*(5), 908–910.

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*(5217), 1632–1634.

Thompson, C. K., & Choy, J. J. (2009). Pronominal Resolution and Gap Filling in Agrammatic Aphasia: Evidence from Eye Movements. *Journal of Psycholinguistic Research*, *38*(3), 255–283.

Wulfeck, B., & Bates, E. (1991). Differential Sensitivity to Errors of Agreement and Word Order in Broca's Aphasia. *Journal of Cognitive Neuroscience*, *3*(3), 258–272.

Yarbay Duman, T., Altinok, N., Özgirgin, N., & Bastiaanse, R. (2011). Sentence comprehension in Turkish Broca's aphasia: An integration problem. *Aphasiology*, *25*(8), 908–926.

List of Figures.

Figure 1: Sample pictures for the target example in Experiment 1 given in (4-a) (*Der Arzt schubst den Dieb 'The doctor is pushing the thief'*). (a) matching picture. (b) foil picture.

Figure 2: Sample pictures for two filler items in Experiment 1. (a) Matching picture and foil (object distractor) for the filler *Der Schmied zieht den Sack ('The smith is pulling the sack').* (b) Matching picture and foil (subject distractor) for the filler *Der Wald wird vom Clown gefilmt ('The wood is being filmed by the clown').*

Figure 3: Controls' and IWA's gaze proportions to the target picture in Experiment 1. For IWA, data is separated for correct and incorrect offline response accuracy (a correct, b incorrect). For controls, only gaze data of correct responses is shown (as this constituted most of their data due to ceiling effects).

Figure 4: Sample pictures for the target example in Experiment 2 given in (5-a) (*Das Kind faengt die Frauen. 'The child is catching the women'*). (a) matching picture. (b) foil picture.

Figure 5: Sample pictures for two filler items in Experiment 2. (a) Matching picture and foil (object distractor) for the filler *Das Paket wird von den Frauen geschoben ('The bundle is being pushed by the women').* (b) Matching picture and foil (subject distractor) for the filler *Das Heu wird von den Ziegen getreten ('The hay is being kicked by the goats').*

Figure 6: Controls' and IWA's gaze proportions to the target picture in Experiment 2. For IWA, data is separated for correct and incorrect offline response accuracy (a correct, b incorrect). For controls, only gaze data of correct responses is shown (as this constituted most of their data due to ceiling effects).

Table 1

Demographic and neurological data of participants with aphasia.

Patient	Gender	Age	Years post-onset	Etiology	Localization
A01	M	60	18	cerebral hemorrhage	right
A02	F	39	5	cerebral infarction	left
A03	F	71	11	cerebral infarction	left
A04	M	72	15	cerebral infarction	left
A05	F	41	12	cerebral infarction	left
A06	M	59	8	cerebral infarction	left
A07	M	46	12	cerebral infarction	left
A08	M	54	8	cerebral infarction	left

Table 2

Participants with aphasia: Results of pre-tests.

	Aachen Aphasia Test		LeMo (% correct)		Edinburgh Handedness Inventory	
Patient	Aphasia classification	Severity (Standard nine)	Auditory Discrimination	Auditory Word-Picture-Matching	Score	Interpretation
A01	Broca	5.0	94	95	-100	left
A02	Broca	5.4	100	95	+69	right
A03	Broca	4.0	99	100	+60	right
A04	Broca	4.8	97	100	+54	right
A05	Anomic	6.6	99	100	+89	right
A06	Anomic	5.6	96	95	+88	right
A07	Anomic	5.6	94	95	+100	right
A08	Wernicke	5.2	97	100	+79	right

Table 3

Participants with aphasia: Results on comprehension of semantically irreversible sentences and semantically reversible case-marked and number-marked active sentences in the pre-test.

	Sentence comprehension (% Correct)				
Patient	Irreversible (n = 22)	Reversible SVO Case (n = 20)	Reversible SVO Number (n = 20)	Reversible OVS Case (n = 20)	Reversible OVS Number (n = 20)
A01	100	95	85	65	90
A02	86	90	90	85	80
A03	86	85	75	40	75
A04	100	95	100	60	60
A05	100	100	100	75	80
A06	91	90	85	70	40
A07	100	90	90	35	45
A08	96	95	80	40	30
Mean	95	93	88	59	62.5

Table 4

Mean constituent durations (in milliseconds) for auditory sentences in Experiment 1.

