Research Article

Recovery of Online Sentence Processing in Aphasia: Eye Movement Changes Resulting From Treatment of Underlying Forms

Jennifer E. Mack[a] and Cynthia K. Thompson[a,b,c]

Purpose: The present study tested whether (and how) language treatment changed online sentence processing in individuals with aphasia.
Method: Participants with aphasia (*n* = 10) received a 12-week program of Treatment of Underlying Forms (Thompson & Shapiro, 2005) focused on production and comprehension of passive sentences. Before and after treatment, participants performed a sentence-picture matching task with active and passive sentences as eye movements were tracked. Twelve age-matched controls also performed the task once each.
Results: In the age-matched group, eye movements indicated agent-first predictive processing after hearing the subject noun, followed by rapid thematic reanalysis after

hearing the verb form. Pretreatment eye movements in the participants with aphasia showed no predictive agent-first processing, and more accurate thematic analysis in active compared to passive sentences. After treatment, which resulted in improved offline passive sentence production and comprehension, participants were more likely to respond correctly when they made agent-first eye movements early in the sentence, showed equally reliable thematic analysis in active and passive sentences, and were less likely to use a spatially based alternative response strategy.
Conclusions: These findings suggest that treatment focused on improving sentence production and comprehension supports the emergence of more normal-like sentence comprehension processes.

Individuals with aphasia frequently miscomprehend sentences, particularly when reaching the correct interpretation requires sensitivity to morphosyntactic information (Caramazza & Zurif, 1976; Schwartz, Saffran, & Marin, 1980). Sentences with noncanonical argument mapping (e.g., passive sentences such as *The man was lifted by the woman*) are typically more impaired than sentences with canonical mapping (e.g., active sentences such as *The man was lifting the woman*). Such effects have been noted frequently in individuals with agrammatic aphasia, which is characterized by nonfluent and grammatically impaired language production in addition to sentence comprehension deficits (Grodzinsky, Piñango, Zurif, & Drai, 1999; Thompson et al., 2013), but have also been noted in groups with other

aphasic patterns (Caplan, Waters, & Hildebrandt, 1997). Considerable attention has been paid to the source(s) of these deficits (see review in Patil, Hanne, Burchert, De Bleser, & Vasishth, 2016), with some accounts proposing a partial loss of syntactic representations (Grodzinsky, 1986, 2000); others positing deficits in specific aspects of language processing, such as impaired thematic integration or mapping of thematic information (Meyer, Mack, & Thompson, 2012; Schwartz, Linebarger, Saffran, & Pate, 1987; Thompson & Choy, 2009); slowed lexical or syntactic processing (Burkhardt, Piñango, & Wong, 2003; Love, Swinney, Walenski, & Zurif, 2008; Piñango, 2000); or a general reduction in processing resources (Caplan, Waters, Dede, Michaud, & Reddy, 2007; Hanne, Sekerina, Vasishth, Burchert, & De Bleser, 2011; Patil et al., 2016).

In order to understand the underlying mechanisms of sentence comprehension performance, several accounts have considered not only the source of the comprehension deficit, but also alternative mechanisms used for interpretation. For example, the observation that semantically nonreversible sentences are typically less impaired than reversible sentences suggests unusually heavy reliance on semantic or plausibility information when interpreting sentences (Caramazza & Zurif, 1976; Schwartz et al., 1980;

[a]Department of Communication Sciences and Disorders, Northwestern University, Evanston, IL
[b]Cognitive Neurology and Alzheimer's Disease Center, Northwestern University, Evanston, IL
[c]Department of Neurology, Northwestern University, Evanston, IL
Correspondence to Jennifer Mack: Jennifer-mack-0@northwestern.edu

cf. Gibson, Sandberg, Federenko, Bergen, & Kiran, 2016). In order to interpret reversible sentences, some have argued that listeners with aphasia adopt an "agent-first" (or "linear") strategy in thematic role assignment, assigning the first noun phrase (NP) the agent role and the second the theme role, when syntactic representations are impaired (Grodzinsky, 1986, 2000) or are slow to form (Burkhardt et al., 2003; Love et al., 2008; Piñango, 2000). A few patients have been described in the literature for whom this appears to be the most prominent means of sentence interpretation (e.g., BL in Schwartz et al., 1980; JQ in Mitchum, Haendiges, & Berndt, 2004); these individuals show above-chance performance for canonical sentences and below-chance performance for noncanonical sentences, putatively reflecting erroneous assignment of the agent role to the first NP encountered (which, in noncanonical structures such as passives, is always a theme). Clark (2012) recently used computational modeling to examine sentence comprehension patterns in 42 people with aphasia (data from Caplan et al., 2007). The results suggested that some use an agent-first response strategy, though considerable variability was observed across individuals. Individuals with relatively mild syntactic deficits were shown to make agent-first responses more frequently as compared to those with severe deficits, suggesting that the agent-first strategy may reflect relatively preserved language abilities.

Indeed, the agent-first strategy is a common pattern seen in normal sentence parsing, reflecting parsimonious thematic role assignment in canonical sentences and requiring reanalysis for correct interpretation of noncanonical sentences in English and typologically similar languages. Unimpaired adults typically comprehend monoclausal active and passive sentences correctly; however, misinterpretations sometimes occur for noncanonical forms and reflect agent-first interpretations (Ferreira, 2003). Ferreira (2003) argued that such misinterpretations might emerge from a fast, heuristically driven agent-first parse of the sentence that, under certain circumstances, is selected over a slower, syntactically driven parse (cf. Townsend & Bever, 2001). Consistent with this hypothesis, several eye-tracking studies have demonstrated that unimpaired listeners predictively form agent-first interpretations, which are subsequently confirmed or disconfirmed by morphosyntactic information (Hanne, Burchert, De Bleser, & Vasishth, 2015; Kamide, Scheepers, & Altmann, 2003; Knoeferle, Crocker, Scheepers, & Pickering, 2005; Meyer et al., 2012). For example, in a study by Meyer et al. (2012), participants listened to active and passive sentences and selected between two pictures with reversed thematic roles. After hearing the subject noun, unimpaired older adult listeners tended to fixate the picture in which the subject was the agent, showing evidence of predictive assignment of the agent role to the sentential subject. However, after hearing the disambiguating verb morphology (i.e., *lifted/lifting*), participants rapidly fixated the correct picture, reflecting confirmation of their agent-first prediction in active sentences but disconfirmation of the prediction followed by rapid thematic reanalysis in passive sentences. Hanne et al. (2015; experiment 2) also found that

unimpaired German listeners made agent-first predictions when presented with a case-ambiguous subject NP. After presentation of the disambiguating verb agreement information, participants fixated the correct picture.

In contrast, online studies of aphasic sentence comprehension have shown an absence of agent-first eye movements in sentence–picture matching tasks, suggesting impaired thematic prediction processes (Hanne et al., 2015, experiment 2; Meyer et al., 2012). Several other aspects of the eye movements of people with aphasia have also been shown to differ from those of unimpaired controls in this task. On canonical sentence trials, they show delayed fixation to target pictures (compared to unimpaired listeners), even on correctly answered trials (Hanne et al., 2011, 2015; Meyer et al., 2012). On noncanonical trials, they fail to show a reliable fixation pattern, with variable looks to the target and distractor picture across sentence regions, consistent with impaired offline performance. It is interesting to note that, in some studies, individuals with aphasia have shown a tendency to fixate the distractor picture early in the sentence on incorrectly answered trials (e.g., Hanne et al., 2011, for canonical sentences; the results of Meyer et al., 2012, also show a trend in this direction for actives). This suggests use of a response strategy guided by eye movements in which subjects select the picture that is fixated early in the sentence.

In addition, it has been shown that spatial strategies may be used by people with aphasia in sentence comprehension tasks. For example, Chatterjee, Maher, Gonzalez Rothi, and Heilman (1995) described the performance of an individual with aphasia (WH) on a sentence–picture matching task with active and passive sentences. Two pictures with reversed thematic roles were presented (one above the other), and within each picture, the spatial location of the agent and theme (left vs. right) was manipulated. Regardless of sentence type or picture placement (top vs. bottom), WH most often selected the picture in which the sentence subject appeared on the left side. Using a similar task, Mitchum et al. (2004) described a gentleman with aphasia who had a tendency to select the picture on the bottom of the array. Some spatial strategies, particularly the strategy of left-to-right sentence interpretation described by Chatterjee et al. (1995), are also utilized in normal sentence processing. For example, studies have shown that unimpaired young adult English and Italian listeners respond more quickly in a sentence–picture verification task when the agent/subject is located on the left versus on the right side of an event picture (Chatterjee, Southwood, & Basilico, 1999; Maass & Russo, 2003).[1] These effects may be related to eye movement patterns (i.e., a tendency to inspect event pictures from left to right; see e.g., Scheepers & Crocker, 2004).

In the present study, we examined online passive sentence processing in people with aphasia and unimpaired

---

[1]These effects may be related to the left-to-right direction of written English and Italian; in contrast, native speakers of Arabic, which is written right-to-left, show facilitated sentence–picture verification when the agent appears on the right side of an event picture (Maass & Russo, 2003).

listeners by measuring eye movements as participants listened to sentences and performed a sentence–picture matching task (following Meyer et al., 2012). Individuals with aphasia also received a 12-week course of Treatment of Underlying Forms (TUF) and following treatment, online sentence processing was tested to examine changes in processing patterns associated with treatment. The treatment provided was focused on comprehension and production of passive sentences, which emphasizes thematic mapping between noncanonical and canonical forms (e.g., passive and active structures) and metalinguistic processes involved in building noncanonical structures from canonical ones (see Thompson & Shapiro, 2005, for details). It is notable that this and similar approaches (e.g., Mapping Therapy; Schwartz, Saffran, Fink, Myers, & Martin, 1994) have been shown to induce improved offline sentence production and comprehension (e.g., Jacobs & Thompson, 2000; Thompson, den Ouden, Bonakdarpour, Garibaldi, & Parrish, 2010). However, little is known about how (or if) treatment impacts mechanisms normally engaged during online sentence processing. We are aware of only one study that has addressed this issue: Using a "stop-making-sense" task, Dickey and Thompson (2004) found that listeners with aphasia who had completed TUF (with treatment-induced improvement in sentence comprehension) compared to a separate no-treatment group (with poor sentence comprehension) showed performance patterns similar to those of unimpaired listeners, suggesting that treatment may lead to changes in online sentence processing. However, in that study, performance on the task was not measured prior to treatment.

