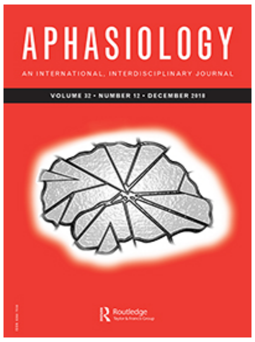

Aphasiology

ISSN: 0268-7038 (Print) 1464-5041 (Online) Journal homepage: informahealthcare.com/journals/paph20

What happens when they think they are right? Error awareness analysis of sentence comprehension deficits in aphasia

Miren Arantzeta, Janet Webster, Itziar Laka, Maite Martínez-Zabaleta & David Howard

To cite this article: Miren Arantzeta, Janet Webster, Itziar Laka, Maite Martínez-Zabaleta & David Howard (2018) What happens when they think they are right? Error awareness analysis of sentence comprehension deficits in aphasia, Aphasiology, 32:12, 1418-1444, DOI: 10.1080/02687038.2017.1423270

To link to this article: https://doi.org/10.1080/02687038.2017.1423270

APHASIOLOGY
2018, VOL. 32, NO. 12, 1418–1444
https://doi.org/10.1080/02687038.2017.1423270

What happens when they think they are right? Error awareness analysis of sentence comprehension deficits in aphasia

Miren Arantzeta[a], Janet Webster[b], Itziar Laka[c], Maite Martínez-Zabaleta[d] and David Howard[b]

[a]International Doctorate for Experimental Approaches into Brain and Language (IDEALAB), Universities of Groningen (NL), Newcastle (UK), Potsdam (DE), Trento (IT) & Macquarie University Sydney (AU), Groningen, The Netherlands; [b]Centre for Research in Linguistics and Language Sciences, Newcastle University, Newcastle upon Tyne, UK; [c]Department of Linguistics and Basque Studies, University of the Basque Country (UPV/EHU), Vitoria-Gasteiz, Spain; [d]Biodonostia Health Research Institute, San Sebastian University Hospital, Donostia-San Sebastian, Spain

ABSTRACT

Background: Comprehension of non-canonical sentences is frequently characterised by chance level performance in people with aphasia (PWA). Chance level performance has been interpreted as guessing, but online data does not support this rendering. It is still not clear whether the incorrect sentence processing is guided by the compensatory strategies that PWA might employ to overcome linguistic difficulties.

Aims: We aim to study to what extent people with non-fluent aphasia are aware of their sentence comprehension deficits.

Methods & Procedures: This study combined offline and online data to investigate the effect of word order and error-awareness on sentence comprehension in a group of PWA and non-brain damaged (NBD) participants. The offline tasks involved auditory sentence picture-matching immediately followed by a confidence rating (CR). Participants were asked to judge the perceived correctness of their previous answer. Online data consisted of eye-tracking.

Outcomes & Results: Replicating previous findings, PWA had significantly worse comprehension of Theme-Agent order compared to Agent-Theme order sentences. Controls showed ceiling level sentence comprehension. CR was a poor predictor of response accuracy in PWA, but moderate-good in NBD. A total of 6.8% of judgements were classified as “guessing” by PWA. Post hoc gaze data analysis indicated that CR was a predictor of the fixation pattern during the presentation of the linguistic stimuli.

Conclusions: Results suggest that PWA were mostly unaware of their sentence comprehension errors and did not consciously employ strategies to compensate for their difficulties.

ARTICLE HISTORY
Received 1 March 2017
Accepted 28 December 2017

KEYWORDS
Aphasia; sentence comprehension; error awareness; eye-tracking; anosognosia

CONTACT Miren Arantzeta m.arantzeta.perez@rug.nl International Doctorate for Experimental Approaches into Brain and Language (IDEALAB), Universities of Groningen (NL), Newcastle (UK), Potsdam (DE), Trento (IT) & Macquarie University Sydney (AU), Groningen, The Netherlands

Introduction

Word order and sentence comprehension deficits in people with aphasia (PWA)

Sentence comprehension deficits in agrammatic aphasia have been well established in the literature over the past few decades. In the absence of lexical comprehension deficits, semantically reversible sentences presented in non-canonical word orders (the terms canonical vs. non-canonical refer to the base word order in a certain language (e.g., SVO in Spanish) vs. the rest of word orders allowed in the language (e.g., OVS, VSO, VOS)) are frequently misunderstood by people with non-fluent aphasia (PWA) (e.g., Caplan & Hildebrandt, 1988; Caramazza & Zurif, 1976; Grodzinsky, 2000; Schumacher et al., 2015; Thompson et al., 2013; see Grodzinsky, Piñango, Zurif, & Drai, 1999; for a review). Both semantic reversibility and word order are key in understanding this difficulty. When both Determiner Phrases (DPs) of a sentence are animate and potential agents of the action (i.e., the sentence is reversible), PWA show better comprehension of sentences presented in Agent-Theme order (henceforth A-T) (e.g., *The girl calls the teacher*) than in Theme-Agent order (henceforth T-A) (e.g., *The teacher has been called by the girl*). This is because in the former, listeners may rely solely on word order information to disentangle the thematic-role assignment, while in the latter they are forced to process morphological information to parse and reach the correct interpretation of the sentence.

Typologically, the more (case and agreement) morphology a language has, the greater freedom it displays in sentence word order. However, deficits in comprehending sentence in T-A argument order have been found in PWA speakers of languages with more rigid word order (i.e., in English, Meyer, Mack, & Thompson, 2012; Schwartz, Saffran, & Marin, 1980; in Dutch, Bastiaanse & Edwards, 2004), flexible word order (i.e., in Spanish, Juncós-Rabadán, Pereiro, & Souto, 2009; in German, Burchert, De Bleser, & Sonntag, 2003; in Italian, Garraffa & Grillo, 2008; in Turkish, Duman, Altinok, Özgirgin, & Bastiaanse, 2011; in Swahili, Abuom, Shah, & Bastiaanse, 2013; cf. in Indonesian, Jap, Martínez-Ferreiro, & Bastiaanse, 2016), and free word order (i.e., in Basque, Arantzeta et al., 2017).

To date, most studies on sentence comprehension have used offline methods to address the ways in which PWA and non-brain-damaged (NBD) individuals process sentences in order to assign grammatical functions and thematic roles. The sentence–picture matching task has typically been used in both experimental and clinical settings. The participant needs to choose, within a set of two (or more) visual stimuli, the picture that best matches the target sentence. Although the results are easy to quantify, it is necessary to consider that the odds of picking the target picture by chance are relatively large. Thus, to compare results against chance, accuracy scores at chance level have been traditionally attributed to guessing by the Trace Deletion Hypothesis (TDH; Grodzinsky, 1986, 1995, 2000; see Drai & Grodzinsky, 2006a, 2006b for a later revision). Based on the tenets of the Principles and Parameters model of generative grammar (Chomsky, 1981), the TDH states that agrammatism precludes the creation of a chain between the moved element and the trace in its original position. According to the TDH, traces of movement are not available to PWA. Consequently, when presented with passive sentences with non-canonical T-A order, PWA cannot assign thematic roles to a moved argument and apply instead a default strategy that assigns the role of Agent to the first DP in the sentence. The argument in the by-phrase gets the correct thematic role, and therefore, the sentence appears to have two agents. Accordingly, the TDH states that PWA resolve this conflict by choosing randomly between the two potential interpretations of the sentence. Thus, the TDH predicts that PWAs will perform at chance level in the comprehension of sentences with non-canonical A-T word order.

The above-mentioned guessing interpretation is related to one important limitation of offline sentence comprehension tasks. Offline methods are static in the sense that they measure how participants interpret a sentence once its presentation has concluded, but they do not provide information related to the timing involved on specific interpretations. Online methodologies address this limitation, allowing monitoring of performance as it takes place.

Eye-tracking studies on sentence comprehension deficits

To examine real-time sentence processing, several studies have used eye-tracking technology in the Visual World Paradigm (VWP). The VWP is based on the idea that the linguistic stimuli mediate visual attention shifts within a visual display due to referentially driven processes (Cooper, 1974; Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995; see Boland, 2004; Tanenhaus, 2007; for an overview). Participants' eye fixations on pictures in a scene are monitored as participants listen to sentences. Changes in the location and timing of fixations during the sentence presentation reveal how visual attention shifts in response to the continuous auditory stimuli. Some findings suggest that language-mediated eye gaze tends to be unconscious, but not fully automatic (Mishra, Olivers, & Huettig, 2013). Current data suggest that the VWP provides a sensitive measure of word order processing in sentence comprehension.

Dickey, Choy, and Thompson (2007) introduced the VWP in aphasiology research in a study on *wh*-questions. They found that the comprehension of non-canonical object *wh*-questions was at chance level in PWA. Interestingly, gaze data analysis did not converge with the traditional interpretation of chance level accuracy (i.e., guessing) (TDH; Grodzinsky, 1986; 1995, 2000; see Drai & Grodzinsky, 2006a, 2006b for a later revision; see Burchert, Hanne, & Vasishth, 2013; for a review). PWA showed distinct fixation patterns in the correctly and incorrectly answered trials. In the correctly answered trials, PWA showed the same gaze fixation pattern of NBD participants, indicating rapid and automatic thematic role assignment. In contrast, in the incorrectly answered trials, PWA showed a completely different gaze-fixation pattern, characterised by a progressive proportion of fixations towards the foil picture. Similar results have been reported in subsequent studies combining VWP with a sentence–picture matching task in English (Dickey & Thompson, 2009), German (Hanne, Sekerina, Vasishth, Burchert, & De Bleser, 2011; Schumacher et al., 2015), and Basque (Arantzeta et al., 2017) PWA. Altogether, evidence from online data does not support the existence of a guessing pattern in sentence processing in PWA, but distinctive parsing routines that determine the interpretation of the sentence. In successful interpretations of the sentence, PWA show a processing pattern comparable to NBD, but it is still unclear what processes underlie sentence misinterpretation.

Online data analysis of incorrect answers might provide insights regarding the intermittent and hardly predictable (i.e., stochastic) failure shown by PWA in processing sentences with non-canonical word orders. However, this analysis is often challenging due to the small sample sizes and noisy data (see Caplan, Waters, DeDe, Michaud, & Reddy, 2007 for an alternative across-task/measures approach). It has been suggested that PWA fail to comprehend sentences with non-canonical word order due to their inability to inhibit conflicting information (Dickey et al., 2007; Dickey & Thompson, 2009; Schumacher et al., 2015), for example, the influence of agent-first heuristics. This refers to the tendency to assign the agent role to the first DP in the sentence (Bever, 1970; Bornkessel-Schlesewsky & Schlesewsky, 2013). This heuristic achieves correct interpretation of sentences presented in the canonical word order, also in healthy speakers (Ferreira & Patson, 2007; Townsend & Bever, 2001; Van Herten, Chwilla, & Kolk, 2006), but it fails to correctly assign thematic roles in derived word orders. In those cases, listeners must revise the initial parsing by applying more effortful analytical computations. Hanne et al. (2011) suggested that PWA have an early preference for interpreting sentences as canonical or non-canonical based on a "deterministic parsing", followed by the inability to revise the initial parsing computation, even when they detect the need for reanalysis. Still, it is an open question whether the use of heuristics by PWA is a consciously learned and self-initiated procedure to compensate for their linguistic deficits, or whether it reflects an unconscious breakdown of parsing routines.