Condition	NP1	Verb	NP2
	M (*SD*)	*M* (*SD*)	*M* (*SD*)
SVO	716 (130)	408 (83)	781 (76)
OVS	731 (121)	411 (91)	767 (86)

Table 5

Participants with aphasia: Single-case accuracy in Experiment 1.

	Accuracy (% correct)		Individual Word-order Effect	
Patient	SVO	OVS	χ^2	p-value
A01	80	30	10.1	.001
A02	70	60	0.44	.254
A03	60	55	0.1	.375
A04	80	50	3.96	.023
A05	95	70	4.33	.019
A06	70	50	1.67	.093
A07	70	35	4.91	.013
A08	90	20	19.8	.000

Table 6

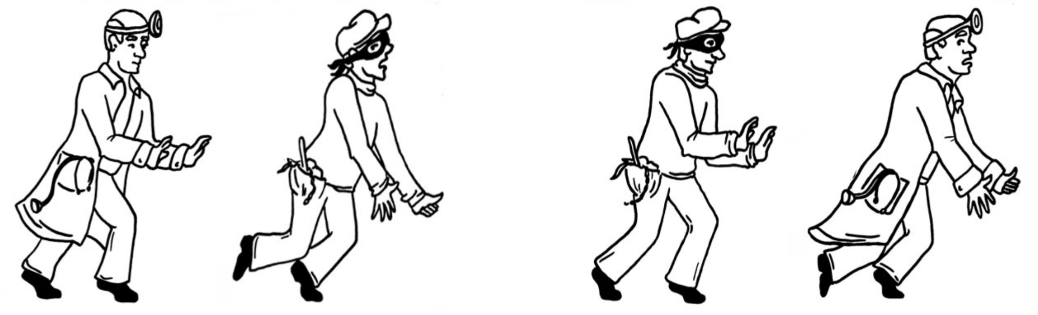
Mean constituent durations for auditory sentences in Experiment 2, in ms.

Condition	NP1	Verb	NP2
	M (*SD*)	*M* (*SD*)	*M* (*SD*)
SVO	730 (124)	446 (91)	793 (55)
OVS	745 (135)	461 (86)	803 (63)

Table 7

Participants with aphasia: Single-case accuracy in Experiment 2.

	Accuracy (% correct)		Individual Word-order Effect	
Patient	SVO	OVS	χ^2	p-value
A01	85	75	0.625	.215
A02	80	80	0.0	.5
A03	60	70	0.44	.254
A04	80	65	1.129	.144
A05	95	75	3.137	.038
A06	65	70	0.114	.368
A07	75	50	2.667	.051
A08	85	25	14.545	.001