We expected that treatment would result in improved comprehension of passive sentences, as seen in previous studies (e.g., Jacobs & Thompson, 2000). In addition, we anticipated more normal-like online processing, including predictive agent-first eye movements and/or successful thematic (re-)analysis in noncanonical (as well as canonical) sentences, together with a reduction in the use of alternative (i.e., compensatory) strategies for sentence–picture interpretation following treatment (cf. Mitchum et al., 2004).

Method

Participants

Ten adults with aphasia (six men, four women) and 12 age-matched (AM) control participants without aphasia (seven men, five women) took part in the study. The AM participants were native English speakers with self-reported normal or corrected vision and hearing, and no history of speech, language, or learning disorders. The participants with aphasia were also native English speakers, and passed a screening for visual acuity as well as a pure-tone audiometric screening (40 dB, 1000 Hz).[2] All had sustained a left-hemisphere stroke at least one year prior to enrolling in the study. The two groups did not differ with respect

to age (aphasia, $M = 46.6$, $SD = 11.6$; control, $M = 52.1$, $SD = 18.0$); $p = .42$; t test) or years of education (aphasia $M = 16.4$, $SD = 2.5$; control $M = 17.1$, $SD = 2.4$; $z = 0.498$; $p = .52$; t test). The study was approved by the Institutional Review Board at Northwestern University and all participants provided informed consent.

The demographic and language testing scores for the participants with aphasia are summarized in Table 1. All exhibited language profiles consistent with mild-to-moderate agrammatic aphasia. Aphasia severity, as measured by the Aphasia Quotient from the Western Aphasia Battery–Revised (Kertesz, 2006), ranged from mild to moderately severe. Word comprehension was relatively intact, as indicated by measures of noun and verb comprehension from the Northwestern Naming Battery (Thompson & Weintraub, 2014). However, grammatical aspects of sentence production and comprehension were substantially impaired. On the Sentence Production Priming Test of the Northwestern Assessment of Verbs and Sentences (Thompson, 2011), all showed impaired production of noncanonical compared to canonical sentences (scores were unavailable for A10). Eight of 10 participants also showed greater impairments in noncanonical versus canonical sentence comprehension, as indicated by scores on the Sentence Comprehension Test of the Northwestern Assessment of Verbs and Sentences. We also obtained measures of fluency and grammatical sentence production from a narrative language sample (Cinderella story); impaired performance was defined as scores greater than two standard deviations below the mean from a normative data set from 13 AM controls (Thompson et al., 2012). All participants except one (A06) produced nonfluent speech, as indicated by reduced words per minute (AM control $M = 132.2$, $SD = 18.8$), and all participants except one (A02) also produced a reduced proportion of grammatical sentences (AM control $M = 93.0\%$; $SD = 4.4\%$).

Treatment Methods

Participants received sentence production and comprehension training within the TUF framework (Thompson & Shapiro, 2005). Ten reversible full passive sentences were selected as training targets (e.g., *The man was saved by the woman at the lake*). The sentences contained a locative adjunct to increase syntactic complexity: adjuncts have been shown to increase sentence processing difficulty in unimpaired adults and adults with aphasia (see Lee & Thompson, 2011, and references therein). Following the Complexity Account of Treatment Efficacy (Thompson, Shapiro, Kiran, & Sobecks, 2003), we hypothesized that training complex passive structures (with adjuncts) would generalize to simpler structures (e.g., passive sentences without adjuncts). Each sentence had a unique, morphologically regular verb.

In each training trial, participants were presented with an action picture (e.g., a man being saved by a woman at the lake). During the production training, the experimenter first guided the participant in building an active sentence with word cards (e.g., *The woman was saving the man at the lake*), and then identifying the thematic roles within the active

---

[2]No participants reported difficulty hearing the stimuli, which were presented at a conversational volume (60–70 dB).

Table 1. Demographic and language testing measures for people with aphasia.

ID	Age	Gender	Education (years)	Months post-onset	WAB-R AQ	NNB (AC) nouns	NNB (AC) verbs (%)	NAVS SPPT C (%)	NAVS SPPT NC (%)	NAVS SCT C (%)	NAVS SCT NC (%)	WPM	% GS
A01	51	M	16	82	69.7	100.0	100.0	66.7	0.0	66.7	66.7	42.3	9.1
A02	35	F	19	58	83.7	100.0	100.0	80.0	46.7	66.7	73.3	54.4	93.3
A03	52	F	16	73	75.8	96.7	100.0	73.3	46.7	86.7	53.3	42.8	64.7
A04	53	F	13	104	53.5	96.7	93.3	80.0	46.7	93.3	33.3	36.4	0.0
A05	53	M	21	39	74.1	100.0	100.0	60.0	0.0	80.0	26.7	32.2	6.7
A06	41	M	16	16	89.0	100.0	100.0	80.0	33.3	86.7	66.7	120.0	78.1
A07	48	M	16	17	85	83.3	86.7	66.7	13.3	80.0	40.0	49.7	45.5
A08	22	F	14	31	77.7	93.3	100.0	100.0	53.3	93.3	66.7	46.1	70.6
A09	64	M	18	19	75.6	100.0	100.0	46.7	26.7	86.7	40.0	72.2	46.7
A10	47	M	18	38	57.2	90.0	86.7	n/a	n/a	93.3	40.0	63.0	64.3
M	46.6		16.7	47.7	74.1	96.0	96.7	72.6	29.6	83.3	50.7	55.9	47.9
SD	11.6		2.4	30.4	11.5	5.6	5.7	15.1	20.8	10.1	16.7	25.5	32.6

WAB-R AQ = Western Aphasia Battery–Revised, Aphasia Quotient; NNB (AC) = Northwestern Naming Battery, Auditory Comprehension subtest; NAVS = Northwestern Assessment of Verbs and Sentences; SPPT = Sentence Production Priming Test; NC: noncanonical sentences; SCT = Sentence Comprehension Test; C = canonical sentences; WPM = words per minute; % GS = percent grammatical sentences.

sentence (action [verb], doer [agent], receiver [theme], and location). The experimenter then helped the participant build the passive sentence from the active sentence with word cards, replacing the present-participle verb form (*saving*) with a past-participle verb form (*saved*), and moving the theme to the subject position and agent to the postverbal adjunct position (and adding *by*). The sentence-building procedure thus demonstrated the NP-movement operation that is hypothesized, in some linguistic theories, to derive passive sentences. The participant then identified the thematic roles within the passive sentence, and practiced building the sentence independently with feedback from the experimenter. The steps for comprehension training were similar, except the experimenter built the sentences instead of the participant to exclude overt production processes. Each training trial began and ended with a probe (see below).

Participants were administered sentence production and comprehension probes before and after the treatment program. A priming task was used to probe sentence production in which participants were given an action picture and sentence, and then asked to produce a sentence with the same structure for a picture with reversed thematic roles. Sentence comprehension was tested using a sentence–picture matching task (A01-05) or a sentence–picture verification task (A06-10),[3] in which participants were presented with a target sentence and either selected one of two pictures that matched the sentence or indicated whether or not it matched a single picture presented, respectively. Both tasks used the same pictures depicting transitive action scenes with reversed thematic roles, and required a button press response. The test items included the 10 trained full passive sentences with adjuncts, 10 untrained items of the same structure, untrained simpler passive structures (full passives without adjuncts, e.g., *The man was saved by the woman*; short passives with adjuncts, e.g., *The man was saved at the lake*; *n* = 20 each),

as well as several untrained structures, designed to examine patterns of learning and generalization, for a total of 120 items per probe. For the purposes of the present study, we summed and compared participants' production and comprehension responses for passive sentences (full passives with adjuncts, trained and untrained items, untrained simpler passive structures) elicited in pre- versus posttreatment tests.

Participants received training for 12 weeks (except A01, who achieved training criteria in six weeks), with two sessions per week of approximately 90 minutes each. Each session consisted of approximately equal numbers of comprehension and production trials. The training sentences were rotated across experimental sessions so that all were trained with approximately equal frequency.
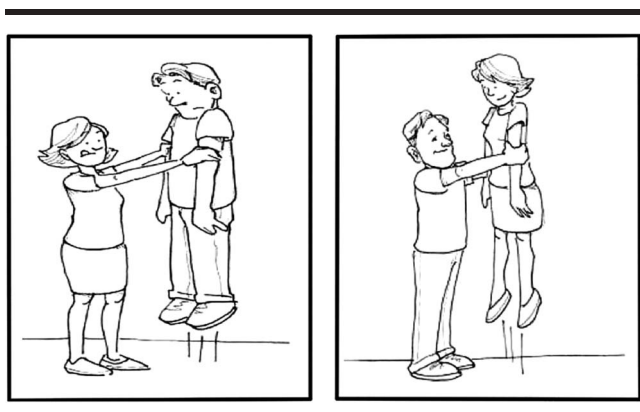
Eye-Tracking: Stimuli

Materials

The eye-tracking stimuli (verbs, sentences, and pictures) did not overlap with the materials used in TUF training. The eye-tracking stimuli consisted of semantically reversible active (e.g., *The man was lifting the woman*) and passive sentences (e.g., *The man was lifted by the woman*). There were 24 items (active/passive sentence pairs), and each contained a unique, morphologically regular verb. The active and passive versions of each item were both tested, resulting in a total of 24 trials per sentence type per participant. The sentences included only four nouns (*man, woman, boy, girl*) in order to reduce lexical processing demands. We counterbalanced the gender of the agent of the sentence across sentence types. In half of the items, the identity of the agent and the theme was reversed within active/passive sentence pairs. A male native speaker of English recorded the sentences at a normal speech rate (*M* = 4.0 syllables/second), and the sentences were matched across sentence types for length in syllables and the length of each NP in ms (*ps* > .3).