Consciousness and compensatory strategies in PWA

Provided that PWA do not show anosognosia (i.e., unawareness of the aphasia) (see Kertesz, 2010; Vuilleumier, 2004; for an overview), PWA can use self-initiated compensatory strategies to overcome their communicative impairments. These might be external (e.g., ask for adaptations to the interlocutor, use of electronic devices) or internal (e.g., self-cuing, verbal repetition, mental association) (e.g., Beeke, Wilkinson, & Maxim, 2009; Oelschlaeger & Damico, 1998; Tompkins, Scharp, & Marshall, 2006; see Simmons-Mackie & Damico, 1997).

Within sentence comprehension, the use of heuristics reduces processing time and effort in parsing routines when compared to analytical processes (Shah & Oppenheimer, 2008; see also Ferreira & Patson, 2007). There is evidence suggesting that the PWA have limited resource availability to process the linguistic stimuli (Caplan et al., 2007; Kolk, 1993; Miyake, Carpenter, & Just, 1994). Thus, PWA may adopt a conscious shortcut to diminish the cognitive load involved in the parsing process. Among the languages studied in sentence processing in PWA, none has Theme-Agent order as its canonical structure. Indeed, this is a rare pattern shown by less than the 4% of the languages worldwide (Dryer, 2005). The frequency of appearance of structures is a primary criterion to determine word order typology across languages; thus, structures with non-canonical word order tend to be less frequent than structures with canonical word order (Dryer, 2007). Hence, if we consider the above factors, reliance on agent-first strategy may be considered as the "best guess" under an arbitrary degree of success. It is unknown to what extent the adoption of an agent-first strategy by PWA is based on a conscious decision that aids comprehension efficiency on an everyday basis. The focus lies in establishing a threshold between the conscious and unconscious processing of language.

Metacognitive tasks assess self-awareness and, hence, consciousness (Seth, Dienes, Cleeremans, Overgaard, & Pessoa, 2008; see Peña-Ayala & Cárdenas, 2015; for an overview). Metacognitive awareness refers to the subjective perception of one's own cognitive process. It can be measured by the subjective Confidence Ratings (CRs) self-reported by the subject in a given task (see Norman & Price, 2015). Participants perform a primary task (e.g., picture matching) and subsequently are asked to rate their confidence in the validity of their decision. The degree of correspondence between the objective performance (i.e., accuracy) and the subjective CR is used to assess the extent to which the primary task is mediated by conscious knowledge (Cheesman & Merikle, 1984, 1986; Dienes, Altmann, Kwan, & Goode, 1995; Overgaard, Timmermans, Sandberg, & Cleeremans, 2010). Conscious cognition is strongly associated with voluntary control. However, voluntary actions become automatic with practice (Shiffrin & Schneider, 1977) and consequently, the cognitive control over them decreases (e.g., Langer & Imber, 1979; Schneider, 2009). Hence, it is important to keep in mind that there is some degree of contribution of conscious and unconscious knowledge across most cognitive tasks.

Note that a judgement based on CR is certainly a constituent of conscious awareness, but it may be the product of an unconscious inference. What is important is that once the judgement is built, it may augment self-control, and therefore the degree of personal regulation over processes that would otherwise influence behaviour directly (Koriat, 2000). Dienes et al. (1995) suggested two criteria to use CRs to distinguish between conscious and unconscious knowledge: the zero-correlation criterion and the guessing criterion. The former refers to the lack of relationship between CR and objective accuracy, while the latter refers to the observed above-chance performance in the primary task when participants say they are guessing. According to the authors, the fulfilment of these two criteria is a strong indicator of unconscious processing. This paradigm has been used mostly in perceptual discrimination and implicit learning tasks (Norman & Price, 2015; Overgaard, 2015; for an overview). In the current study, it is applied for the first time in an aphasiology study.

Studies of metacognitive awareness in comprehension deficits have been mainly on jargon aphasia (Marshall, Rappaport, & Garcia-Bunuel, 1985; Shuren, Hammond, Maher, Rothi, & Heilman, 1995; see Rubens & Garret, 1991; for an overview). To the best of our knowledge, the only study on comprehension error awareness in non-fluent PWA was conducted by Kennedy and Chiou (2008), who studied a group of Broca's and anomic PWA regarding metacognitive awareness on discourse-related questions. Kennedy and Chiou (2008) reported that metacognitive awareness was mainly explained by the linguistic abilities of the PWA (discourse comprehension and design fluency repetition score), but also executive functions, such as switching and perseveration, contributed to a lesser extent (see also Stuss, 1991).

PWA have difficulties comprehending sentences presented in derived word orders. We aimed to study the extent to which they are aware of their comprehension difficulties, and whether they use conscious strategies to compensate them. Comprehension accuracy and self-reported CRs were considered during a comprehension task involving canonical and non-canonical sentences.

Research questions

Behavioural data were collected using a sentence–picture matching task, followed by a CR. Moreover, participants were monitored with an eye-tracker while performing the primary task (i.e., sentence–picture matching task) as an online measure of sentence processing. The latter aimed to further explore, in a post hoc analysis, the interaction between self-awareness and real-time sentence processing. This study poses three research questions:

(1) Are PWA aware of their sentence comprehension errors?

The CRs reported for the incorrectly answered trials were studied to uncover the extent to which PWA performed a correct judgement of their (failed) comprehension accuracy. Correct judgements of incorrectly answered trials were considered as an indicator of error awareness in sentence comprehension. In contrast, incorrect judgements of incorrectly answered trials suggest that PWA were not aware of their comprehension errors.

(2) Is sentence comprehension performance unconsciously mediated in PWA and NBD?

Following Dienes et al. (1995), this research question was answered based on the zero-correlation criterion and the guessing criterion. In relation to the zero-correlation criterion, the absence of correspondence between objective comprehension performance and subjective CR is an indicator of unconscious knowledge. The opposite pattern would be indicative of consciously mediated processing, and therefore voluntary control over thematic-role parsing. Regarding the guessing criterion, above-chance performance in trials where participants believed that they were guessing indicate unconscious knowledge (i.e., implicit knowledge) of sentence processing.

(3) To what extent do PWA answer by guessing on a task for comprehension of sentences in non-canonical argument order?

Descriptive analysis of the data was used to add further evidence on the validity of the TDH (Grodzinsky, 1986, 1995, 2000; see Drai & Grodzinsky, 2006a, 2006b for a later revision). As described before, the TDH states that PWA answer by guessing when confronted with sentences presented in derived word orders. However, the TDH does not explicitly report if the guessing act is a conscious/unconscious process. Based on the definition of guessing as "form(ing) or express(ing) an uncertain estimate or conclusion (about something), based on insufficient information" (Collins English Dictionary, 2012, online version), in the current study we assume that it refers to a conscious act. Hence, the self-rating of PWA will be used to explore whether the "guessing" pattern is conscious.

Methods

This study was approved by the Basque Clinical Research Ethics Committee (CEIC-E). All participants, or legal tutors, gave informed consent according to the declaration of Helsinki.

Participants

A total of 14 individuals (mean age 66.07 years; *SD* = 10.38; range = 55–85; male/female = 11:3) with chronic non-fluent aphasia were tested in this study. Half of them were native Spanish monolingual speakers, whereas the others were L1Basque-L2Spanish bilingual speakers. All bilingual participants had acquired Spanish at an early age (<5 years) and were literate only in Spanish, their language of instruction at school. For more information on sentence comprehension abilities as a function of bilingualism, see Arantzeta (2017). They were all pre-morbidly right-handed, as assessed by the Edinburgh Handedness Inventory (Oldfield, 1971) and presented with aphasia due to cerebrovascular injury. Visual neglect was excluded using the Behavioural Inattention Test (Wilson, Cockburn, & Halligan, 1987). A total of 14 individuals without any history of neurological or sensory impairments composed the NBD group. They were matched on age range and bilingualism with the clinical group (mean age 62.92 years; *SD* = 12.04; range = 44–82; male/female = 8:6). Demographic, linguistic, and clinical information is provided in Table 1.

PWA were screened in Spanish for word and sentence comprehension abilities using the extended version of the Boston Aphasia Test (BDAE; Goodglass, Kaplan, & Barresi, 2005; Adapted to Spanish by García-Albea, 2005). The subsection of conversation and language exposition, comprehension of words, commands, complex ideational materials, and syntactic processing subsections were used. Furthermore, PWA were also assessed for working memory using the digit-span task of the Wechsler Adult Intelligence Scale (WAIS-III; Wechsler, 1997). Scores from the subtests of the BDAE and the WAIS-III are presented in Table 2.

Table 1. Demographic, linguistic, and clinical background information of the participants.

Participant	Age	Gender	Education level	Linguistic profile	Y:M post-onset	Location
A1	68	F	1	Bilingual	0:10	I- LMCA
A2	55	M	2	Bilingual	7:4	I-RMCA
A3	57	M	2	Bilingual	9:10	I- LMCA
A4	61	M	2	Bilingual	1:5	I- LMCA
A5	82	M	1	Bilingual	2:8	I- LMCA
A6	85	M	2	Bilingual	0:9	I- LMCA
A7	64	M	1	Bilingual	2:5	I- LMCA
A8	58	M	2	Monolingual	>8:00	AE
A9	79	F	1	Monolingual	2:9	I-LMCA
A10	59	M	1	Monolingual	2:11	I- LMCA
A11	77	M	2	Monolingual	2:10	I- LMCA
A12	57	F	2	Monolingual	1:00	I-LMCA
A13	65	M	2	Monolingual	5:7	I- LMCA
A14	58	M	1	Monolingual	2:6	I- LMCA
C1	67	M	2	Bilingual	n/a	n/a
C2	54	M	2	Bilingual	n/a	n/a
C3	63	F	1	Bilingual	n/a	n/a
C4	58	F	1	Bilingual	n/a	n/a
C5	77	F	1	Bilingual	n/a	n/a
C6	82	M	2	Bilingual	n/a	n/a
C7	73	M	1	Bilingual	n/a	n/a
C8	44	F	2	Monolingual	n/a	n/a
C9	64	F	1	Monolingual	n/a	n/a
C10	58	M	3	Monolingual	n/a	n/a
C11	79	M	2	Monolingual	n/a	n/a
C12	53	F	2	Monolingual	n/a	n/a
C13	65	M	1	Monolingual	n/a	n/a
C14	44	M	2	Monolingual	n/a	n/a

F = female; M = male; 1 = elementary; 2 = technical; 3 = university; I-LMCA = ischemic-left middle cerebral artery; I-RMCA = ischemic-right middle cerebral artery; AE = anoxic encephalopathy; n/a = not applicable.