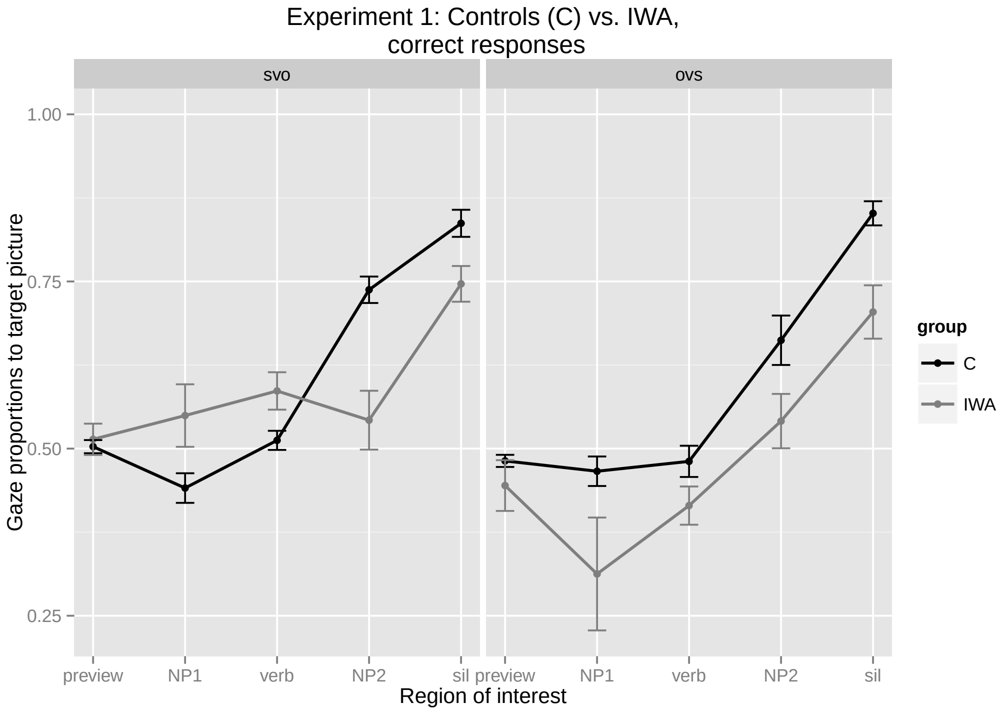


(a) (b)

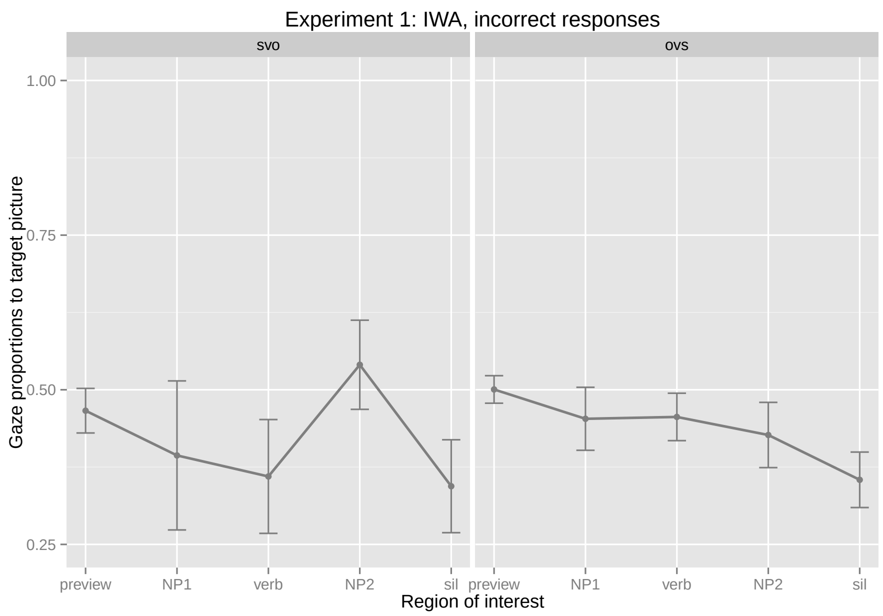
Figure 1

(a) (b)

Figure 2


Experiment 1: Controls (C) vs. IWA,
correct responses
svo
ovs
1.00
0.75
0.50
0.25
Gaze proportions to target picture
preview
NP1
verb
NP2
sil
Region of interest
group
C
IWA

(a)


Experiment 1: IWA, incorrect responses
svo
ovs
1.00
0.75
0.50
0.25
Gaze proportions to target picture
preview
NP1
verb
NP2
sil
Region of interest

(b)