The visual stimuli for each item consisted of a pair of action pictures with reversed thematic roles (e.g., a woman

---

[3] The comprehension probe task was changed starting with A06 in order to coordinate it with a newly designed fMRI task.

Figure 1. Example visual stimulus.

lifting a man; a man lifting a woman; see Figure 1). Each action picture contained one male and one female participant. The pictures were each 5 × 6 in. and placed 4 in. apart on the computer screen. Across sentence types, we counterbalanced the location of the correct picture (left vs. right), as well as the location of the agent (left vs. right) within each picture.

Interspersed with the 48 experimental trials were 24 filler trials, in which participants heard an intransitive sentence (e.g., *The woman was dancing*) and selected between two participants performing the action (e.g., a woman dancing, a man dancing). The trial order was pseudorandomized, with the following constraints: (a) there were no more than three trials in a row of the same sentence type, (b) the active and passive trials from the same item appeared at least 10 trials apart, and (c) the correct picture appeared on the same side of the array for no more than five trials in a row.

Eye-Tracking: Procedure

Participants were seated in a dimly lit room in front of a computer monitor, with their eyes level with the center of the computer screen and their chins placed in a chinrest, to reduce motion. An ASL EYE-TRAC 6000 remote eye-tracker (Applied Science Laboratories, Bedford, MA) was used to record the location of participants' fixations (sampling rate of 60 Hz). The eye-tracker was calibrated at the beginning of each test session, with interim calibration checks every 10 trials. The testing session began with instructions and a five-item practice session. Participants were instructed to listen to each sentence and click on the matching picture. At the beginning of each trial, a fixation cross appeared on the screen. Then, the participant clicked on the cross, and the picture pair was presented. After 500 ms, the auditory sentence was presented. The picture pair remained on the screen throughout the presentation of the sentence, and after the sentence until the participant clicked on one of the pictures (or after 10 seconds if no picture was selected). The picture then disappeared and the next trial began.

For control participants, the eye-tracking experiment was conducted in a single session of approximately 20 min. For participants with aphasia, the eye-tracking experiment was conducted in four sessions, two pretreatment and two posttreatment.

In each session, participants performed the eye-tracking task in its entirety. The two pretreatment sessions were conducted a mean of 3.8 days apart (SD = 3.6) and the posttreatment sessions a mean of 4.3 days apart (SD = 4.8).

Data Analysis

Treatment Data

Mixed-effects logistic regression was used to compare participants' production and comprehension of passive sentences in the pretreatment versus posttreatment probes (lme4 package in R; Bates, Mächler, Bolker, & Walker, 2015; R Core Team, 2015). Fixed effects included treatment phase (pretreatment vs. posttreatment), structure (full passives with adjuncts–trained, full passives with adjuncts–untrained, untrained simpler passive structures), and their interaction. For the comprehension probes, we also tested for any overall effects of the probe task (sentence–picture matching, used for A01–A05, vs. sentence–picture verification, used for A06–A10), and for any interactions with structure and treatment phase. The models included by-participant and by-item intercepts and slopes. For these and all other models used in data analysis, categorical independent and dependent variables were simple-coded, allowing the intercept to be interpreted as reflecting overall significant deviations from at-chance accuracy (or for eye movements, target advantage). We performed a backward stepwise model comparison procedure to identify the best-fitting model of each data set. The model comparison steps were as follows: starting with the full model, we (a) identified the maximal converging random-effects structure, which was used for all subsequent models; (b) determined whether each fixed interaction contributed to model fit (if so, the interaction and its component main effects were retained in the model); and (c) determined whether other fixed main effects contributed to model fit (if so, they were retained in the model). Model comparison was performed with the ANOVA test in R, with a threshold of $p < .2$.

Eye-Tracking Task: Accuracy Data

In the AM controls, we did not model accuracy due to at-ceiling performance. For participants with aphasia, we analyzed the accuracy data using mixed-effects logistic regression (lme4 package in R; Bates et al., 2015; R Core Team, 2015). We first tested stability in performance between the two pretreatment test sessions. The full model of the pretreatment data contained fixed effects for sentence type (active vs. passive), target location (left vs. right), and test session and its interactions (with sentence type and target location), by-participant intercepts and slopes (sentence type, test session, target location; test session × sentence type, test session × target location) and by-item intercepts and slopes (sentence type). We then tested for cross-session stability in the posttreatment data using the same procedures. After confirming stability of performance within the two pretreatment sessions and the two posttreatment sessions, we collapsed across the two test sessions within each study phase in order to test treatment

effects. The full model of the data was the same as described for the pretreatment data, except "test session" was replaced with "phase" (pre- vs. posttreatment).

Eye Movement Data

The eye movements were first assigned to fixations on areas of interest (pictures) in the visual array, using ASL Eyenal (Applied Science Laboratories). A fixation was defined as a gaze of at least 100 ms in duration, within one degree of visual angle. All data, regardless of response accuracy, were included in eye data analyses; however, as described below, we performed analyses that examined the relationship between eye movements and accuracy. For visualization, the data were aggregated into 50-ms bins time-locked to the offset of the verb.

For statistical analyses, the sentences were split into four regions: the subject noun and auxiliary (N1 + Aux; e.g., *man was*; mean length = 646 ms), the verb (V; e.g., *lifting/lifted*; mean length = 513 ms), the postverbal noun phrase/prepositional phrase (NP/PP2; e.g., *(by) the woman*; mean length = 683 ms), and the first 1000 ms after sentence end (S End). Reaction times (RTs) were considered in the selection of the sentence regions, because a response triggered the end of a trial, and we wanted to restrict data analysis to temporal regions in which all participants (even the fastest responders) had sufficient data. Thus, the duration of the S End region approximately corresponded to the mean RT of the fastest-responding participant with aphasia (Mean RT = 1020 ms). We did not analyze the S End region in the AM control group because the mean RT of the fastest responder was soon after sentence offset (Mean RT = 60 ms).

In the first set of analyses, we examined the effects of sentence type, target location, and (for participants with aphasia only) test session and study phase, on the likelihood of fixating the target picture. For each trial and sentence region, we computed a binary measure of *target advantage*, by (a) computing the proportion of time that the participant fixated the target picture (i.e., the summed durations of fixations on the target divided by the summed durations of fixations on either picture), and (b) scoring the sentence region as *target advantage* if the proportion was greater than .5 and *distractor advantage* if it was less than .5.

We used mixed-effects logistic regression to model target advantage in each sentence region. Only regions

with a total fixation time of at least 100 ms were included in the analyses. For AM participants, we tested for effects of sentence type and target location on target advantage in each sentence region; random by-participant intercepts and slopes (sentence type, target location), and by-item intercepts and slopes (sentence type) were included. For the participants with aphasia, we first tested for stability of eye movements between the two pretreatment sessions, and then did the same for the posttreatment data. To examine treatment effects, the data were collapsed across the two sessions within each study phase. The full models of the eye data for each sentence region were the same as described for the accuracy data, and the same model comparison procedures were also used.

Eye Movements as a Predictor of Accuracy

We also examined the relationship between eye movements and accuracy in individuals with aphasia, testing whether online eye movements predicted accuracy, and if so, whether this changed from pre- to posttreatment. The purpose of these analyses was to test for treatment-related changes in response strategies (e.g., systematic associations between eye movement patterns early in the sentence and participants' responses). The pretreatment and posttreatment data were first modeled separately. For each sentence region, we modeled the effects of target advantage, sentence type, and their interaction on response accuracy, using mixed-effects logistic regression. Then, we modeled the pretreatment and posttreatment data together, testing for effects of target advantage, sentence type, study phase, and their interactions. The full model for each sentence region included random by-participant intercepts and slopes (sentence type in the pretreatment and posttreatment models; sentence type, phase, and sentence type × phase in the combined pre- and posttreatment model) and by-item intercepts and slopes (sentence type). The same model comparison procedure was used that was described for the accuracy data. For visualization purposes only, the data were plotted separately for correct and incorrect trials.

Individual Performance Patterns: Relating Sentence Production/Comprehension Abilities to Online Eye Movements

We performed two sets of correlation analyses examining the relationship between sentence processing abilities

Table 2. Treatment results (%).

Tasks	Pretreatment		Posttreatment		Post–Pre	
	M	*SD*	*M*	*SD*	*M*	*SD*
Production						
Full passives with adjuncts—trained items	5.0	12.7	77.0	25.0	72.0	23.0
Full passives with adjuncts—untrained items	7.0	10.6	75.0	31.4	68.0	28.2
Untrained passive structures	12.8	16.0	74.8	19.7	62.0	16.6
Comprehension						
Full passives with adjuncts—trained items	44.0	17.8	67.0	18.9	23.0	17.0
Full passives with adjuncts—untrained items	49.0	19.7	63.0	25.8	14.0	17.6
Untrained passive structures	49.3	15.6	66.3	24.2	17.0	23.1

Table 3. Eye-tracking task accuracy results (%).