Table 2. Summary of the individual scores obtained in the Wechsler Adult Intelligence Scale (WAIS-III; Wechsler, 1997) and the extended version of the Boston Aphasia Test (%) (BDAE; Goodglass, Kaplan, & Barresi, 2005; Adapted to Spanish by García-Albea, 2005).

	WAIS-III	BDAE			Syntactic processing (%)	
Participant	Digit-span (pCTL)	Word comprehension	Commands	Complex ideational material	Touching A with B	Embedded sentences
A1	68.8	100	86.66	41.66	58.33	60
A2	14.7	100	86.66	83.33	66.66	80
A3	68.8	94.59	100	75	58.33	60
A4	6.8	91.89	80	58.33	33.33	30
A5	14	86.48	66.66	41.66	16.66	70
A6	14	94.59	86.66	66.66	50	40
A7	–	91.89	73.33	50	25	50
A8	68.8	94.59	100	75	66.66	60
A9	0.7	94.59	80	83.33	33.33	90
A10	0	89.1	66.66	50	66.66	60
A11	14	94.59	86.66	91.66	33.33	90
A12	0	91.89	80	58.33	33.33	30
A13	–	81.08	–	50	16.66	60
A14	37.1	97.29	100	66.66	41.66	90

pCTL: percentile by age range.

To be included in the present study, PWA showed relatively preserved lexical comprehension (>80%) and impaired sentence processing (<66%) as assessed by the "Touch A with B" sub-test. This sub-test distinguishes a reversible relationship between sentence constituents presented in a variety of word orders.

Design and materials

The materials used in this study were adapted to standard European Spanish from the experiment presented by Arantzeta et al. (2017). They consisted of single sentences provided auditorily in combination with the presentation of two pictures on the screen. Each of the pictures depicted two people taking part in the performance of the same action, but with reverse Agent-Theme thematic roles (see Figure 1). There were 126 trials consisting of 120 experimental items and 6 practice items.

Picture-matching task

Linguistic stimuli

The same 22 transitive verbs used in the original study (Arantzeta et al., 2017) as well as two animate singular DPs assigned to each verb were selected to create declarative sentences in the following structures: (a) active; (b) passive; (c) subject relative; (d) object relative; (e) subject cleft; (f) object cleft. Subsequently, sentence conditions were clustered as Agent-Theme (A-T) and Theme-Agent (T-A) for data analysis. The Agent/Theme assignment in the DPs was randomised and balanced within the six conditions. Regions of interest (ROIs) of the experimental stimuli were individually measured using the Computerized Language Analysis software (MacWhinney, 2000). *t*-Test comparisons of ROI durations between paired conditions showed no differences (see Table 3).

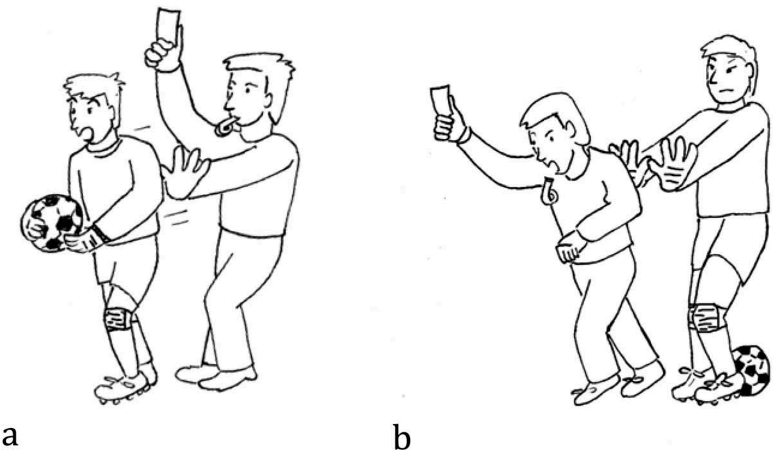

Figure 1. Sample visual display. Target stimulus (Active): "El árbitro ha empujado al portero" (The referee has pushed the goalkeeper). (a) Target picture; (b) Foil.

Table 3. Regions of Interest (ROI), duration (ms) (mean and *SD*), and comparison of length across paired conditions.

	ROI (mean duration and *SD*)					
	ROI 1 Argument 1 999 ms (46)		ROI 2 Verb 931 ms (12)		ROI 3 Argument 2 852 ms (30)	
Paired conditions	*t*	*p*	*t*	*p*	*t*	*p*
Active vs. Passive	0.945	.350	−0.174	.863	−1.376	.177
Subj. vs. Obj. Relative	−0.221	.826	−1.877	.068	1.268	.213
Subj. vs. Obj. Cleft	1.074	.290	−0.397	.693	0.366	.716
A-T vs. T-A	−0.554	.581	−0.027	.177	0.077	.939

ROI: Region Of Interest; A-T: Agent-Theme; T-A: Theme-Agent.

In the linguistic stimuli, the active sentences (1) were constructed with perfect present tense to keep the length of the verb as equal as possible with the counterpart passive sentences (2). In the relative clauses, the antecedent was always introduced by the verbal phrase "I see" and the relative pronoun "que" functioning as subject (3) or object (4). In the cleft sentence constructions, the contrastive element became the complement of the copular verb "ser", and the relativiser pronoun "que" introduced the rest of the sentence (5–6). In both object-cleft and object-relative constructions, the relative pronoun was preceded by the preposition "a", as well as the direct object in the cleft constructions.

(1) La mujer ha peinado a la niña
det woman aux.has comb-PTCP prep det girl
The woman has combed the girl

(2) La niña ha sido peinada por la mujer
det girl aux.has be-PTCP comb-PTCP prep det woman
The girl has been combed by the woman

(3)

Veo	a	la	mujer	que	peina	a	la	niña
see	prep	det	woman	pron-rel	comb	prep	det	girl

I see the woman who combs the girl

(4)

Veo	a	la	niña	a	la	que	peina	la	mujer
see	prep	det	girl	prep	det	pron-rel	comb	det	woman

I see the girl who the woman combs

(5)

Es	la	mujer	la	que	peina	a	la	niña
be	det	woman	det	rel-pron	comb	prep	det	girl

It is the woman who combs the girl

(6)

Es	a	la	niña	a	la	que	peina	la	mujer
be	prep	det	girl	prep	det	rel-pron	comb	det	woman

It is the girl who the woman combs

Visual stimuli

The 44 black-and-white line drawings used by Arantzeta et al. (2017) were adapted for the current study. The drawings constituted 22 pairs, which were presented together divided by a black vertical line in the middle of the screen. The two pictures within each pair showed the same action with an Agent-Theme reversal. See Arantzeta et al. (2017) for detailed information on the visual stimuli and their corresponding normalisation.

The delivery of the visual stimuli was controlled to reduce adoption of particular spatial strategies by PWA, as suggested by Chatterjee, Maher, and Heilman (1995). The presentation of the experimental items in the visual display (i.e., right vs. left) was pseudo-randomised. No more than two target stimuli could occur in a row on the same side of the screen. The direction in which the action was performed was also balanced across the stimuli to avoid rightward-inclined scanning (Scheepers & Crocker, 2004).

Confidence rating

A three-interval CR was used and visually represented by coloured emoticons. From right to left, a green smiley face, an amber neutral face, and a red sad face were presented following a horizontal axis (seeFigure 2). The first (i.e., green smiley emoticon) and last (i.e., red sad emoticon) implied confidence of response to some degree, corresponding to "sure, I answered correctly" and "sure, I answered incorrectly", respectively. The middle response alternative (i.e., amber neutral) corresponded to having no knowledge to report the accuracy judgement, and it was defined as "I don't know/guess".

Procedure

The experiment was programmed using E-Prime 2.0.10 with extensions for Tobii 2.0.2.41. The visual stimulus was presented on a 23″ wide screen LED monitor having 1280 × 720 resolution. The auditory stimuli were delivered binaurally through headphones. Participants were seated 60–70 cm from the screen, with a visual angle under 15° (max. allow 35°), while their eye movements were captured using a Tobii 120 Desktop Eye Tracker (sampling rate 120 Hz).

Figure 2. Response grid used to self-report confidence rate on the accuracy of the sentence–picture matching task.

The emoticons represented the next confidence rates: green smiley face = "sure, I answered correctly"; amber neutral face = "I don´t know/guess"; red sad face = "sure, I answered incorrectly".

The presentation of the experimental stimuli was divided into four blocks of 30 items. Two blocks were administered in each of the two experimental sessions, after the presentation of six practice items. No more than two consecutive items corresponded to the same linguistic condition.

Prior to the start of the presentation of each block of stimuli, a 5-point calibration sequence was used to calibrate the eye-tracker to each subject. Subsequently, written instructions for the task appeared on the screen. The same instructions were also read aloud and verbally explained to all participants.

The experiment consisted of two consecutive tasks. In the primary task, the participants were asked to perform a picture-matching task, while in the secondary task they were asked to report the perceived correctness of their previous answer using a confidence rate.

At the beginning of each trial, a fixation smiley face was centred in the screen. Participants had to fixate on the image for 250 ms before it was replaced by the display of a pair of pictures corresponding to the experimental stimuli. After 1000 ms of pre-visualisation, the auditory stimulus was presented. Participants were asked to select the picture that best depicted the auditorily presented sentence by pressing specific keys on the keyboard with the left hand. As an exception, a participant with crossed aphasia (A2) answered with the right hand. Trials with no response within 8000 ms from the offset of the sentence were counted as having no response. When the decision was made, the second task was introduced. Participants were asked to make a CR as an assessment of the accuracy of their response on the previous task (i.e., picture-matching task). To this end, the three colour emoticons appeared on the screen and participants had to select the one that best represented their judgement. They were allowed to make their choice by pressing specific buttons on the keyboard or pointing directly on the screen. CRs provided by pointing directly on the screen were entered by the researcher using a second keyboard. The participants had a maximum of 12,000 ms to answer. After this time, the CR was registered as not being answered and the next trial was started. Trials answered before the onset of the auditory presentation (i.e., pre-visualisation face) were excluded from subsequent data analysis, corresponding to the 0.25% of the total data.

In both the first and second tasks, trials with no responses were excluded from further analysis, corresponding to 1.28% and 0.58% of the total data, respectively. Fixations with durations less than 90 ms were removed from the analysis to exclude ocular artefacts (e.g., blinks and saccades). We introduced a switch of 200 ms to correct for the time required for planning and executing an eye-movement, and thereby, to time align the gaze-fixation data with the linguistic stimuli (Matin, Shao, & Boff, 1993).