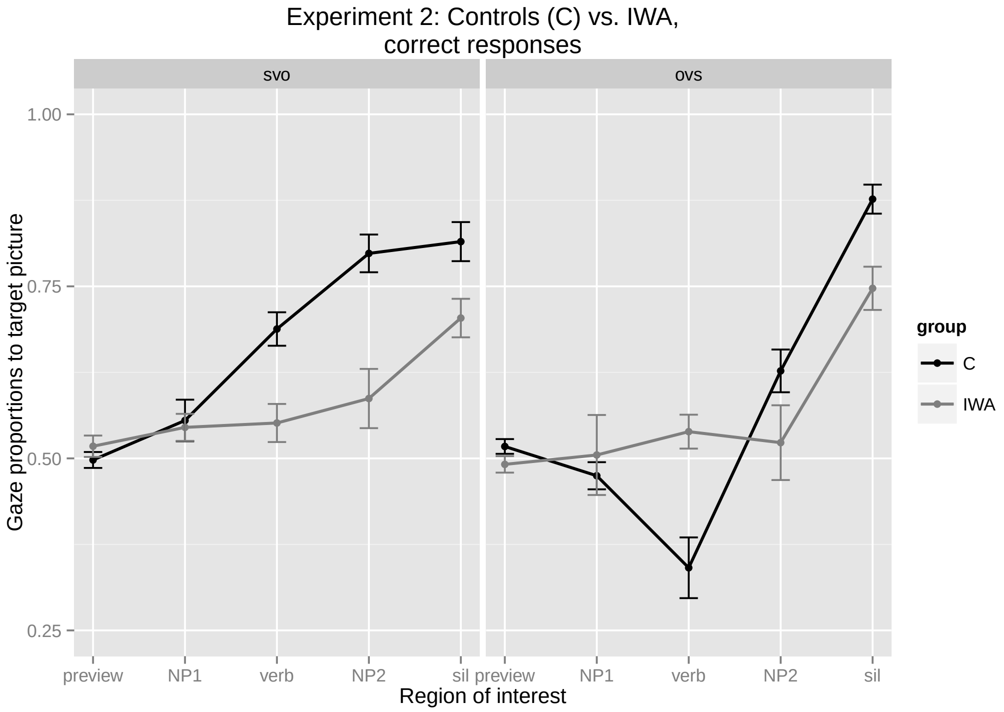
Figure 3

(a) (b)

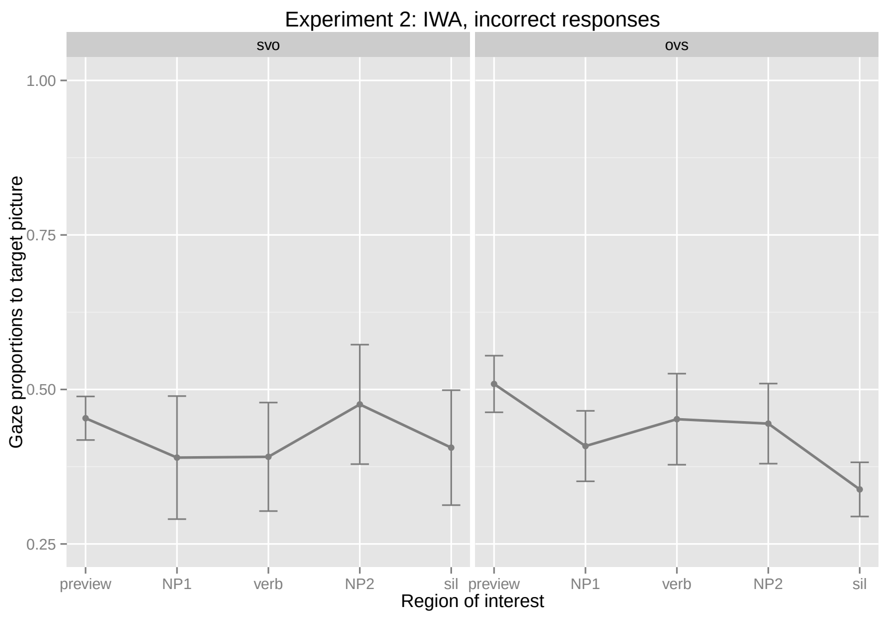
Figure 4

(a) (b)

Figure 5


Experiment 2: Controls (C) vs. IWA,
correct responses
svo
ovs
Gaze proportions to target picture
1.00
0.75
0.50
0.25
preview
NP1
verb
NP2
sil
Region of interest
group
C
IWA

(a)