Participants	Active		Passive	
	M	**SD**	**M**	**SD**
Age-matched controls	100.0	0.0	98.3	2.8
People with aphasia, pretreatment	77.3	12.3	52.9	16.6
People with aphasia, posttreatment	74.4	15.9	64.0	16.9

Table 4. Eye-tracking task: models of accuracy in people with aphasia.

| Variable	**z**	**p**
Pretreatment		
Intercept	5.283	< .001
Sentence type	−4.725	< .001
Target location	−1.432	.152
Test session	1.248	.212
Test session × sentence type	x	x
Test session × target location	x	x
Posttreatment		
Intercept	5.175	< .001
Sentence type	−1.513	.13
Target location	1.282	.2
Test session	x	x
Test session × sentence type	x	x
Test session × target location	x	x
Pre- vs. posttreatment		
Intercept	5.772	< .001
Sentence type	−3.308	.001
Target location	−0.907	.365
Phase	1.249	.212
Phase × sentence type	2.549	.011
Phase × target location	2.089	.037

Note. Reference levels are as follows: sentence type = active; target location = right; test session = session 1; phase = pretreatment. Cells with x indicate predictors that were not found to improve model fit and were excluded from the final, best-fitting model.

and online eye movements in participants with aphasia. The first set of correlations related pretreatment language profiles (the proportion of grammatical sentences produced in narratives, comprehension and production accuracy for all passives on the TUF probes, comprehension accuracy for passives in the eye-tracking task) to eye movement measures. The second set of correlations related treatment-related (post–pre) changes in sentence processing (comprehension and production accuracy for passives on the TUF probes, comprehension accuracy for passives in the eye-tracking task) to treatment-related changes in eye movement measures. We selected eye movement measures based on the results of the analyses of treatment effects: the proportion of target fixations for all passives in the S End region (first 1000 ms after sentence end), and the proportion of agent-first (distractor) fixations for correctly comprehended passives in the V region. Two-tailed nonparametric (Spearman) correlations were used, with correction for multiple comparisons using false discovery rate.

Results

Treatment Data

The participants with aphasia showed improved comprehension and production of passive sentences in the probe tasks posttreatment as compared to pretreatment (see Table 2). Comprehension of passive sentences significantly improved ($z = 3.375$, $p < .001$) from at-chance performance pretreatment ($p > .07$) to above-chance performance after treatment ($z = 2.518$, $p < .05$).[4] The best-fitting model of the data did not include sentence structure, indicating gains of similar magnitude with respect to acquisition of the trained items, generalization to untrained items of the same structure, and generalization to untrained simpler passive structures. The nature of the comprehension probe task (sentence–picture verification vs. sentence–picture matching) also did not contribute to model fit, indicating that this did not affect accuracy or response to treatment. Production of passives also significantly improved from pre- to posttreatment ($z = 4.567$; $p < .001$). There was no main effect of structure, or interaction between structure and phase (ps > .07), indicating a similar amount of improvement for the trained sentences, untrained items, and untrained simpler passive structures.

Eye-Tracking Task: Accuracy Data

Table 3 summarizes accuracy on the eye-tracking task, across participant groups, sentence types, and, for the participants with aphasia, study phases. The AM controls performed with at-ceiling accuracy on both active and passive sentences ($M > 98\%$), and thus we did not model accuracy for this group. Table 4 summarizes the results of the best-fitting mixed-effects regression models of the accuracy data for the participants with aphasia. Parameter estimates are provided only for variables that were found to contribute to model fit and were thus included in the best-fitting model (although not all were statistically significant predictors). Pretreatment, overall performance was above chance ($z = 5.283$, $p < .001$), and accuracy was lower for passive vs. active sentences ($z = −4.725$, $p < .001$). Models performed for each sentence type separately demonstrated that performance for active sentences was above chance ($z = 7.849$, $p < .001$), whereas accuracy for passive sentences was statistically at chance ($z = 1.039$, $p > .07$).[5] There were no other significant or marginally significant effects (ps > .07); notably, there were no significant differences in performance across test sessions, indicating stable performance. Posttreatment, participants continued to perform above chance ($z = 5.175$, $p < .001$) with no significant differences between active and passive sentences ($p > .07$) and no significant

---

[4]Due to the use of simple coding, a significant effect at the intercept indicates performance significantly different from chance, whereas no significant effect at the intercept indicates chance-level performance.

[5]The structure of these models was the same as that of the best-fitting model of the pretreatment data, except that fixed and random effects of sentence type were excluded.

Figure 2. Eye movement data in age-matched controls. N1 + Aux = subject noun and auxiliary; V = verb; NP/PP2 = postverbal noun phrase/prepositional phrase; S End = sentence end.
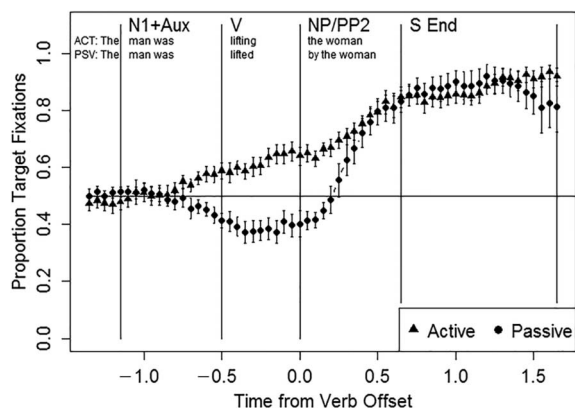
effects or interactions involving test session (*p*s > .07), indicating stable performance in the posttreatment accuracy data.

In the combined model of the pre- and posttreatment data, we focus on the effects of phase and its interactions. There was no overall effect of phase (*p* > .07). However, there was an interaction between phase and sentence type such that from pre- to posttreatment, accuracy increased

more for passives than actives (*z* = 2.549, *p* < .05). A post hoc analysis demonstrated that accuracy significantly increased for passive sentences from pre- to posttreatment (paired *t* test, *p* = .002), whereas accuracy for active sentences did not change (paired *t* test, *p* > .07). There was also an interaction between phase and target location such that from pre- to posttreatment, accuracy increased more when the target picture was on the right vs. the left (*z* = 2.089, *p* < .05). The interaction reflects the presence of a nonsignificant pretreatment trend to select the left picture, which was not evident in the posttreatment data.

Eye Movement Data

Figure 2 illustrates the proportion of target fixations over the course of the sentence in AM controls. The results of statistical analyses of the target advantage in each sentence region appear in Table 5. AM listeners were equally likely to fixate the target versus distractor pictures in the N1 + Aux and V regions (*p*s > .07). In the N1 + Aux region, they fixated the target picture more frequently when it was located on the left side of the array (*z* = −3.714, *p* < .001) and there was no effect of sentence type (*p* > .07). In the V region, they showed an agent-first fixation pattern, in which they tended to fixate the target picture in active sentences and the distractor picture in passive sentences (main effect of sentence type, *z* = −4.649, *p* < .001), with

Table 5. Models of eye movement data.

Variable	N1 + Aux		V		NP/PP2		S End	
	z	*p*	*z*	*p*	*z*	*p*	*z*	*p*
Age-matched controls								
Intercept	0.552	.581	0.784	.433	5.218	< .001	Not modeled in	
Sentence type	x	x	−4.649	< 0.001	x	x	AM participants	
Target location	−3.714	< .001	x	x	x	x		
People with aphasia, pretreatment								
Intercept	−0.987	.324	0.535	.593	1.102	.270	1.828	.068
Sentence type	−1.467	.142	−1.283	.200	−1.586	.113	−4.255	< .001
Target location	−2.938	.003	−2.426	.015	−3.298	.001	−1.189	.234
Test session	0.780	.436	0.488	.625	x	x	1.505	.132
Test session × sentence type	x	x	1.781	.075	x	x	x	x
Test session × target location	1.671	.095	1.745	.081	x	x	1.735	.083
People with aphasia, posttreatment								
Intercept	1.074	.283	1.235	.217	0.008	.993	3.902	< .001
Sentence type	−1.460	.144	x	x	x	x	x	x
Target location	−3.064	.002	−2.778	.005	−3.657	< .001	x	x
Test session	x	x	x	x	−1.435	.151	x	x
Test session × sentence type	x	x	x	x	x	x	x	x
Test session × target location	x	x	x	x	1.471	.141	x	x
People with aphasia, pre- vs. posttreatment								
Intercept	−0.190	.849	0.870	.384	0.962	.336	2.793	.005
Sentence type	x	x	x	x	−1.718	.086	−3.420	.001
Target location	−3.173	.002	−1.818	.069	−2.865	.004	−1.005	.315
Phase	1.638	.101	0.749	.454	x	x	1.233	.218
Phase × sentence type	x	x	x	x	x	x	2.720	.007
Phase × target location	x	x	−1.845	.0650	x	x	1.478	.139

Note.   Reference levels are as follows: sentence type = active; target location = right; test session = session 1; phase = pretreatment. Cells with x indicate predictors that were not found to improve model fit and were excluded from the final, best-fitting model. N1 + Aux = subject noun and auxiliary; V = verb; NP/PP2 = postverbal noun phrase/prepositional phrase; S End = first 1000 ms after sentence end.