The difference in the proportion of fixations between the target and foil pictures was computed across different temporal ROIs. ROI 1 corresponded to the subject/object of the main clause or the antecedent of the relative clause; ROI 2 corresponded to the (relative pronoun +) verbal phrase; and ROI 3 corresponded to the subject/object of the verbal phrase. In addition, a post-offset ROI 4 with a duration of 1120 ms was included. Missing gaze data were imputed in the trials answered before the offset of ROI 4 (i.e., mean RT <3902 ms) based on the accuracy of the response.

Data analysis

Comprehension accuracy data and post hoc gaze data analysis were conducted using Generalized Linear Mixed-effects Models (GLMM) with logit function, and Linear Mixed-effects Models (LMM), respectively. GLMM combine both fixed and random effects of known variables in a single model (see Bates, Maechler, Bolker, & Walker, 2015). To the specific case of our study, GLMM are suitable to address the variability across participants/stimuli, outliers, and missing data, a common characteristic of PWA, small sample sizes, and time-series data. For a detailed discussion on these statistical aspects, see Diggle, Heagerty, Liang, and Zeger (2002) and Verbeke and Molenberghs (2000).

Model building was conducted by progressively introducing random effects, fixed effects, and corresponding interactions. Akaike's Information Criterion (Akaike, 1974) was used to measure the goodness of fit and compare models with each other. The numerical predictors Age and Trial number were centred. Least squares means (LSMeans) were used for comparing LSMean differences on the basis of the mixed model.

Specific procedures were followed to test the zero-correlation and guessing criterions. Zero-correlation criterion was analysed with simple Logistic Regression Models (LRM) fitted with a logit link function, in addition to the estimation of the area under the receiver operating characteristic (ROC) curve. In relation to the zero-correlation criterion, the absence of correspondence between objective comprehension performance and subjective CR is an indicator of unconscious knowledge. Thus, the area under the curve (AUC) is the probability that participants correctly rate their accuracy. The nonparametric approach of DeLong (DeLong, DeLong, & Clarke-Pearson, 1988) was used to compare the AUCs between groups. The guessing criterion was examined using GLMM, by comparing the linear predictor with zero. LSMeans and 95% CIs were used for this purpose.

Tukey and Bonferroni correction was used to control the error rate in testing multiple comparisons on behavioural (i.e., accuracy and CR) and gaze-fixation data, respectively. The statistical software R was used for this analysis (R Core Team, 2015, v.3.2.3.).

Results

Comprehension accuracy

Sentence conditions were clustered as Agent-Theme (A-T) and Theme-Agent (T-A). The former included active, subject cleft and subject relative conditions, while the latter included the counter pairs: passives, object clefts, and object relatives.

PWA correctly comprehended 79% (Standard Error, *SE* = 1.42) of the sentences in A-T order and 65% (*SE* = 1.67) of the sentences presented in T-A order. The NBD group performed at ceiling level obtaining 96% (*SE* = 0.69) accuracy in A-T and 96% (*SE* = 0.64) accuracy in T-A argument orders. Detailed statistics on sentence comprehension scores by argument orders are presented in Table 4. Individual participants' scores are in Appendix A.

In the final GLMM, the predictors of interest were the group (PWA, NBD), argument order (A-T, T-A), and trial number (from 1 to 120, the position of the presentation of a given trial in relation to the others). There was a two-way interaction between group and argument order, and group and trial number, as well as three-way interaction between group, argument order, and trial number. By means of random effects, we accounted for variability within-participants (i.e., participants that answered one trial correctly would be more likely to answer other trials correctly) and within stimuli (i.e., some stimuli may be more difficult than other).

The PWA performed significantly poorer than the NBD across both A-T ($\beta = -2.029$; $SE = 0.324$; $p = < .0001$) and T-A ($\beta = -2.703$; $SE = 0.305$; $p = < .0001$) argument orders. The PWA comprehended sentences presented in A-T better than T-A argument order ($\beta = 0.754$; $SE = 0.120$; $p = < .0001$), while the people with NBD showed no effect of word order ($\beta = 0.080$; $SE = 0.277$; $p = .7728$). In addition, there was an interaction between group, argument order, and trial number. The NBD participants became more accurate in the sentence-picture matching task across the course of the experiment, but this improvement was restricted to sentences presented in A-T argument order (A-T; $\beta = -2.43$; $SE = 0.704$; $p = .0006$. T-A; $\beta = -0.736$; $SE = 0.701$; $p = .2942$). The PWA did not show any effect of trial number across any of the argument orders (A-T; $\beta = -0.414$; $SE = 0.343$; $p = .2278$. T-A; $\beta = -0.327$; $SE = 0.330$; $p = .3229$).

Confidence ratings

Participants expressed their confidence about the correctness of their answer in the sentence–picture matching task by choosing one of the three options. The distribution of the CRs is presented in Table 5, separately for correctly and incorrectly comprehended trials. See Figure 3 for an overall distribution of responses in each group. Strikingly, both NBD participants and PWA believed they were correct on almost all trials (97% for NBD and 87% for PWA). On the incorrect trials, PWA were sure that they were right on 78% of occasions.

Table 4. Comprehension accuracy (%) and standard error (*SE*) as a function of group and sentence condition.

	Condition					
	Agent-Theme			Theme-Agent		
Group	Active	Subj. Cleft	Subj. Rel.	Passive	Obj. Cleft	Obj. Rel.
PWA	81.88% (2.32)	79.92% (2.44)	74.81% (2.60)	74.34% (2.66)	60.44% (2.99)	61.62% (2.95)
NBD	98.17% (0.80)	96.05% (1.16)	93.18% (1.51)	97.45% (0.95)	96.78% (1.05)	95.00% (1.30)

PWA: People With Aphasia; NBD: Non-Brain Damaged.

Table 5. Distribution (%) of the confidence ratings within correctly and incorrectly answered trials in both PWA and NBD groups.

	PWA			NBD		
Response accuracy	"sure, I answered correctly"	"I don't know/ guess"	"Sure, I answered incorrectly"	"Sure, I answered correctly"	"I don't know/ guess"	"Sure, I answered incorrectly"
Correct answers	90.44% (*n* = 1060)	5.71% (*n* = 67)	3.83% (*n* = 45)	98.37% (*n* = 1576)	1.43% (*n* = 23)	0.18% (*n* = 3)
Incorrect answers	77.60% (*n* = 350)	9.53% (*n* = 43)	12.86% (*n* = 58)	49.23% (*n* = 32)	15.38% (*n* = 10)	35.38% (*n* = 23)

PWA: People With Aphasia; NBD: Non-Brain Damaged.

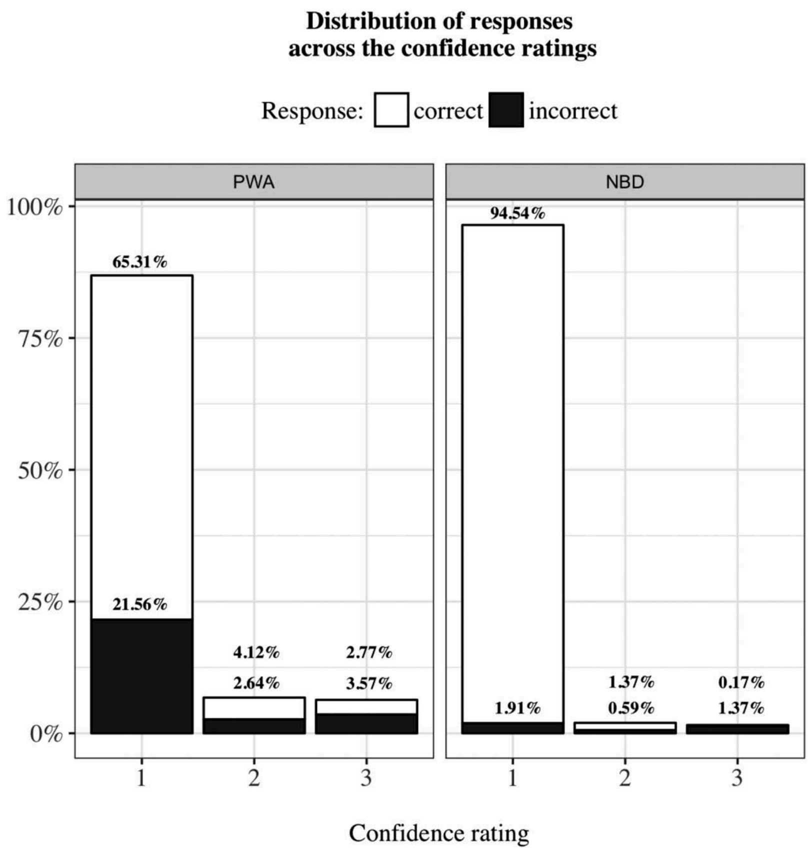

Figure 3. Distribution (%) of the self-reported judgement of the sentence–picture matching task as a function of response accuracy in both NBD and PWA groups.

Confidence rating 1 = "sure, I answered correctly", 2 = "I do not know, guessing", 3 = "sure, I answered incorrectly". NBD: non-brain damaged; PWA: people with aphasia.

Zero-correlation criterion

Data were sorted into correct/incorrect judgements. Correct judgements refer to the instances in which participants reported as correct the trials answered correctly and reported as incorrect the trials answered incorrectly. Thus, regardless of the accuracy of their response in the primary task (i.e., sentence–picture matching task), participants made a correct judgement of their performance. Incorrect judgements refer to the cases in which participants' subjective self-report (i.e., CR) did not match with the correctness of their answer in the sentence–picture matching task.