Experiment 2: IWA, incorrect responses
svo
ovs
Gaze proportions to target picture
1.00
0.75
0.50
0.25
preview
NP1
verb
NP2
sil
Region of interest

(b)

Figure 6

Electronic supplementary material to: Sentence comprehension in aphasia: Eye-tracking reveals delayed morphological cue integration and late parsing commitments

Sandra Hanne

Department of Linguistics, University of Potsdam, Germany

Frank Burchert

Department of Linguistics, University of Potsdam, Germany

Ria De Bleser

Department of Linguistics, University of Potsdam, Germany

Shravan Vasishth

Department of Linguistics, University of Potsdam, Germany

Electronic supplementary material to: Sentence comprehension in aphasia: Eye-tracking reveals delayed morphological cue integration and late parsing commitments

Rating Studies for Experiment 1

Sentence Rating Study

In order to construct the experimental sentences for the sentence rating study, we selected 26 transitive German verbs describing simple depictable events. For each verb, we designed three sentences with different NPs taking the role of agent or theme, respectively. At this point, the NPs were chosen based on intuitive judgement of equal semantic plausibility. In addition, we selected additional verbs and constructed sentences, in which one of the two nouns had a much higher probability of being the agent (e.g. The doctor vaccinates the man). These sentences served as fillers in the rating study.

Overall, 86 sentences entered the rating study and they were arranged in two lists, so that in the second list subject and object for each sentence were reversed. The sentences of each list were pseudo-randomized in four different orders in order to control for familiarization and fatigue effects during the rating. The task was to rate the plausibility of the subject in a given sentence on a 5-point scale. Participants were instructed to judge how reasonable it is that the person mentioned in the sentence is actually doing the action, i.e., is the agent of the action. The scores of the scale were verbalized as follows: (1)–not plausible at all, (2)–not plausible, (3)–more plausible, (4)–plausible, (5)–very plausible. Two example sentences were provided to clarify the task.

Altogether, 42 undergraduate students (13 male, 29 female, mean age: 23 years, range: 19-42 years) participated in the rating study receiving one of the four different versions of either list one or two. Participants signed informed consent about the study and were given course credit. For each sentence, we calculated the mean rating across all participants and compared it to the mean rating of the respective sentence in list two using an unpaired t-test. We then identified items for which the t-test was significant because this would mean that one of the nouns had a bias towards being the subject of the sentence. This led to 35 items being excluded from further analyses and two verbs (*to vaccinate, to baptise*) were no longer present in the target items. We then applied an additional criterion to the remaining sentences and excluded sentences with a mean rating of less than 3 on the 5-point-scale. Finally, 23 sentences each containing a different verb and specified NPs passed this criterion and entered the final item set.

Rating of Sentence Recordings

In order to rule out any effects of the NP1 manipulation on processing of the auditory sentences during experiment 1, a rating study was carried out with 10 participants (all female, age range: 21 - 26 years). The raters heard 20 pairs of recorded sentences and were asked to decide whether they had heard the same two recordings or whether there was a difference between the two. For the 10 recordings containing a manipulation, the original file and the manipulated file were played (in randomized order). For the other 10 pairs, the same recording was played twice so that these served as fillers. Participants performed significantly above chance in rating the filler pairs (i.e., the same pairs) correctly ($M = 68$, $\chi^2(1) = 5.97$, $p < .05$, 2-tailed). For the pairs containing manipulated (i.e., different) sounds, performance was not different from chance ($M = 44$, $\chi^2(1) = 0.5$, $p > .05$, two-tailed). Moreover, same pairs were recognized correctly as being same significantly more often than different pairs were recognized as being different ($t(9) = 2.84$, $p < .05$, paired, two-tailed). Therefore, we assume that the manipulations of the recorded sentences would not be recognized by participants in the experiments and that having elongated some of the recordings would not affect the goal of the experiment.