Figure 3. Eye movement data in people with aphasia, all trials pre- and posttreatment. N1 + Aux = subject noun and auxiliary; V = verb; NP/PP2 = postverbal noun phrase/prepositional phrase; S End = sentence end.
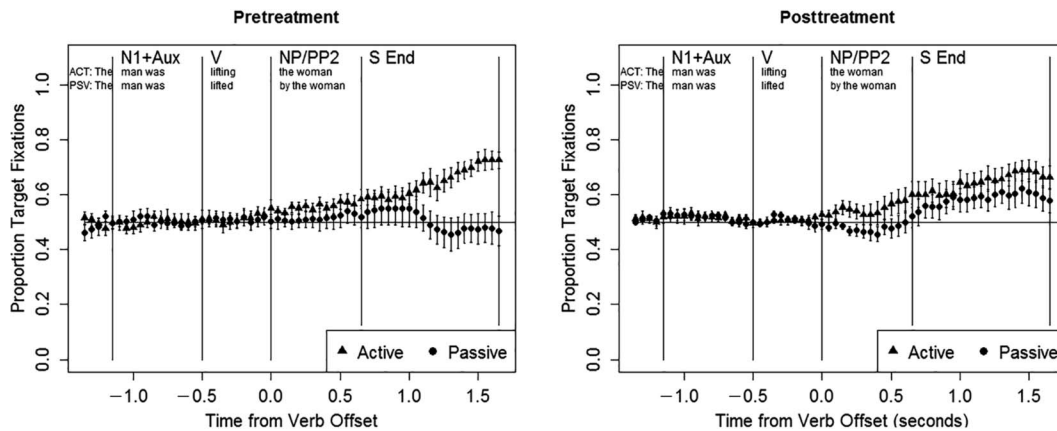
no effect of target location ($p > .07$). In the NP/PP2 region, the overall target advantage was above chance (intercept: $z = 5.218$, $p < .001$), with no significant effects of sentence type or target location ($ps > .07$).

Figure 3 illustrates the eye movement patterns for individuals with aphasia pretreatment and posttreatment, including all responses (correct and incorrect). Table 5 summarizes the statistical analyses of these data. Pretreatment, the same pattern of results emerged in all three regions of the sentence itself (N1 + Aux, V, and NP/PP2). In these regions, the target advantage did not differ from chance, with no effects of sentence type ($ps > .07$), and participants showed a tendency to fixate the left picture (N1 + Aux: $z = -2.938$, $p < .01$; V: $z = -2.426$, $p < .05$; NP/PP2: $z = -3.298$, $p = .001$), as did the AM controls in the N1 + Aux region. In the S End region (the first 1000 ms after sentence offset), the overall target advantage was marginally above chance ($z = 1.828$, $p < .07$), with a significantly greater target advantage for active than passive sentences ($z = -4.255$, $p < .001$). Models performed for each sentence type separately indicated that the target advantage was significantly above chance for active sentences ($z = 4.182$, $p < .001$) and at chance for passive sentences ($z = -0.686$, $p > .1$).[6] No other significant effects were observed ($ps > .07$). The absence of any significant effects of test session (or interactions between test session and other variables) indicates that eye movement patterns were stable across the two baseline sessions.

In the posttreatment data, the pattern of eye movements observed during the sentence was similar to the pretreatment data: the target advantage did not differ from chance, with no effects of sentence type ($ps > .07$), and participants tended to fixate the left picture (N1 + Aux: $z = -3.064$, $p < .01$; V: $z = -2.778$, $p < .01$; NP/PP2: $z = -3.657$, $p < .001$). However, in the S End region a

_____

[6]These models had the same structure as the best-fitting model of the data for this region, except that fixed and random effects of sentence type were excluded.

different pattern emerged: the target advantage was significantly above chance ($z = 3.902$, $p < .001$), with no significant effects of sentence type ($ps > .07$). There were no significant main effects or interactions involving test session ($ps > .07$), indicating that eye movement patterns were stable across the two posttreatment test sessions.

In the direct comparison of the pretreatment to the posttreatment eye movements, focusing on the effects of phase and its interactions, significant effects emerged only in the S End region. In that region, there was a significant interaction between phase and sentence type, such that from pre- to posttreatment, the target advantage increased more for passive than active sentences ($z = 2.720$, $p < .01$). In addition, there was a marginally significant interaction between phase and target location in the V region ($z = -1.845$, $p = .065$), indicating a greater tendency to fixate the leftmost picture posttreatment. There were no other significant or marginally significant effects ($ps > .07$).

Eye Movements as Predictors of Accuracy

The next set of analyses examined eye movements during the sentence as a predictor of accuracy. The results of statistical analyses appear in Table 6, and for visualization purposes, the time course of eye movements is presented for correct trials (Figure 4) and incorrect trials (Figure 5). Because the effects of sentence type, study phase, target location, and their interactions on accuracy were already reported (Table 4), here we focus just on the effects of eye movements (target advantage) and their interactions with sentence type and study phase. Pretreatment, the target advantage was a significant or marginally significant predictor of accuracy in all regions (N1 + Aux: $z = 1.911$, $p = .056$; V: $z = 2.165$, $p < .05$; NP/PP2: $z = 5.783$, $p < .001$). In other words, participants were more likely to select the target picture when they fixated on it, even early in the sentence. There were no interactions between target advantage and sentence type ($ps > .07$). These effects are evident in the left panels of Figures 4 and 5; starting from the beginning of

Table 6. Eye movements as a predictor of accuracy in people with aphasia.

Variable	N1 + Aux		V		NP/PP2	
	z	p	z	p	z	p
Pretreatment						
Intercept	5.251	< .001	5.140	< .001	5.045	< .001
Target advantage	1.911	.056	2.165	.030	5.783	< .001
Sentence type	−3.675	< .001	−4.427	< .001	−4.477	< .001
Target advantage × sentence type	x	x	x	x	x	x
Posttreatment						
Intercept	5.551	< .001	5.762	.000	5.683	.000
Target advantage	−0.206	.836	0.674	.500	3.369	.001
Sentence type	−1.483	.138	−1.520	.129	−1.580	.114
Target advantage × sentence type	−2.367	.018	−2.931	.003	−1.510	.131
Pre- vs. Posttreatment						
Intercept	5.629	< .001	5.728	< .001	5.770	< .001
Target advantage	2.456	.014	2.473	.013	6.594	< .001
Sentence type	−2.207	.027	−2.920	.004	−3.031	.002
Phase	0.866	.387	1.312	.189	0.959	.338
Target advantage × sentence type	−1.697	.090	−1.792	.073	−1.419	.156
Target advantage × phase	−2.001	.045	−1.325	.185	−1.662	.097
Sentence type × phase	2.708	.007	2.156	.031	1.954	.051
Target advantage × sentence type × phase	x	x	−1.844	.065	x	x

Note. Reference levels are as follows: target advantage = 0 (distractor advantage); sentence type = active; phase = pretreatment. Cells with x indicate predictors that were not found to improve model fit and were excluded from the final, best-fitting model. N1 + Aux = subject noun and auxiliary; V = verb; NP/PP2 = postverbal noun phrase/prepositional phrase.

the sentence, participants tended to fixate the target picture in correct trials and the distractor picture in incorrect trials.

Posttreatment, the relationship between online eye movements and accuracy was quite different. In the N1 + Aux and V regions, there were no main effects of target advantage ($ps > .07$), but significant interactions were observed between target advantage and sentence type (N1 + Aux: $z = −2.367$, $p < .05$; V: $z = −2.931$, $p < .01$). These effects indicate that participants were more likely to respond correctly when they fixated the target picture in active sentences, or the distractor picture in passive sentences (i.e., when they fixated the picture in which the

subject was the agent). These effects are illustrated in the right panels of Figures 4 and 5. On correct trials, participants tended to initially fixate the target picture in active sentences and the distractor in passive sentences, whereas the reverse was true in incorrect trials. In the NP/PP2 region, the overall target advantage was a significant predictor of accuracy ($z = 3.369$, $p = .001$), but there was no interaction between target advantage and sentence type ($p > .07$).

In the combined model of the pre- and posttreatment eye movements as predictors of accuracy, we focused on interactions between target advantage and study phase. In the N1 + Aux region, the target advantage more strongly

Figure 4. Eye movement data in people with aphasia, correct trials pre- and posttreatment. N1 + Aux = subject noun and auxiliary; V = verb; NP/PP2 = postverbal noun phrase/prepositional phrase; S End = sentence end.
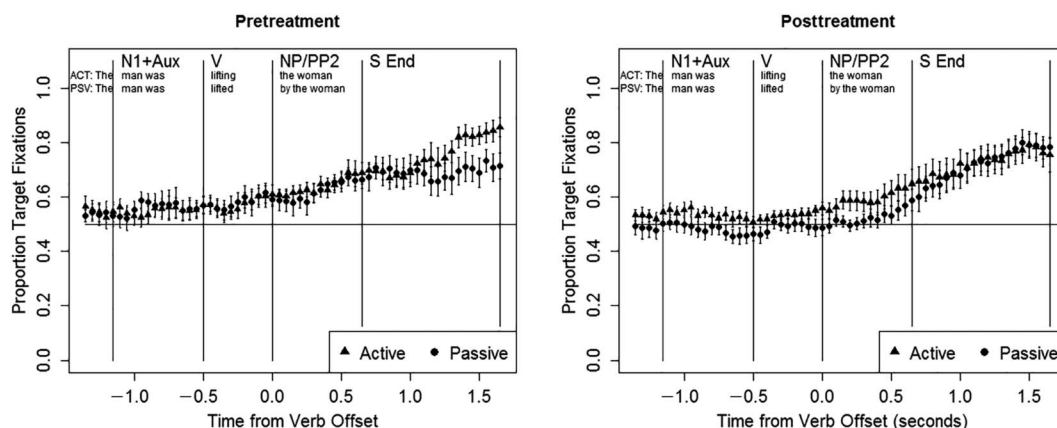
Figure 5. Eye movement data in people with aphasia, incorrect trials pre- and posttreatment. N1 + Aux = subject noun and auxiliary; V = verb; NP/PP2 = postverbal noun phrase/prepositional phrase; S End = sentence end.
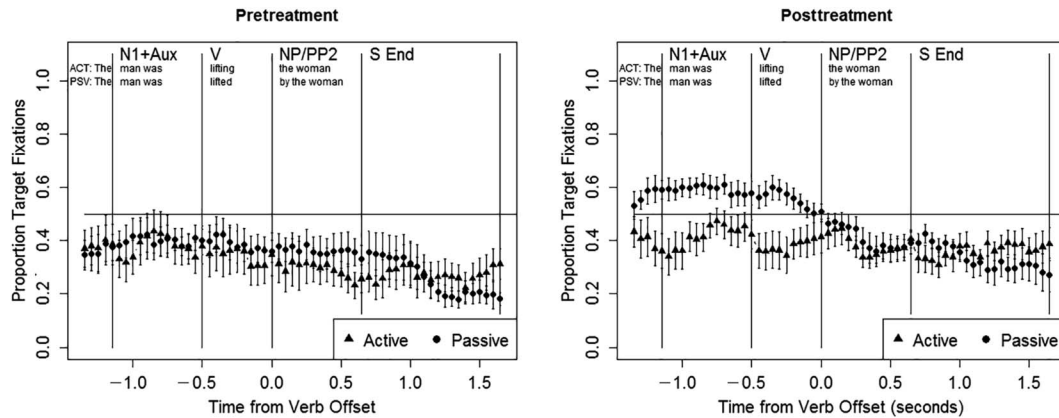
predicted accuracy pretreatment as compared to post-treatment (target advantage × phase interaction, $z = -2.001$, $p < .05$). In other words, participants were more likely pretreatment than posttreatment to select the picture that they were fixating during the N1 + Aux region. In the V region, there was a marginally significant three-way interaction between target advantage, sentence type, and phase ($z = -1.844$, $p = .065$). This interaction reflects the fact that pretreatment, the target advantage was an overall predictor of accuracy, but posttreatment, target advantage interacted with sentence type (see preceding paragraphs). There were no other significant interactions involving target advantage and study phase ($ps > .07$).