LRM analysis was conducted to explore the relationship between the outcome of sentence comprehension accuracy and the CR provided by the participants. Separate logit models were fitted for each group of participants to analyse the probability of each group to correctly rate the accuracy of their response. ROC analysis considered together the sensitivity (percentage of correctly identified incorrect answers) and the 1-specificity (percentage of incorrectly identified incorrect answers) across a range of values, for the ability to predict the sentence accuracy outcome. The results are illustrated in Figure 4. In a hypothetical scenario, in which participants would use the confidence rates 100% accurately to classify the accuracy of their response, the AUC would be 1. In our data, the CR had an AUC of 0.56 (95% CI: 0.54–0.58) in the PWA group, just better than random (i.e., 0.50), while the group with NBD had an AUC of 0.74 (95% CI: 0.68–0.81). A

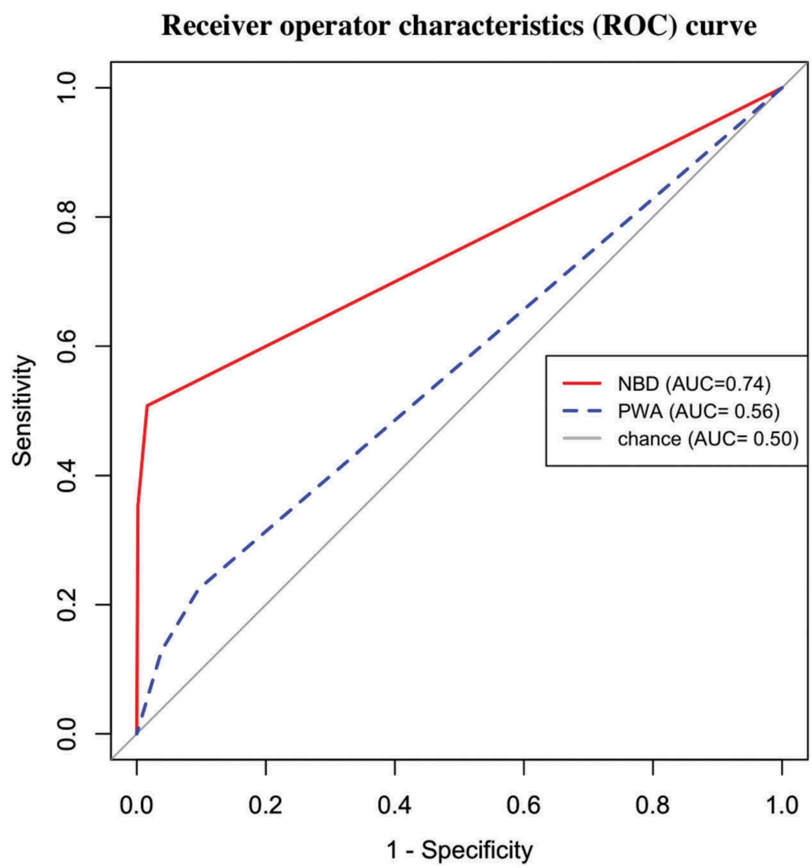

Figure 4. Predicted probability of the confidence rating on the comprehension accuracy, illustrated by the sensitivity (i.e., true positive rate) against 1-specificity (i.e., false positive rate).

NBD: non-brain damaged; PWA: people with aphasia. The area under the ROC curve (AUC) is a measure of how well confidence ratings can distinguish between correctly and incorrectly answered trials.

comparison of AUCs conducted by DeLong's method uncovered significant differences between the groups ($p = < .0001$).

Guessing criterion

The analysis considered if there was above chance accuracy in the sentence–picture matching task in trials where participants declared to guess. In the GLMM, the predictors of interest were the group (PWA, NBD) and CR (1, 2, 3) in a two-way interaction, and the group (PWA, NBD) and argument order (A-T, T-A) in a two-way interaction. The random effects accounted for the variability across participants and stimuli. In addition, we took into consideration that there was differential variability in the effect of argument order and sentence condition in each participant and stimulus. For example, the difficulty imposed by argument order T-A was larger in some participants than in others, or the difficulty imposed by object cleft sentence condition in some stimulus was greater than in others. Note that although only LSMeans related to CR = 2 (i.e., "I don't know/guess") were extracted to get insight the guessing criterion, we fitted a full model including CR as a variable with three levels to avoid convergence problems. The use of a full model provides statistical power to the results and decreases the possibility to obtain false negatives.

The guessing criterion was studied by comparing the linear predictor with zero in the subset of items ranked CR = 2. Significantly positive values refer to instances in which participants showed an above-chance accuracy rate when they believed that they were guessing. The regression intercept was significantly positive in PWA (LSMean = 0.640; *SE* = 0.305; 95% CI = 0.04–1.23; $t(109) = 2.098$; $p = .0382$) and those with NBD (LSMean = 1.254; *SE* = 0.490; 95% CI = 0.29–2.21: $t(32) = 2.557$; $p = .0155$), indicating above-chance accuracy. The same analysis was conducted separately as a function of argument order. In the PWA group, the lower limit of the 95% CI was significantly positive in sentences presented in A-T argument order (LSMean = 1.085; *SE* = 0.336; 95% CI = 0.42–1.74; $t(44) = 2.339$; $p = .0239$), but not in sentences presented in T-A argument order (LSMean = 0.196; *SE* = 0.325; 95% CI = –0.44–0.833; $t(64) = 0.605$; $p = .5469$). In the group of people with NBD, the intercept was significantly positive across sentence presented in A-T argument order (LSMean = 1.150; *SE* = 0.532; 95% CI = 0.10–2.19; $t(14) = 2.16$; $p = .0486$), as well as in T-A argument order (LSMean = 1.358; *SE* = 0.526; 95% CI = 0.32–2.39; $t(17) = 2.579$; $p = .0195$). In sum, in the trials in which PWA declared to have guessed, the accuracy was better than expected by chance in A-T sentences, but not in T-A sentences. Conversely, NBD performed better than at chance level in the two argument orders.

Gaze data analysis

Post hoc analysis of the gaze data was conducted exclusively on the PWA, since the accuracy and CR data were almost completely confounded in the NBD group, as shown in Table 5. The empirical model included response accuracy (correct, incorrect), ROI (ROI 1, 2, 3, 4), and CR (1, 2, 3) as predictors. There was a two-way interaction between response accuracy and ROI, and between response accuracy and confidence rate, as well as a three-way interaction between response accuracy, confidence rate, and ROI. In addition, the model included random effects for subject (i.e., the tendency of each participant to answer correctly or incorrectly) and stimulus (i.e., the degree of difficulty that a specific stimulus may impose).

Previous studies using eye-tracking with the VWP have shown that gaze-fixations diverge during the auditory presentation of the linguistic stimuli in correctly and incorrectly comprehended trials (e.g., Arantzeta et al., 2017; Dickey & Thompson, 2009; Hanne et al., 2011). Thus, the interaction between response accuracy and ROI was expected. The detailed analysis of these results is reported by Arantzeta (2017).

In the current study we focus on the distinctive gaze-fixation pattern as a function of response accuracy and CR (see Figure 5). The following results are based on gaze-fixation differences along the visual display during the presentation of the linguistic stimuli as a whole.

In the correctly answered trials, PWA showed distinctive gaze-fixation routines in sentences judged correctly (i.e., CR = 1) and in sentences judged incorrectly (i.e., CR = 3) (β = 0.084; *SE* = 0.030; *p* = .0060). Trials rated with CR = 1 (i.e., "sure, I answered correctly") also corresponded to a different fixation pattern compared to trials in which PWA reported to answer by guessing (i.e., CR = 2) (β = 0.071; *SE* = 0.024;

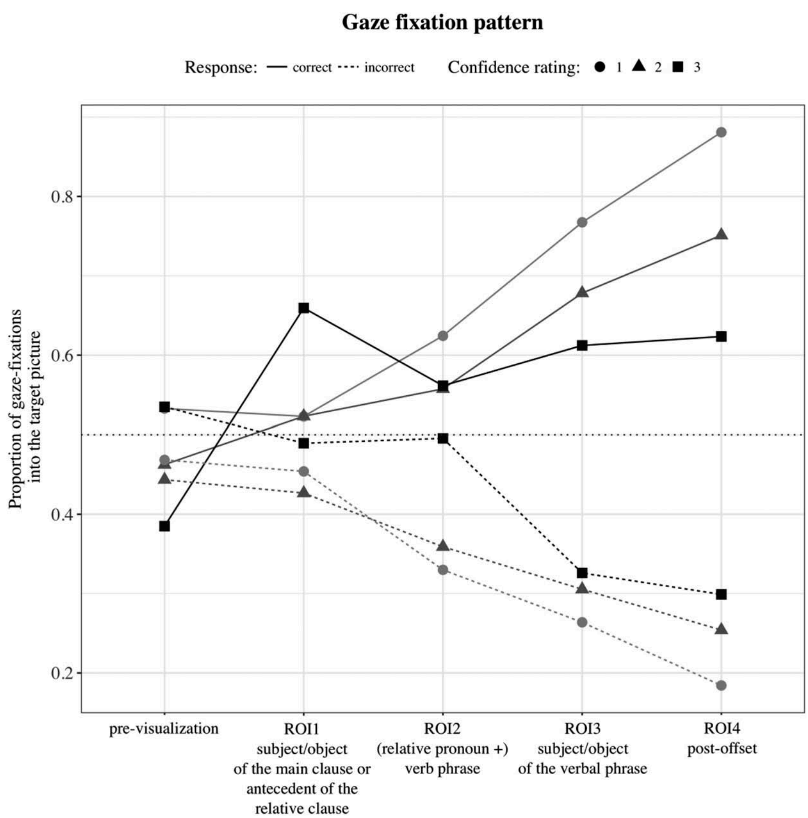

Figure 5. Fixation pattern of PWA across the visual display along the presentation of the auditory stimuli.

The gaze data is presented as a function of response accuracy (correct, incorrect) and confidence rate (1, 2, 3). Confidence rate, 1 = "sure, I answered correctly", 2 = "I do not know, guessing", 3 = "sure, I answered incorrectly". Values above .5 indicate more fixations into the target picture, whereas values below .5 indicate more fixations into the foil picture.

p = .0033). Nonetheless, the fixation pattern of the latter was indistinguishable from incorrectly judged trials (i.e., CR = 3) ($\beta = 0.013$; $SE = 0.038$; $p = .7305$). In the incorrectly answered trials, gaze-fixation pattern was significantly different between trials rated as correctly understood (i.e., CR = 1) and trials rated as incorrectly understood (i.e., CR = 3) by PWA ($\beta = -0.93$; $SE = 0.028$; $p = .0008$). There was no difference in the fixation patterns of the trials where PWA reported they had guessed (i.e., CR = 2) versus trials answered with certainty (perceived as correct, CR = 1; $\beta = -0.027$; $SE = 0.031$; $p = .3847$; or those perceived as incorrect, CR = 3; $\beta = -0.066$; $SE = 0.039$; $p = .0959$).

The results of the three-way interaction between response accuracy, ROI, and CR are detailed in Table 6.

To sum up, trials answered by guessing consistently showed a midpoint in the proportion of fixations into the target/foil picture in relation to correctly and incorrectly judged trials. This pattern was symmetric in correctly and incorrectly answered trials and it was progressive during the presentation of the linguistic stimuli.

Discussion

The questions addressed in this study are: (a) Are PWA aware of their sentence comprehension errors?; (b) Is sentence comprehension performance unconsciously mediated in PWA and NBD?; (c) To what extent do PWA answer by guessing on a task for comprehension of sentences in non-linear argument order? In addition, we conducted a post hoc gaze data analysis to explore the relationship between real-time sentence processing and the metacognitive awareness reported by PWA. This section will first discuss the sentence comprehension accuracy data to contextualise the subsequent discussion. Subsequently, the research questions will be addressed attending to the zero-correlation criterion and guessing criterion (Dienes et al., 1995).