Rating Studies for Experiment 2

Sentence Rating Study

Overall, 83 sentences entered the rating study for sentences to used in Experiment 2. These comprised three different sentences for each of 25 transitive German verbs (most of which had also been selected for the sentence set for Experiment 1) and filler sentences constructed around additional verbs. In the fillers, one of the nouns had a bias to be the agent (e.g., The nuns baptize the child.). Analogous to the rating for the other experiment, two lists were created, whereby in the second list subject and object were reversed for each sentence. For each list, four pseudo-randomizations existed. Sentences were rated by 40 participants (8 male, 32 female, mean age: 27, range: 19-41), all of which signed informed consent. Most of them were undergraduate students who received course credit. The task, rating scale, general procedure and analysis was identical to the other rating study. We identified 34 sentences for which the mean rating in list one was significantly different from that in list two and excluded these in the first step. By applying the additional criterion (excluding sentences with a mean rating of less than 3 on the 5-point-scale), we identified 23 sentences with different verbs that entered the final item set.

Rating of Sentence Recordings

The rating study for sentence recordings of Experiment 2 was carried out with 10 participants. As for the rating for Experiment 1, the raters were asked to decide for pairs of recorded sentences whether they had heard the same two recordings or whether they had heard any difference between the two. For 20 pairs, the original and the manipulated file were played. The remaining pairs consisted of two identical recordings and thus served as fillers. Participants performed significantly above chance in rating the fillers correctly ($M = 154$, $\chi^2(1) = 30.3$, $p < .05$, 2-tailed), whereas for the pairs of manipulated sounds performance was below chance ($M = 72$, $\chi^2(1) = 7.44$, $p < .05$, two-tailed). A direct comparison of correct identifications of same and different pairs revealed that same pairs were recognized correctly significantly more often than different pairs ($t(19) = 9.02$, $p < .05$, paired, two-tailed). Thus, we concluded that the manipulation of the verb duration would not be perceived by the participants of the experiment

Model parameters for results of Experiments 1 and 2

Table 1

Model parameters for accuracy measures in Experiment 1.

Effect	Model comparisons			contrast	Model values			
	χ^2	df	p-value		b	SE	z	p-value
Condition	**34.92**	**4**	**<.05**	OVSvs.SVO	**-0.637**	**0.34**	**-1.86**	**=.06**
Group	**33.36**	**5**	**<.05**	IWAvs.C	**-2.193**	**0.41**	**-5.288**	**<.05**
Condition:Group	**3.73**	**6**	**<.05**	Cond:IWAvs.C	**-0.825**	**0.42**	**-1.95**	**=.052**

Table 2

Model parameters for reaction time measures in Experiment 1.

Effect	Model comparisons			contrast	Model values		
	χ^2	df	p-value		b	SE	t
Condition	**31.74**	**5**	$<$**.05**	OVSvs.SVO	**421.1**	**72.9**	**5.77**
Group	**27.01**	**6**	$<$**.05**	IWAvs.C	**1762.9**	**257.6**	**6.84**
Condition:Group	1.88	7	$>$ .05	Cond:IWAvs.C	-229.8	167.7	-1.37

Table 3

Model parameters for the effect of response accuracy on IWA's RTs in Experiment 1.

Effect	Model comparisons			Model values		
	χ^2	df	p-value	b	SE	t
Accuracy	**11.77**	**5**	**<.05**	**-680.2**	**195.7**	**-3.48**
Condition	1.32	6	> .05	226.5	197.0	1.15
Condition:Accuracy	0.33	7	> .05	-242.9	418.1	-0.58

Table 4

Model parameters for comparing IWA's gaze proportions to the target picture in correct trials against controls in Experiment 1.

Effect	Model comparisons			contrast	Model values		
	χ^2	df	p-value		b	SE	t
Group	**5.28**	**5**	**<.05**	IWAvs.C	-0.017	0.02	-0.86
Condition	**9.13**	**6**	**<.05**	OVSvs.SVO	-0.018	0.01	-1.5
Roi	**562.98**	**11**	**<.05**	NP1	**-0.063**	**0.03**	**-2.33**
				Verb	**0.073**	**0.03**	**2.69**
				NP2	**0.227**	**0.03**	**8.67**
				Sil	**0.095**	**0.03**	**3.42**
Group:RoI:Condition	**50.13**	**23**	**<.05**	Group:NP1	0.1	0.05	1.84
				Group:Verb	-0.041	0.05	-0.74
				Group:NP2	**-0.263**	**0.05**	**-4.95**
				Group:Sil	**0.115**	**0.05**	**2.15**
				Group:NP1:Cond	**-0.219**	**0.09**	**-2.5**
				Group:Verb:Cond	0.119	0.09	1.34
				Group:NP2:Cond	**0.208**	**0.08**	**2.44**
				Group:Sil:Cond	-0.124	0.09	-1.44