Individual Performance Patterns: Sentence Production/Comprehension Abilities and Online Eye Movements

The results of correlation analyses relating pretreatment language profiles to eye-tracking patterns appear

in Table 7. The proportion of target fixations for passives in the S End region was positively correlated with accuracy on the TUF passive sentence comprehension probes ($r = .83$, uncorrected $p < .01$) and accuracy for passive trials on the eye-tracking task ($r = .8$, uncorrected $p = .01$); the latter correlation, however, was not significant with FDR correction for multiple comparisons. No other significant or marginally significant correlations between pretreatment language profiles and eye movement patterns were found.

Table 7 also summarizes the correlations between treatment-related changes in sentence production and comprehension accuracy and treatment-related changes in eye movements. A marginally significant positive correlation was found between increased agent-first fixations in the V region for passives (correct trials) and increased accuracy on the TUF passive sentence comprehension probes ($r = .62$, $p = .05$). In addition, increased target fixations in the S End region for passives (all trials) were positively correlated with increased accuracy on passive trials in the eye-tracking task

Table 7. Correlations between language task performance and eye-tracking measures.

	Eye-tracking measures			
	V: A1 fixations (correct passive trials)		S End: target fixations (all passive trials)	
Language measures	*r*	*p*	*r*	*p*
Pretreatment correlations				
% grammatical sentences	−.18	.63	.22	.53
TUF passive probes: comprehension	−.1	.79	.83	< .01*
TUF passive probes: production	−.13	.72	.35	.31
Eye-tracking task accuracy (passives)	−.42	.23	.8	.01
Post- and Pretreatment correlations				
TUF passive probes: comprehension	.62	.05	.16	.65
TUF passive probes: production	.28	.44	.15	.67
Eye-tracking task accuracy (passives)	.26	.46	.67	.03

V = verb; A1 = agent-first; S End = first 1000 ms after sentence end; TUF = Treatment of Underlying Forms. Uncorrected *p*-values are listed;

* = significant ($p < .05$) with FDR-correction for multiple comparisons.

Figure 6. Individual treatment results: Accuracy and eye movements. The *x*-axis indicates the change in the proportion of correct passive sentences in the eye-tracking task from pre- to posttreatment; the *y*-axis indicates the change in target advantage for passive sentences in S End from pre- to posttreatment. The gray vertical and horizontal lines indicate cutoffs for improvement in accuracy (to the right of the line) and target advantage (above the line), respectively; the regression line appears in black.
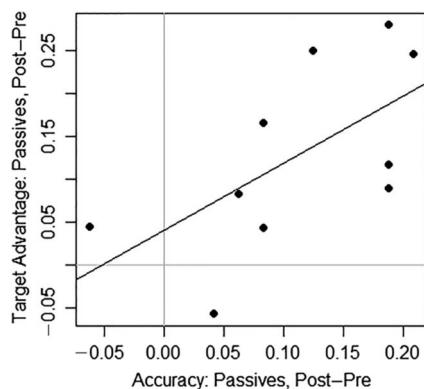
($r = .67$, $p = .03$). Neither correlation, however, was significant when adjusted for multiple comparisons. No other significant or marginally significant correlations were found. Figure 6 plots individual treatment-related changes in passive sentence comprehension accuracy (eye-tracking task) and eye movements for passive sentences. In the eye-tracking task, nine of 10 participants showed an increase in comprehension accuracy for passive sentences from pre- to posttreatment (range: .04 to .21 increase in the proportion of correct responses); and one participant showed a small decline in performance (.06). For eye movement measures, nine of 10 participants similarly showed an increase in target advantage for passives during the S End region (range: .04 to .28 increase in the proportion of trials with a positive target advantage), whereas one participant showed a decrease in the target advantage (.06).

Discussion

The aim of the present study was to examine the effects of language treatment on sentence comprehension accuracy and online sentence processing in aphasia. Ten people with aphasia received language treatment focusing on the production and comprehension of passive sentences, and performed a sentence–picture matching task pre- and posttreatment. We tested treatment-related changes in accuracy and online eye movements. Our main research question was whether treatment-induced improvements in sentence comprehension would be reflected in eye movement patterns and, if so, whether eye movement patterns at post- compared to pretreatment would show a re-emergence of more normal-like sentence processing and/or the reduction or elimination of alternative strategies used prior to treatment.

In order to identify normal online processing patterns, 12 AM adults performed the eye-tracking task. The AM

participants showed at-ceiling accuracy for both active and passive sentences ($M > 98\%$). Spatial factors guided their eye movements in the earliest region of the sentence (N1 + Aux), in which they tended to fixate the left-most picture. This finding is consistent with the observation that English listeners tend to process visual representations of events from left-to-right (Chatterjee et al., 1999; Maass & Russo, 2003). Starting in the V region, linguistic rather than spatial factors were the only significant predictors of eye movements. In this region, AM listeners showed an online agent-first effect, preferentially fixating the picture in which the subject noun was depicted as the agent (the "agent-first picture"). This replicates several previous findings of online agent-first effects in unimpaired listeners (Hanne et al., 2015; Kamide et al., 2003; Knoeferle et al., 2005; Meyer et al., 2012). In the NP/PP2 region, after presentation of the verbal morphology that distinguishes active from passive sentences, participants showed a strong tendency to fixate the target picture, with no residual effects of sentence type. This indicates that unimpaired speakers can rapidly map from morphosyntactic to thematic information, confirming the initial agent-first interpretation in the case of active sentences and performing a rapid thematic reanalysis in the case of passive sentences.

Before treatment, the people with aphasia showed impaired production and comprehension of passive sentences. On the eye-tracking sentence–picture matching task, participants performed more accurately on active as compared to passive sentences, consistent with several previous studies (Bastiaanse & Edwards, 2004; Burchert & De Bleser, 2004; Caplan et al., 1997; Grodzinsky et al., 1999; Meyer et al., 2012; Thompson et al., 2013). This suggests that before treatment, the people with aphasia were better at mapping thematic roles in canonical versus noncanonical structures. Accuracy was above chance for active sentences but statistically at chance for passive sentences, suggesting failure to apply an across-the-board agent-first strategy, which would have resulted in below-chance performance on passives. The people with aphasia were also numerically, but not significantly, more accurate when the target picture appeared in the left position, suggesting that some individuals may have been guided by spatial considerations in their responses (cf. Chatterjee et al., 1995; Mitchum et al., 2004).

The pretreatment eye movement data indicated a tendency to fixate the left-most picture throughout the sentence (N1 + Aux, V, and NP/PP2 regions), with an overall equal likelihood of fixating the target and distractor pictures during these regions in both sentence types. Thus, individuals with aphasia showed spatially guided online looking patterns similar to those seen in AM controls in the N1 + Aux region but, unlike AM controls, the pattern persisted throughout the sentence. The absence of an effect of sentence type during the sentence indicates that the participants with aphasia did not show an online agent-first strategy pretreatment, consistent with previous findings (Hanne et al., 2015; Meyer et al., 2012). A different pattern of eye movements was observed in the S End region (the

first 1000 ms after sentence end), which mirrored accuracy patterns: the target advantage was stronger in active sentences (significantly above chance) than in passive sentences (at chance). This pattern of eye movements is similar to results of previous studies, in which people with aphasia have demonstrated a delayed but ultimately reliable target advantage in canonical sentences (Hanne et al., 2011, 2015; Meyer et al., 2012), but no reliable target advantage in noncanonical sentences (Meyer et al., 2012).

It is notable that there were no significant differences in accuracy or eye movement patterns between the two pretreatment test sessions, suggesting that performance was stable in the baseline phase. Moreover, in a previous study we examined the test–retest reliability of accuracy and eye movement patterns for unimpaired adults and listeners with aphasia, using the same task as in the present study (i.e., sentence–picture matching; Mack, Wei, Gutierrez, & Thompson, 2016). For the people with aphasia, we found generally good-to-excellent reliability for accuracy and online eye movements, suggesting that the task elicited stable performance patterns across test sessions.