As expected, PWA taking part in this study performed worse than NBD in the sentence comprehension task. PWA comprehended sentences with a linear Agent-Theme order better than the derived Theme-Agent order, and that applied across all the three sentence sets. The NBD group comprehended sentences in both A-T and T-A argument orders at ceiling level. Altogether, the present findings converge with those of Arantzeta et al. (2017), whose linguistic and visual materials were adapted for the present study, as well as with previous studies involving Spanish speakers with aphasia (Juncós-Rabadán et al., 2009), and other languages (e.g., Abuom et al., 2013; Burchert et al., 2003; Duman et al., 2011; Garraffa & Grillo, 2008; Meyer et al., 2012; Schwartz et al., 1980; cf. Jap et al., 2016).

Metacognitive awareness of sentence comprehension accuracy differed significantly between correct and incorrect responses. In the former, both NBD and PWA usually made a correct judgement of their answer. That is, they perceived as correct the trials answered correctly. PWA frequently rated incorrect trials as correct (i.e., false negative detection). Hence, PWA were not aware of their sentence comprehension difficulties in 78% of the incorrectly answered trials. NBD made few comprehension errors, but similarly only a few of these errors were correctly rated (i.e., 35% of the incorrectly answered trials). This indicates that not even NBD participants are always aware of their comprehension failures, possibly due to an age-related executive decline (see Martins,

Table 6. Comparison of proportion of gaze-fixations in between correctly and incorrectly answered trials across ROIs. Analysis of the gaze-data of people with aphasia as a function of confidence with which the trials were rated.

Incorrect answers	ROI 1				ROI 2				ROI 3				ROI 4			
CR = 1	β	*SE*	*t*-ratio	*p*	β	*SE*	*t*-ratio	*p*	β	*SE*	*t*-ratio	*p*	β	*SE*	*t*-ratio	*p*
ROI 1	−0.068	0.023	−2.893	.0038	−0.170	0.023	−7.166	**<.0001**	−0.313	0.023	−13.172	**<.0001**	−0.426	0.023	−17.969	**<.0001**
ROI 2	−0.192	0.023	8.123	**<.0001**	−0.294	0.023	−12.401	**<.0001**	−0.437	0.023	−18.407	**<.0001**	−0.550	0.023	−23.217	**<.0001**
ROI 3	−0.258	0.023	10.907	**<.0001**	−0.360	0.023	−15.186	**<.0001**	−0.503	0.023	−21.188	**<.0001**	−0.616	0.023	−26.002	**<.0001**
ROI 4	−0.338	0.023	14.299	**<.0001**	−0.439	0.023	−18.588	**<.0001**	−0.582	0.023	−24.599	**<.0001**	−0.696	0.023	−29.428	**<.0001**
CR = 2	β	*SE*	*t*-ratio	*p*	β	*SE*	*t*-ratio	*p*	β	*SE*	*t*-ratio	*p*	β	*SE*	*t*-ratio	*p*
ROI 1	−0.097	0.076	−1.269	.2045	−0.132	0.076	−1.712	.0869	−0.252	0.076	−3.288	**.001**	−0.325	0.076	−4.252	**<.0001**
ROI 2	−0.164	0.076	−2.165	.0304	−0.199	0.076	−2.609	.0091	−0.319	0.076	−4.200	**<.0001**	−0.392	0.075	−5.174	**<.0001**
ROI 3	−0.218	0.075	−2.891	.0039	−0.252	0.075	−3.336	**.0009**	−0.373	0.076	−4.941	**<.0001**	−0.446	0.075	−5.926	**<.0001**
ROI 4	−0.269	0.075	−3.569	**.0004**	−0.304	0.075	−4.013	**.0001**	−0.424	0.075	−5.620	**<.0001**	−0.497	0.075	−6.606	**<.0001**
CR = 3	β	*SE*	*t*-ratio	*p*	β	*SE*	*t*-ratio	*p*	β	*SE*	*t*-ratio	*p*	β	*SE*	*t*-ratio	*p*
ROI 1	−0.171	0.078	−2.168	.0302	−0.072	0.078	−0.923	.3563	−0.123	0.078	−1.568	.1169	−0.135	0.078	−1.723	.0849
ROI 2	−0.164	0.079	−2.070	.0385	−0.066	0.079	−0.834	.4045	−0.117	0.079	−1.474	.1405	−0.128	0.079	−1.627	.1037
ROI 3	−0.334	0.079	−4.199	**<.0001**	−0.235	0.079	−2.964	.0031	−0.286	0.079	−3.604	**.0003**	−0.298	0.079	−3.772	**.0002**
ROI 4	−0.361	0.079	−4.523	**<.0001**	−0.263	0.079	−3.292	**.001**	−0.314	0.079	−3.930	**.0001**	−0.325	0.079	−4.099	**<.0001**

C.R.: Confidence rate; 1 = "sure, I answered correctly"; 2 = "I don't know/guess"; 3 = "sure, I answered incorrectly". ROI = Region Of Interest; ROI 1 = subject/object of the main clause or the antecedent of the relative clause; ROI 2 = (relative pronoun +) verbal phrase; ROI 3 = subject/object of the verbal phrase; ROI 4 = Post-offset region. Significance level $p < .003$ with Bonferroni's correction for multiple comparisons.

Joanette, & Monchi, 2015, for a review) that impacts awareness of error. This is a potentially interesting experimental question, but it is beyond the immediate scope of this study.

Zero-correlation criterion

Following Dienes et al. (1995), the degree of correspondence between the accuracy and the CR was used to assess the conscious versus unconscious threshold in sentence comprehension. The presence or absence of conscious awareness is a key element to distinguish between explicit and implicit language processing, respectively. The consistency between the measures varied across the two groups, suggesting that the extent to which sentence comprehension is mediated by explicit processing varies as a consequence of neurological injury. In the PWA group, the strength of the subjective perception (i.e., CRs) in predicting the comprehension accuracy was slightly above chance, while in the NBD it was significantly larger, although not perfect.

The results suggest that PWA and NBD do not share the same consciousness threshold in sentence processing. In the PWA group, sentence comprehension is mainly mediated by implicit knowledge, suggesting little voluntary control. Consequently, PWA do not consciously perceive failure in the parsing routine. In contrast, in the NBD group, a moderate relationship between accuracy and CR suggests that explicit and implicit knowledge appear to function together in thematic-role mapping in NBD. Note that the lack of control shown in the PWA group cannot be explained by practice (Langer & Imber, 1979; Shiffrin & Schneider, 1977), since agent-first heuristics have also been proven to be deployed in healthy speakers (Bornkessel-Schlesewsky & Schlesewsky, 2013; Ferreira & Patson, 2007; Van Herten et al., 2006). Still, NBD show a moderate-good perception of the parsing mechanism, and reanalyse the sentence based on analytical routines (i.e., by processing morphological cues) when the use of heuristics fails to correctly interpret the sentence.

Guessing criterion

The analysis of the instances in which participants said they were guessing uncovered two main things. First, both PWA and NBD rarely reported that their answer in the comprehension task was based on uncertainty. This finding does not support the TDH (Grodzinsky, 1986, 1995, 2000; see Drai & Grodzinsky, 2006a, 2006b for a later revision), which states that PWA choose randomly when thematic roles need to be assigned in non-linear order, as far as we assume that "guessing" is a conscious act. In contrast to the predictions of the TDH, PWA nearly always tended to provide CRs (correct or incorrectly) based on certainty (i.e., "sure, I answered right/wrong") – they did not report they were guessing. Second, according to the guessing criterion (Dienes et al., 1995), the two groups performed above chance in the trials in which they said they had guessed, indicating unconscious knowledge of sentence processing. This effect was found in the NBD group across all argument orders and in the PWA group in sentences presented in A-T order. It suggests that PWA lack the required unconscious knowledge, or access to it, to guide the processing of sentences presented in T-A argument order.

Gaze-fixation data

Gaze data analysis showed that the CR provided by the participants trial-by-trial was closely related to the gaze-fixation pattern. The interaction between the response accuracy, confidence rate, and ROI suggests that in the trials that participants believed to have answered correctly, correct and incorrect answers show distinctive gaze-fixation pattern already during the presentation for the (relative pronoun +) verbal phrase. Thus, positive confidence is characterised by an early commitment towards one of the two potential interpretations of the sentence. In contrast, in the trials perceived as guessed or incorrectly answered, the time window in which participants commit for one of the interpretations is delayed until the noun phrase following the verb. Altogether, CR of correctness, guessing, and incorrectness show a progressive delay on sentence resolution. These results go against the TDH in that in the trials rated with positive certainty (i.e., "sure, I answered correctly"), where most of the trials are clustered, gaze-fixation pattern shows that PWA make a commitment towards the target or foil picture already at verb position. Thus, opposite to the TDH, PWA do not wait to hear the second DP to make a random choice between the two potential interpretations of the sentence.

It is not possible to identify the causal direction of the relationship between real-time sentence processing and the subjective perception of correctness in sentence comprehension. That is, we do not know whether the gaze fixations have an effect on the subsequent individual perception of the correctness of the answer, or whether different parsing routines, corresponding to different degrees of confidence, are reflected distinctively in the gaze data of PWA. Certainly, the results of the post hoc analysis on the gaze-data suggest that PWA did not randomly rate the perceived accuracy of their sentence comprehension responses, for example as consequence of misinterpretation of the task. Contrarily, the association shown between the CRs and the gaze-fixation pattern shows that the rating of perceived correctness was guided by certain unconscious and automatic regularities. So far, the data show that CRs are a sensitive measure of metacognitive awareness.

Conclusions

The study of metacognitive awareness, and hence consciousness of processing, has been widely neglected in aphasiology research regarding comprehension deficits. The current study introduced behavioural and online methods in consciousness research of PWA to get insight into the extent to which the participants' responses are mediated by conscious versus unconscious knowledge.

The relationship between comprehension accuracy and CR was studied to get insight into the degree to which sentence processing was unconsciously mediated (i.e., zero-correlation criterion; Dienes et al., 1995). We have gathered evidence suggesting that NBD use both conscious and unconscious knowledge in sentence comprehension, since their ability to perceive the correctness/incorrectness of their answer is high, although not perfect. In contrast, in the PWA group, the CR is a very poor predictor of the response accuracy, particularly because PWA do not tend to be aware of their misinterpretation of incorrectly answered trials. In contrast to NBD, PWA's sentence processing is implicit and involves little voluntary control over the assignment of thematic roles. The data show that PWA do not systematically overuse an explicit (i.e., conscious) strategy such as agent-first heuristics, as the best probabilistic guess to reach the correct interpretation of the sentences.