Table 5

Model parameters for IWA's gaze proportions in incorrect trials in Experiment 1. Note: We do not provide coefficients for the three-way-interaction as the respective model could not significantly account for any variance and all model coefficients were non-significant.

Effect	Model comparisons			contrast	Model values		
	χ^2	df	p-value		b	SE	t
Group	**38.33**	**5**	**<.05**	IWAvs.C	**-0.16**	**0.02**	**-7.82**
Condition	2.38	6	> .05	OVSvs.SVO	-0.02	0.01	-1.54
Roi	**392.43**	**11**	**<.05**	NP1	**-0.039**	**0.02**	**-2.03**
				Verb	**0.042**	**0.02**	**2.19**
				NP2	**0.209**	**0.02**	**11.18**
				Sil	**0.136**	**0.02**	**6.91**
Group:Roi	**147.65**	**13**	**<.05**	NP1	-0.016	0.05	-0.32
				Group:Verb	-0.045	0.05	-0.88
				Group:NP2	**-0.221**	**0.05**	**-4.4**
				Group:Sil	**-0.234**	**0.05**	**-4.66**
Group:RoI:Condition	12.57	23	> .05				

Table 6

Model parameters for separate model of only IWA's gaze proportions in incorrect trials in Experiment 1. Note: We do not provide coefficients for the Condition:RoI interaction as the respective model could not significantly account for any variance and all model coefficients were non-significant.

Effect	Model comparisons			contrast	Model values		
	χ^2	df	p-value		b	SE	t
Condition	0.02	9	. > 05	OVSvs.SVO	0.005	0.04	0.13
Roi	**12.64**	**8**	**<.05**	NP1	-0.056	0.05	-1.12
				Verb	-0.003	0.05	-0.06
				NP2	-0.012	0.05	-0.24
				Sil	**-0.0097**	**0.05**	**-2.02**
Condition:RoI	1.17	13	> .05				

Table 7

Model parameters for accuracy measures in Experiment 2.

Effect	Model comparisons			contrast	Model values			
	χ^2	df	p-value		b	SE	z	p-value
Condition	**28.45**	**4**	**<.05**	OVSvs.SVO	**-1.815**	**0.43**	**-4.2**	**<.05**
Group	**25.76**	**5**	**<.05**	IWAvs.C2	**-2.909**	**0.51**	**-5.68**	**<.05**
Condition:Group	**5.23**	**6**	**<.05**	Cond:IWAvs.C2	**1.07**	**0.5**	**2.14**	**<.05**

Table 8

Model parameters for reaction time measures in Experiment 2.

Effect	Model comparisons			contrast	Model values		
	χ^2	df	p-value		b	SE	t
Condition	**77.29**	**5**	<**.05**	OVSvs.SVO	**575.47**	**63.83**	**9.02**
Group	**20.75**	**6**	<**.05**	IWAvs.C	**1540.21**	**259.8**	**5.93**
Condition:Group	**5.22**	**7**	<**.05**	Cond:IWAvs.C	**-307.81**	**134.52**	**-2.29**

Table 9

Model parameters for the effect of response accuracy on IWA's RTs in Experiment 2.

Effect	Model comparisons			Model values		
	χ^2	df	p-value	b	SE	t
Accuracy	**7.63**	**5**	**<.05**	**-454.5**	**163.4**	**-2.78**
Condition	0.23	6	> .05	69.46	145.97	0.48
Condition:Accuracy	3.67	7	> .05	646.8	334.9	1.93

Table 10

Model parameters for comparison of controls' and IWA's gaze proportions to the target picture in correct trials in Experiment 2.