It is interesting to note that pretreatment eye movements during the sentence were related to response accuracy. On correct trials, participants tended to fixate the target picture early in the sentence, whereas on incorrect trials, they tended to fixate the distractor picture. A tendency to fixate the distractor picture early in the sentence on incorrect trials has been observed in previous studies (Hanne et al., 2011; trend for actives in Meyer et al., 2012). One possible interpretation of this effect is that the people with aphasia tended to judge the picture they were fixating (often the left picture) to match the sentence, perhaps on the basis of lexical content, without reliably detecting thematic mismatches. This *initial fixation bias* effect may be related to the *yes bias* effect, which has been reported for some listeners with aphasia in sentence–picture verification tasks, in which participants are asked to indicate whether the meaning of a sentence matches a single picture (e.g., Mitchum et al., 2004).

The pretreatment eye movement patterns are not in line with the trace deletion hypothesis, which suggests that people with aphasia adapt to their syntactic (representational) deficit by using an agent-first strategy (Grodzinsky, 1986, 2000), which precludes ability to comprehend noncanonical sentences accurately. Further, the results are inconsistent with the hypothesis that slowed dependency formation underlies impaired comprehension of noncanonical sentences (e.g., Burkhardt et al., 2003; Love et al., 2008). These studies have observed delays in dependency formation in listeners with aphasia ranging from 150 ms (unaccusatives; Burkhardt et al., 2003) to 500 ms (object-relatives; Love et al., 2008, experiment 1), as compared to unimpaired controls (but see Dickey, Choy, & Thompson, 2007, who found no delays in dependency formation in listeners with aphasia). Thus, if slowed dependency formation were the source of passive comprehension deficits, we would expect listeners with aphasia to show an increase in target fixations for passives within 500 ms of controls. However, in the present study, controls showed a rapid increase in target fixations for passives at the beginning of NP/PP2, whereas participants with aphasia did not do so within 1000 ms of sentence offset—beyond the window in which one would expect slowed dependency formation to take place.

Instead, the present results are in line with the thematic integration account of sentence comprehension impairments in aphasia (e.g., Mack, Ji, & Thompson, 2013; Meyer et al., 2012; Thompson & Choy, 2009, in which the term *lexical integration* is used). Like mapping deficit accounts of impaired sentence comprehension (e.g., Schwartz, Linebarger, Saffran, & Pate, 1987), the thematic integration account entails deficits in the mapping between thematic roles and syntactic representations. In addition, the thematic integration account focuses on the specific online processes that underlie thematic mapping deficits, such as thematic prediction. On this account, the pretreatment eye movement patterns in the present study can be explained as follows: Listeners with agrammatic aphasia are impaired in making thematic predictions online, which results in an absence of agent-first fixation patterns and a prolongation of the normal tendency to fixate the left-most picture early in the sentence. After presentation of the verb in active sentences, listeners with aphasia largely are accurate in thematic role mapping; however, this process is delayed due to the absence of thematic prediction, resulting in an increase in target fixations only after sentence end (see, e.g., Van Petten & Luka, 2012, for discussion of the benefits of prediction in terms of increasing sentence processing speed). For passive sentences, however, thematic role mapping is substantially impaired, resulting in at-chance fixation patterns throughout the sentence. The initial fixation bias effect (i.e., the tendency to select the picture fixated early in the sentence, even if it mismatches the sentence) may also be related to impaired thematic integration. In normal sentence processing, prediction has shown to be an essential component of detecting unexpected linguistic material and linguistic violations (e.g., Kuperberg & Jaeger, 2016). Thus, impaired thematic prediction in aphasia may contribute to deficits in detecting thematic mismatches, contributing to the initial fixation bias effect.

After receiving 12 weeks of treatment within the TUF framework (Thompson & Shapiro, 2005), participants showed improved production and comprehension of the trained structure (passives) on the TUF probes. For sentence production, the magnitude of improvement for trained passives (post–pre treatment $M = 72\%$) and untrained passives (untrained full passives with adjuncts, 68%; untrained passive structures, 62%) was similar to the results of previous TUF studies. A meta-analysis of TUF production studies by Dickey and Yoo (2010) indicated a mean gain of 76.9% ($SD = 20.7\%$) for trained items and 57.9% ($SD = 27.2\%$) for untrained items. Relatively few studies have examined the effects of TUF treatment for sentence comprehension, particularly for passives. However, the magnitude of improvement for comprehension of passives (trained full passives with adjuncts, $M = 23\%$; untrained full passives with adjuncts, $M = 14\%$; untrained passive structures, $M = 17\%$) was similar

to that reported in a previous TUF study in which passive sentence comprehension was trained (Jacobs & Thompson, 2000; an overall gain in sentence comprehension accuracy of approximately 18% for P2 and 25% for P4[7]). Although these gains in comprehension are relatively modest (e.g., smaller than the threshold of 33% that Kiran et al. (2012) propose as a meaningful improvement in sentence comprehension), notably, comprehension of passives improved from at-chance pretreatment to above-chance posttreatment, indicating substantial changes in sentence processing ability (also suggested by Kiran et al., 2012). Within each domain (production and comprehension), the magnitude of improvement was similar between trained and untrained items/passive structures, indicating successful acquisition and generalization of passives.

On the eye-tracking task, accuracy also improved for passive sentences, with improvement noted for nine of 10 participants, whereas performance on active sentences did not change at the group level. This indicates that training was successful in promoting improvement in the assignment of thematic roles within noncanonical sentences. The lack of change in comprehension accuracy (and eye movements) for actives may reflect the fact that thematic mapping was relatively preserved pretreatment for this structure. It is notable that the items used for training and those included in the eye-tracking task did not overlap and thus, changes in the eye-tracking task for passives reflect response (across-item) generalization. In addition, from pre- to posttreatment, the people with aphasia showed an increase in accuracy when the target picture was located on the right side of the array, suggesting a reduction in the use of a spatial bias evident in the pretreatment data (a trend, though nonsignificant, to select the left-most picture). This supports the idea, suggested by Mitchum et al. (2004), that training thematic role mapping can reduce the use of nonlinguistic response strategies.

From pre- to posttreatment, changes were also noted in eye movement patterns. Because we found stable performance for both accuracy and eye movements during the pretreatment phase (as well as within the posttreatment phase), these changes can be attributed to treatment, rather than practice effects. The most notable change in the overall pattern of eye movements, combining correct and incorrect trials, was evident in the first 1000 ms after sentence end. During the sentence (N1 + Aux, V, and NP/PP2 regions), eye-movement patterns showed little change. At posttreatment participants continued to direct fixations to the left picture during the N1 + Aux, V, and NP/PP2 regions, with no effects of sentence type and an overall equal likelihood of fixating the target and distractor pictures. However, treatment-related changes in eye movements were observed within 1000 ms of sentence end (i.e., in the S End region): posttreatment, participants showed an overall

tendency to fixate the target picture at the end of the sentence, with no effects of sentence type. We observed an increase in target fixations for passives in the S End region for nine of 10 participants, and these changes were closely related to improvements in passive comprehension accuracy on the eye-tracking task: the individuals who showed larger gains in accuracy also exhibited a greater increase in target fixations. Treatment-related changes were also observed in the relationship between online eye movements and accuracy. Following treatment, the initial fixation bias noted in the pretreatment data disappeared, again suggesting a treatment-induced reduction in the use of nonlinguistic response strategies (Mitchum et al., 2004). Instead, on correct trials, participants tended to make online fixations to the agent-first picture (N1 + Aux and V regions), whereas on incorrect trials, participants tended to fixate the picture in which the subject was the theme.

What do these patterns reveal about the mechanisms of TUF? The eye-movement results suggest that TUF changed online processing in two ways. First, treatment resulted in an emergence of thematic prediction in correct (but not incorrect) trials, as indicated by a treatment-related increase in agent-first fixation patterns during the V region, as seen in healthy control participants. TUF emphasizes canonical thematic role mapping and the relation between noncanonical structures and canonical structures, thereby encouraging canonical thematic role order as the starting point for successful sentence comprehension. It is notable that treatment-related increase in agent-first fixation patterns was associated with improvement on the TUF sentence comprehension probes for passive structures, indicating the importance of thematic prediction for supporting gains in sentence comprehension.

Second, for passive sentences, the participants with aphasia evinced an increase in target fixations in the S End region. Although occurring prior to the S End region in healthy listeners (i.e., in the NP/PP2 region), this pattern reflects successful thematic analysis once the verb is encountered. One primary component of TUF focuses on thematic mapping between morphosyntactic structure (e.g., past-participle verb for passive structures and *by*-phrase in full passives) and thematic structure. The delay seen in the participants with aphasia, however, may reflect residual thematic prediction deficits (cf. the finding that agent-first eye movements were observed only in correct trials posttreatment), which slow subsequent thematic mapping processes. Nevertheless, treatment not only resulted in eye-movement shifts reflecting improved thematic mapping processes, it also was associated with increased passive sentence comprehension accuracy in the eye-tracking task. In future work, it may be possible to develop modules of TUF that directly train thematic prediction with the aim of fostering more robust improvements in sentence processing.