Regarding the guessing criterion (Dienes et al., 1995), PWA had above-chance performance in the comprehension of sentences presented in A-T order, even when they claimed to be guessing, but not in sentences presented in T-A order. This suggests that, unlike NBD participants, PWA do not show signs of the implicit knowledge required for correct thematic role assignment. Note that participants believed themselves to be guessing in very few responses. On the one hand, this indicates that cautious interpretation of the guessing criterion is needed. On the other hand, the pattern of responses is not compatible with the predictions of the TDH (Grodzinsky, 1986; 1995, 2000; see Drai & Grodzinsky, 2006a, 2006b for a later revision), which claims that PWA answer by guessing when confronted with reversible sentences in non-canonical word orders. Gaze data analysis contradicts this interpretation by showing the correctly and incorrectly trials have distinctive gaze-fixation pattern from very early on the presentation of the linguistic stimuli (see Arantzeta et al., 2017; Dickey & Thompson, 2009; Hanne et al., 2011). The current study complements this evidence by showing that the time point in which PWA make a commitment towards one or the other interpretation is related to the confidence with which the trial is rated. Although the causal direction of this relationship cannot be determined in the current study, it demonstrates the validity of CR for obtaining insight into unconscious processes in PWA.

In summary, PWA showed anosognosic behaviour towards sentence comprehension deficits, even when they were aware of their aphasic condition. These results converge with previous studies in jargon aphasia (see Rubens & Garrett, 1991). The lack of conscious cognition, which is strongly linked with voluntary control, does not suggest that they use explicit strategies to overcome their comprehension difficulties. Thus, awareness cannot be taken as a self-regulatory mechanism for therapeutic applications in PWA, as suggested for healthy individuals (Koriat, 2000). Further research is required to determine the relationship between executive functions and sentence comprehension deficits in PWA.

Acknowledgements

We are thankful to the Association of Acquired Brain Injury of Gipuzkoa (ATECE-Gipuzkoa) for their contribution in the development of this study, and especially to the speech and language therapist M. Conception Trecet. Last but not least, we are grateful to all participants and their families for their willingness to participate in the study and for their co-operation throughout.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research has been supported by the Erasmus Mundus Joint International Doctorate for Experimental Approaches to Language and Brain (IDEALAB) of the Universities of Groningen (NL), Newcastle (UK), Potsdam (DE), Trento (IT) and Macquarie University, Sydney (AU) under grant n°: 2012-1713/001-001 EMII EMJD (Framework Partnership Agreement 2012-0025), Advancing the European Multilingual Experience (AThEME FP7/grant no. 613465), Spanish Ministerio de Economía y Competitividad (FFI2015-64183), and the Basque Government (IT665-13).

References

Abuom, T. O., Shah, E., & Bastiaanse, R. (2013). Sentence comprehension in Swahili-English bilingual agrammatic speakers. *Clinical Linguistics and Phonetics*, *27*, 355–370. doi:10.3109/02699206.2013.775346

Akaike, H. (1974). A new look at the statistical model identification. *IEE Transactions on Automatic Control AC*, *19*, 716–723. doi:10.1109/TAC.1974.1100705

Arantzeta, M. (2017). *Sentence comprehension in monolingual and bilingual aphasia: Evidence from behavioral and eye-tracking methods*. Groningen, NL: Groningen Dissertations in Linguistics (GRODIL). ISBN: 978-90-367-9635-4.

Arantzeta, M., Bastiaanse, R., Burchert, F., Wieling, M., Martínez-Zabaleta, M., & Laka, I. (2017). Eye-tracking the effect of word order in sentence comprehension in aphasia: Evidence from Basque, a free word order ergative language. *Language, Cognition and Neuroscience.*, *32*, 1320–1343. doi:10.1080/23273798.2017.1344715

Bastiaanse, R., & Edwards, S. (2004). Word order and finiteness in Dutch and English Broca's and Wernicke's aphasia. *Brain and Language*, *89*, 91–107. doi:10.1016/S0093-934X(03)00306-7

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48. doi:10.18637/jss.v067.i01

Beeke, S., Wilkinson, R., & Maxim, J. (2009). Prosody as a compensatory strategy in the conversations of people with agrammatism. *Clinical Linguistics & Phonetics*, *23*, 133–155. doi:10.1080/02699200802602985

Bever, T. G. (1970). The influence of speech performance on linguistic structure. In G. B. Flores d'Arcais & W. J. M. Levelt (Eds.), *Advances in psycholinguistics*. 4–30. Amsterdam: North-Holland Publishing Co. ISBN: 072046031X.

Boland, J. (2004). Linking eye movements to sentence comprehension in reading and listening. In M. Carreiras & C. Clifton (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERPs and beyond*. 51–76. Hove, US: Psychology Press. ISBN: 1-84169-400-2.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2013). Neurotypology: Modeling crosslinguistic similarities and differences in the neurocognition of language comprehension. In M. Sanz, I. Laka, & M. K. Tanenhaus (Eds.), *Language down the garden path. The cognitive and biological basis for linguistic structures*. 241–252. Oxford: Oxford University Press. ISBN: 978-0-19-967713-9.

Burchert, F., De Bleser, R., & Sonntag, K. (2003). Does morphology make the difference? Agrammatic sentence comprehension in German. *Brain and Language*, *87*, 323–342. doi:10.1016/S0093-934X(03)00132-9

Burchert, F., Hanne, S., & Vasishth, S. (2013). Sentence comprehension disorders in aphasia: The concept of chance performance revisited. *Aphasiology*, *27*, 112–125. doi:10.1080/02687038.2012.730603

Caplan, D., & Hildebrandt, N. (1988). *Disorders of syntactic comprehension*. MIT Press: Cambridge, MA. ISBN: 0-262-03132-9.

Caplan, D., Waters, G., DeDe, G., Michaud, J., & Reddy, A. (2007). A study of syntactic processing in aphasia I: Behavioral (psycholinguistic) aspect. *Brain and Language*, *101*, 103–150. doi:10.1016/j.bandl.2006.06.225

Caramazza, A., & Zurif, E. (1976). Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and Language*, *3*, 572–582. doi:10.1016/0093-934X(76)90048-1

Chatterjee, A., Maher, L. M., & Heilman, K. M. (1995). Spatial characteristics of thematic role representation. *Neuropsychologia*, *33*, 643–648. doi:10.1016/0028-3932(94)00134-B

Cheesman, J., & Merikle, P. M. (1984). Priming with and without awareness. *Perception and Psychophysics*, *36*, 387–395. doi:10.3758/BF03202793

Cheesman, J., & Merikle, P. M. (1986). Distinguishing conscious from unconscious perceptual processes. *Canadian Journal of Psychology/Revue Canadienne De Psychologie*, *40*, 343. doi:10.1037/h0080103

Chomsky, N. (1981). *Lectures on government and binding: The Pisa Lectures*. Foris: Dordrecht. ISBN: 9070176289.

Collins English Dictionary (2012). Retrieved 2016, July 27th, from: http://www.collinsdictionary.com/dictionary/english/guess

Cooper, R. M. (1974). The control of eye spoken fixation by the meaning language. *Cognitive Psychology*, *107*, 84–107. doi:10.1016/0010-0285(74)90005-X

DeLong, E. R., DeLong, D. M., & Clarke-Pearson, D. L. (1988). Comparing the areas under two or more receiver operating characteristic curves: A nonparametric approach. *Biometrics*, *44*, 837–845. doi:10.2307/2531595

Dickey, M. W., Choy, J. J., & Thompson, C. K. (2007). Real-time comprehension of wh- movement in aphasia: Evidence from eyetracking while listening. *Brain and Language*, *100*, 1–22. doi:10.1016/j.bandl.2006.06.004

Dickey, M. W., & Thompson, C. K. (2009). Automatic processing of wh- and NP-movement in agrammatic aphasia: Evidence from eyetracking. *Journal of Neurolinguistics*, *22*, 563–583. doi:10.1016/j.jneuroling.2009.06.004

Dienes, Z., Altmann, G. T. M., Kwan, L., & Goode, A. (1995). Unconscious knowledge of artificial grammars is applied strategically. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*, 1322–1338. doi:10.1037/0278-7393.21.5.1322

Diggle, P., Heagerty, P., Liang, K., & Zeger, S. (2002). *Analysis of longitudinal data* (2nd ed.). Oxford: Oxford University Press. ISBN: 978-0-19-852484-7.

Drai, D., & Grodzinsky, Y. (2006a). A new empirical angle on the variability debate: Quantitative neurosyntactic analyses of a large data set from Broca's Aphasia. *Brain and Language*, *96*, 117–128. doi:10.1016/j.bandl.2004.10.016

Drai, D., & Grodzinsky, Y. (2006b). The variability debate: More statistics, more linguistics. *Brain and Language*, *96*, 157–170. doi:10.1016/j.bandl.2005.05.004

Dryer, M. S. (2005). Order of subject, object, and verb. In M. Haspelmath, M. Dryer, D. Gil, & B. Comrie (Eds.), *The world atlas of language structures*. 330–333. Oxford: Oxford University Press. ISBN: 0-19-925591-1.

Dryer, M. S. (2007). Word order. In T. Shopen (Ed.), *Clause structure, language typology and syntactic description*. Vol. 1, 61–131. Cambridge: Cambridge University Press. ISBN: 978-0-521-58156-1.

Duman, T. Y., Altinok, N., Özgirgin, N., & Bastiaanse, R. (2011). Sentence comprehension in Turkish Broca's aphasia: An integration problem. *Aphasiology*, *25*, 908–926. doi:10.1080/02687038.2010.550629

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research*, *24*, 409–436. doi:10.1007/BF02143160

Ferreira, F., & Patson, N. D. (2007). The good enough approach to language comprehension. *Language and Linguistic Compass*, *1*, 71–83. doi:10.1111/j.1749-818X.2007.00007.x

Garraffa, M., & Grillo, N. (2008). Canonicity effects as grammatical phenomena. *Journal of Neurolinguistics*, *21*, 177–197. doi:10.1016/j.jneuroling.2007.09.001

Goodglass, H., Kaplan, H., & Barresi, B. (2005). *The assessment of aphasia and related disorders*. (3rd ed.). (J. E. García-Albea, Spanish adaptation). Madrid:Médica Panamericana. ISBN: 84-7903-785-7.