Effect	Model comparisons			contrast	Model values		
	χ^2	df	p-value		b	SE	t
Group	**5.43**	**5**	**<.05**	IWAvs.C	**-0.09**	**0.02**	**-5.01**
Condition	**59.57**	**6**	**<.05**	OVSvs.SVO	**-0.109**	**0.01**	**-9.17**
Roi	**487.15**	**11**	**<.05**	NP1	**0.05**	**0.02**	**2.04**
				Verb	**0.145**	**0.02**	**5.9**
				NP2	**0.111**	**0.02**	**4.5**
				Sil	0.013	0.03	0.43
Group:Condition:RoI	**237.85**	**23**	**<.05**	Cond:NP1	**-0.098**	**0.04**	**-2.76**
				Cond:Verb	**-0.306**	**0.04**	**-8.6**
				Cond:NP2	**0.219**	**0.04**	**6.17**
				Cond:Sil	**0.228**	**0.04**	**5.66**
				Group:Cond:NP1	0.087	0.07	1.2
				Group:Cond:Verb	**0.323**	**0.07**	**4.42**
				Group:Cond:NP2	**-0.29**	**0.07**	**-3.95**
				Group:Cond:Sil	0-0.102	0.07	-1.36

Table 11

Model parameters for comparing IWA's gaze proportions to the target picture in incorrect trials against controls in Experiment 2.

Effect	Model comparisons			contrast	Model values		
	χ^2	df	p-value		b	SE	t
Group	**57.25**	**5**	**<.05**	IWAvs.C	**-0.246**	**0.03**	**-8.83**
Condition	**64.6**	**6**	**<.05**	OVSvs.SVO	**-0.109**	**0.01**	**-9.33**
Roi	**378.64**	**11**	**<.05**	**NP1**	**0.05**	**0.02**	**2.07**
				Verb	**0.146**	**0.02**	**6.0**
				NP2	**0.11**	**0.03**	**4.57**
				Sil	0.012	0.02	0.45
Group:Condition:RoI	**313.77**	**23**	**<.05**	Group:Cond:NP1	0.022	0.11	0.2
				Group:Cond:Verb	**0.32**	**0.11**	**2.99**
				Group:Cond:NP2	-0.206	0.11	-1.9
				Group:Cond:Sil	**-0.225**	**0.11**	**-2.05**

Table 12

Model parameters for separate model of IWA's gaze proportions in incorrect trials in the SVO condition in Experiment 2.

Effect	Model comparisons			contrast	Model values		
	χ^2	df	p-value		b	SE	t
Group	**33.0**	**5**	**<.05**	IWAvs.C	**-0.244**	**0.03**	**-7.28**
Roi	**202.99**	**9**	**<.05**	NP1	**0.049**	**0.02**	**2.06**
				Verb	**0.146**	**0.02**	**6.07**
				NP2	**0.111**	**0.02**	**4.62**
				Sil	0.019	0.03	0.66
Group:RoI	**44.06**	**13**	**<.05**	NP1	-0.036	0.08	-0.44
				Verb	**-0.169**	**0.08**	**-2.06**
				NP2	-0.16	0.08	-1.91
				Sil	-0.086	0.09	-1.0

Table 13

Model parameters for separate model of IWA's gaze proportions in incorrect trials in the OVS condition in Experiment 2.

Effect	Model comparisons			contrast	Model values		
	χ^2	df	p-value		b	SE	t
Group	**18.81**	**5**	**<.05**	IWAvs.C	**-0.141**	**0.03**	**-4.98**
Roi	**343.31**	**9**	**<.05**	NP1	-0.048	0.02	-1.93
				Verb	**-0.161**	**0.02**	**-6.44**
				NP2	**0.33**	**0.02**	**13.29**
				Sil	**0.246**	**0.03**	**9.25**
Group:RoI	**118.62**	**13**	**<.05**	NP1	-0.015	0.07	-0.23
				Verb	**0.151**	**0.07**	**2.28**
				NP2	**-0.366**	**0.07**	**-5.52**
				Sil	**-0.309**	**0.07**	**-4.6**