Grodzinsky, Y. (1986). Language deficits and the theory of syntax. *Brain and Language*, *27*, 135–159. doi:10.1016/0093-934X(86)90009-X

Grodzinsky, Y. (1995). A restrictive theory of agrammatic comprehension. *Brain and Language*, *50*, 27–51. doi:10.1006/brln.1995.1039

Grodzinsky, Y. (2000). The trace deletion hypothesis and the tree-pruning hypothesis: Still valid characterizations of Broca's aphasia. *The Behavioral and Brain Sciences*, *23*, 1–21. discussion 55-64. 10.1017/S0140525X00002399

Grodzinsky, Y., Piñango, M., Zurif, E., & Drai, D. (1999). The critical role of group studies in neuropsychology: Comprehension regularities in Broca's aphasia. *Brain and Language.*, *67*, 134–147. doi:10.1006/brln.1999.2050

Hanne, S., Sekerina, I. A., Vasishth, S., Burchert, F., & De Bleser, R. (2011). Chance in agrammatic sentence comprehension: What does it really mean? Evidence from eye movements of German agrammatic aphasic patients. *Aphasiology*, *25*, 221–244. doi:10.1080/02687038.2010.489256

Jap, B. A., Martínez-Ferreiro, S., & Bastiaanse, R. (2016). The effect of syntactic frequency on sentence comprehension in standard Indonesian Broca's aphasia. *Aphasiology*, *11*, 1325–1340. doi:10.1080/02687038.2016.1148902

Juncós-Rabadán, O., Pereiro, A. X., & Souto, M. (2009). Manifestaciones de la afasia en gallego. Datos preliminares de pacientes bilingües gallego-castellano. *Revista De Logopedia, Foniatría Y Audiología*, *29*, 21–29. doi:10.1016/S0214-4603(09)70140-8

Kennedy, M. R. T., & Chiou, H. H. (2008). *What explains metacomprehension in adults with aphasia?* Jackson, WY: Clinical Aphasiology Conference. http://aphasiology.pitt.edu/archive/00001937/01/viewpaper.pdf

Kertesz, A. (2010). Anosognosia in Aphasia. In G. P. Prigatano (Ed.), *The study of anosognosia*. 113–122. New York: Oxford University Press. ISBN: 978-0-19-537909-9.

Kolk, H. H. J. (1993). A time-based approach to agrammatic production. *Brain and Language*, *50*, 282–303. doi:10.1006/brln.1995.1049

Koriat, A. (2000). The feeling of knowing: Some metatheoretical implications for consciousness and control. *Consciousness and Cognition*, *9*, 149–171. doi:10.1006/ccog.2000.0433

Langer, E. J., & Imber, L. G. (1979). When practice makes imperfect: Debilitating effects of over-learning. *Journal of Personality and Social Psychology*, *37*, 2014–2024. doi:10.1037/0022-3514.37.11.2014

MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk*. Mahwah: Lawrence Erlbaum Associates. http://childes.talkbank.org/clan/

Marshall, R. C., Rappaport, B. Z., & Garcia-Bunuel, L. (1985). Self-monitoring behavior in a case of auditory agnosia with aphasia. *Brain and Language*, *24*, 297–313. doi:10.1016/0093-934x(85)90137-3

Martins, R., Joanette, Y., & Monchi, O. (2015). The implications of age-related neurofunctional compensatory mechanisms in executive function and language processing including the new Temporal Hypothesis for Compensation. *Frontiers in Human Neuroscience*, *9*, 221. doi:10.3389/fnhum.2015.00221

Matin, E., Shao, K. C., & Boff, K. R. (1993). Saccadic overhead: Information-processing time with and without saccades. *Perception and Psychophysics*, *53*, 372–380. doi:10.3758/BF03206780

Meyer, A. M., Mack, J. E., & Thompson, C. K. (2012). Tracking passive sentence comprehension in agrammatic aphasia. *Journal of Neurolinguistics*, *25*, 31–43. doi:10.1016/j.jneuroling.2011.08.001

Mishra, R. K., Olivers, C. N. L., & Huettig, F. (2013). Spoken language and the decision to move the eyes: To what extent are language-mediated eye movements automatic? In V. S. C. Pammi & N. Srinivasan (Eds.), *Progress in brain research: Decision making - Neural and behavioural approaches* (pp. 135–149). New York: Elsevier. doi: 10.1016/B978-0-444-62604-2.00008-3.

Miyake, A., Carpenter, P. A., & Just, M. A. (1994). A capacity approach to syntactic comprehension disorder: Making normal adults perform like aphasic patients. *Cognitive Neuropsychology*, *11*, 671–717. doi:10.1080/02643299408251989

Norman, E., & Price, M. C. (2015). Measuring consciousness with confidence ratings. In M. Overgaard (Ed.), *Behavioral methods in consciousness research*. 159–180. Oxford: Oxford University Press. ISBN: 978-0-19-968889-0.

Oelschlaeger, M., & Damico, J. S. (1998). Spontaneous verbal repetition: A social strategy in aphasic conversation. *Aphasiology*, *12*, 971–988. doi:10.1080/02687039808249464

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*, 97–113. doi:10.1016/0028-3932(71)90067-4

Overgaard, M. (2015). *Behavioral methods in consciousness research*. Oxford University Press: Oxford. ISBN: 978-0-19-968889-0.

Overgaard, M., Timmermans, B., Sandberg, K., & Cleeremans, A. (2010). Optimizing subjective measures of consciousness. *Consciousness and Cognition*, *19*, 682–684. doi:10.1016/j.concog.2009.12.018

Peña-Ayala, A., & Cárdenas, L. (2015). A conceptual model of metacognitive activity. In A. Peña-Ayala (Ed.), *Metacognition: Fundaments, applications, and trends* (pp. 39–72). New York, NY: Springer International Publishing. ISBN: 978-3-319-11061-5.

R Core Team. (2015). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/

Rubens, A. B., & Garrett, M. F. (1991). Anosognosia of linguistic deficits in patients with neurological deficits. In G. P. Prigatano & D. L. Schacter (Eds.), *Awareness of deficit after brain injury: Clinical and theoretical issues*. 40–52. New York: Oxford University Press. ISBN: 0-195-05941-7.

Scheepers, C., & Crocker, M. (2004). Constituent order priming from reading to listening: A visual-world study. In M. Carreiras & C. Clifton Jr. (Eds), *The on-line study of sentence comprehension: Eyetracking, ERP and beyond*. 167–399. Hove, US: Psychology Press. ISBN: 1-84169-400-2.

Schneider, W. (2009). Automaticity and consciousness. In W. P. Banks (Ed.), *Encyclopedia of consciousness* (pp. 83–92). Elsevier:Amsterdam. ISBN: 978-0-123-73864-6

Schumacher, R., Cazzoli, D., Eggenberger, N., Preisig, B., Nef, T., Nyffeler, T., . . . Müri, R. (2015). Cue recognition and integration – Eyetracking evidence of processing differences in sentence comprehension in aphasia. *PloS ONE*, *10*, e0142853. doi:10.1371/journal.pone.0142853

Schwartz, M. F., Saffran, E., & Marin, O. S. M. (1980). The word order problem in agrammatism. *Brain and Language*, *10*, 249–262. doi:10.1016/0093-934X(80)90055-3

Seth, A. K., Dienes, Z., Cleeremans, A., Overgaard, M., & Pessoa, L. (2008). Measuring consciousness: Relating behavioral and neurophysiological approaches. *Trends in Cognitive Sciences*, *12*, 314–321. doi:10.1016/j.tics.2008.04.008

Shah, A. K., & Oppenheimer, D. M. (2008). Heuristic made easy: An effort-reduction framework. *Psychological Bulletin*, *134*, 207–222. doi:10.1037/0033-2909.134.2.207

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information precessing: II. Perceptual learning, automatic attending, and a general theory. *Psychological Review*, *84*, 127–190. doi:10.1037/0033-295X.84.2.127

Shuren, J. E., Hammond, C. S., Maher, L. M., Rothi, L. J., & Heilman, K. M. (1995). Attention and anosognosia: The case of a jargon aphasic patient with unawareness of language deficit. *Neurology*, *45*, 376–378. doi:10.1212/WNL.45.2.376

Simmons-Mackie, N. N., & Damico, J. S. (1997). Reformulating the definition of compensatory strategies in aphasia. *Aphasiology*, *11*, 761–781. doi:10.1080/02687039708250455

Stuss, D. T. (1991). Disturbance of self-awareness after frontal system damage. In G. P. Prigatono & D. L. Schacter (Eds.), *Awareness of deficits after brain injury*. 64–83. New York: Oxford University Press. ISBN: 0-195-05941-7.

Tanenhaus, M. K. (2007). Eye movements and spoken language processing. In R. P. G. Van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain*. 443–469. Amsterdam; Boston: Elsevier. ISBN: 0-080-44980-8.

Thompson, C. K., Meltzer-Asscher, A., Cho, S., Lee, J., Wieneke, C., Weintraub, S., & Mesulam, M. M. (2013). Syntactic and morphosyntactic processing in stroke-induced and primary progressive aphasia. *Behavioural Neurology*, *26*, 35–54. doi:10.1155/2013/749412

Tompkins, C. A., Scharp, V. L., & Marchall, R. (2006). Communicative value of self cues in aphasia: A re-evaluation. *Aphasiology*, *20*, 684–704. doi:10.1080/02687030500334076

Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of rules and habits*. Cambridge, MA: MIT Press. ISBN: 978-0-262-70080-1.
van Herten, M., Chwilla, D. J., & Kolk, H. H. J. (2006). When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception. *Journal of Cognitive Neuroscience*, *18*, 1181–1197. doi:10.1162/jocn.2006.18.7.1181
Verbeke, G., & Molenberghs, G. (2000). *Linear mixed models for longitudinal data*. New York: Springer. ISBN: 978-1-4419-0299-3.
Vuilleumier, P. (2004). Anosognosia: The neurology of beliefs and uncertainties. *Cortex*, *40*, 9–17. doi:10.1016/S0010-9452(08)70918-3
Wechsler, D. A. (1997). *Wechsler Adult Intelligence Scale-III (WAIS-III)*. The Psychological Corporation: San Antonio. ISBN: 0-158-98104-9.
Wilson, B. A., Cockburn, J., & Halligan, P. W. (1987). *Behavioral Inattention Test (BIT)*. England: Thames Valley Test Company. ISBN: 978-0-749-12997-2.

Appendix A.

Table A1. Individual scores (%) in the sentence–picture matching task attending to group and argument order.

	Argument order			Argument order	
Group = PWA participant	A-T	T-A	Group = NBD participant	A-T	T-A
A1	76.27	72.88	**C1**	96.67	98.31
A2	81.36	66.67	**C2**	93.33	96.67
A3	94.92	65.52	**C3**	96.61	98.33
A4	96.55	63.79	**C4**	100	91.67
A5	81.36	65.52	**C5**	94.83	95
A6	82.76	72.88	**C6**	93.22	94.92
A7	61.82	50	**C7**	91.53	96.49
A8	98.31	85.19	**C8**	96.67	100
A9	50	53.85	**C9**	96.55	100
A10	89.83	50.88	**C10**	95	96.67
A11	96.67	86.44	**C11**	96.67	96.67
A12	55.56	50.85	**C12**	96.67	96.67
A13	50	44.83	**C13**	95	95
A14	85	86.44	**C14**	98.31	93.33

PWA: people with aphasia; NBD: non-brain damaged; A-T: Agent-Theme; T-A: Theme-Agent.
For individual scores on sentence comprehension across conditions, we refer the reader to Arantzeta (2017).