The results of the present study also suggest that the mechanisms underlying thematic mapping may be at least somewhat dissociable across language domains (production and comprehension). The participants in the present study

---

[7]These approximate values were computed by comparing the mean passive sentence comprehension accuracy for the last two pretreatment baseline probes to the last two probes during treatment.

exhibited considerable variability with respect to the severity of language impairments in each domain, as well as in response to TUF treatment, allowing us to examine the relationship between sentence comprehension and production abilities and online eye movements. Although we found strong relationships between sentence comprehension and online eye movements in the pretreatment data, we did not observe any significant relationships between sentence production ability and eye movements during comprehension. A similar pattern of results was observed for the analyses examining treatment effects: treatment-related gains in sentence production were not correlated with changes in eye movements. These results are consistent with previous findings indicating that TUF may have different effects across modalities (Jacobs & Thompson, 2000; Schröder, Burchert, & Stadie, 2015). This suggests that although one of the principles of TUF is to train abstract linguistic representations and processes (which should be shared across domains), TUF may also have domain-specific effects (e.g., promoting thematic prediction in sentence comprehension). Further work can investigate this question by comparing changes in online processing during sentence comprehension and production in response to TUF (or similar treatment approaches).

Conclusion

The present study replicated the findings of previous studies, that online comprehension of noncanonical sentences is impaired in agrammatic aphasia, and has provided preliminary evidence that language treatment affects this ability. Further research is clearly needed in order to understand which processes are likely to change (and which are not) as a result of treatment, the effects of different types of treatment on online processing, and individual differences with respect to treatment-related changes in sentence processing. In addition, future research considering the relation between online processing and neural changes associated with treatment will likely provide a fuller picture of neuro-cognitive mechanisms underlying recovery. Few studies have examined treatment-related neural changes in sentence processing and, notably, results of these studies have shown variability in posttreatment neural recruitment patterns (see Thompson et al., 2010; Wierenga et al., 2006). However, to our knowledge, no studies have examined the relationship between treatment-induced changes in online computational strategies and neural changes. Indeed, one of the driving questions pertaining to language recovery is whether treatment results in emergence of more normal-like online computational strategies or if it induces use of alternative, compensatory, processing strategies (Kiran, 2012; Saur & Hartwigsen, 2012; Thompson & den Ouden, 2008; Turkeltaub, Messing, Norise, & Hamilton, 2011). Thus, in future research, it will be important to examine individual differences in neural recruitment patterns and how (or if) they reflect changes in online processing strategies.

References

Bastiaanse, R., & Edwards, S. (2004). Word order and finiteness in Dutch and English Broca's and Wernicke's aphasia. *Brain and Language, 89*(1), 91–107.

Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software, 67*(1), 1–48.

Burchert, F., & De Bleser, R. (2004). Passives in agrammatic sentence comprehension: A German study. *Aphasiology, 18,* 29–45.

Burkhardt, P., Piñango, M. M., & Wong, K. (2003). The role of the anterior left hemisphere in real-time sentence comprehension: Evidence from split intransitivity. *Brain and Language, 86,* 9–22.

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

Caplan, D., Waters, G. S., & Hildebrandt, N. (1997). Determinants of sentence comprehension in aphasic patients in sentence–picture matching tasks. *Journal of Speech, Language, and Hearing Research, 40,* 542–555.

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

Chatterjee, A., Maher, L. M., Gonzalez Rothi, L. J., & Heilman, K. M. (1995). Asyntactic thematic role assignment: The use of a temporal–spatial strategy. *Brain and Language, 49,* 125–139.

Chatterjee, A., Southwood, M. H., & Basilico, D. (1999). Verbs, events and spatial representations. *Neuropsychologia, 37,* 395–402.

Clark, D. G. (2012). Storage costs and heuristics interact to produce patterns of aphasic sentence comprehension performance. *Frontiers in Psychology, 3*(135), 1–19.

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

Dickey, M. W., & Thompson, C. K. (2004). The resolution and recovery of filler-gap dependencies in aphasia: Evidence from on-line anomaly detection. *Brain and Language, 88,* 108–127.

Dickey, M. W., & Yoo, H. (2010). Predicting outcomes for linguistically specific sentence treatment protocols. *Aphasiology, 24,* 787–801.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47,* 164–203.

Gibson, E., Sandberg, C., Federenko, E., Bergen, K., & Kiran, S. (2016). A rational inference approach to aphasic language comprehension. *Aphasiology, 30,* 1341–1360.

Grodzinsky, Y. (1986). Language deficits and the theory of syntax. *Brain and Language, 27,* 135–159.

Grodzinsky, Y. (2000). The neurology of syntax: Language use without Broca's area. *Behavioral and Brain Sciences, 23,* 1–21; discussion 21–71.

Grodzinsky, Y., Piñango, M. 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.

Hanne, S., Burchert, F., De Bleser, R., & Vasishth, S. (2015). Sentence comprehension and morphological cues in aphasia: What eye-tracking reveals about integration and prediction. *Journal of Neurolinguistics, 34,* 83–111.

Hanne, S., 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.

Jacobs, B. J., & Thompson, C. K. (2000). Cross-modal generalization effects of training noncanonical sentence comprehension and production in agrammatic aphasia. *Journal of Speech, Language, and Hearing Research, 43,* 5–20.

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,* 37–55.

Kertesz, A. (2006). *Western Aphasia Battery–Revised (WAB-R)* [Measurement instrument]. San Antonio, TX: Pearson.

Kiran, S. (2012). What is the nature of poststroke language recovery and reorganization? *ISRN Neurology, 2012,* 786872.

Kiran, S., Caplan, D., Sandberg, C., Levy, J., Berardino, A., Ascenso, E., . . . Tripodis, Y. (2012). Development of a theoretically based treatment for sentence comprehension deficits in individuals with aphasia. *American Journal of Speech-Language Pathology, 21,* S88–S102.

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,* 95–127.

Kuperberg, G. R., & Jaeger, T. F. (2016). What do we mean by prediction in language comprehension? *Language, Cognition, and Neuroscience, 31,* 32–59.

Lee, J., & Thompson, C. K. (2011). Real-time production of arguments and adjuncts in normal and agrammatic speakers. *Language and Cognitive Processes, 26,* 985–1021.

Love, T., Swinney, D., Walenski, M., & Zurif, E. (2008). How left inferior frontal cortex participates in syntactic processing: Evidence from aphasia. *Brain and Language, 107,* 203–219.

Maass, A., & Russo, A. (2003). Directional bias in the mental representation of spatial events: Nature or culture? *Psychological Science, 14,* 296–301.

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,* 619–636.

Mack, J. E., Wei, A. Z., Gutierrez, S., & Thompson, C. K. (2016). Tracking sentence comprehension: Test–retest reliability in people with aphasia and unimpaired adults. *Journal of Neurolinguistics, 40,* 98–111.

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

Mitchum, C. M., Haendiges, A. N., & Berndt, R. S. (2004). Response strategies in aphasic sentence comprehension: An analysis of two cases. *Aphasiology, 18,* 675–692.

Patil, U., Hanne, S., Burchert, F., De Bleser, R., & Vasishth, S. (2016). A computational evaluation of sentence processing deficits in aphasia. *Cognitive Science, 40,* 5–50.

Piñango, M. M. (2000). On the proper generalization for Broca's aphasia comprehension pattern: Why argument movement

may not be at the source of the Broca's deficit. *Behavioral and Brain Sciences, 21,* 48–49.

R Core Team. (2015). *R: A language and environment for statistical computing*. Retrieved from http://www.R-project.org/

Saur, D., & Hartwigsen, G. (2012). Neurobiology of language recovery after stroke: Lessons from neuroimaging studies. *Archives of Physical Medicine and Rehabilitation, 93*(1 Suppl), S15–S25.

Scheepers, C., & Crocker, M. W. (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, ERPs, and beyond* (pp. 167–185). New York, NY: Psychology Press.

Schröder, A., Burchert, F., & Stadie, N. (2015). Training-induced improvement of noncanonical sentence production does not generalize to comprehension: Evidence for modality-specific processes. *Cognitive Neuropsychology, 32,* 195–220.

Schwartz, M. F., Linebarger, M. C., Saffran, E. M., & Pate, D. S. (1987). Syntactic transparency and sentence interpretation in aphasia. *Language and Cognitive Processes, 2,* 85–113.

Schwartz, M. F., Saffran, E. M., Fink, R. B., Myers, J. L., & Martin, N. (1994). Mapping therapy: A treatment programme for agrammatism. *Aphasiology, 8,* 19–54.

Schwartz, M. F., Saffran, E. M., & Marin, O. S. M. (1980). The word order problem in agrammatism I: Comprehension. *Brain and Language, 10,* 249–262.

Thompson, C. K. (2011). *Northwestern Assessment of Verbs and Sentences (NAVS)* [Measurement instrument]. Evanston, IL: Northwestern University.

Thompson, C. K., Cho, S., Hsu, C. J., Wieneke, C., Rademaker, A., Weitner, B. B., . . . Weintraub, S. (2012). Dissociations between fluency and agrammatism in primary progressive aphasia. *Aphasiology, 26,* 20–43.

Thompson, C. K., & Choy, J. J. (2009). Pronominal resolution and gap filling in agrammatic aphasia: Evidence from eye movements. *Journal of Psycholinguistic Research, 38,* 255–283.

Thompson, C. K., & den Ouden, D. B. (2008). Neuroimaging and recovery of language in aphasia. *Current Neurology and Neuroscience Reports, 8,* 475–483.

Thompson, C. K., den Ouden, D. B., Bonakdarpour, B., Garibaldi, K., & Parrish, T. B. (2010). Neural plasticity and treatment-induced recovery of sentence processing in agrammatism. *Neuropsychologia, 48,* 3211–3227.

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.

Thompson, C. K., & Shapiro, L. (2005). Treating agrammatic aphasia within a linguistic framework: Treatment of Underlying Forms. *Aphasiology, 19,* 1021–1036.

Thompson, C. K., Shapiro, L. P., Kiran, S., & Sobecks, J. (2003). The role of syntactic complexity in treatment of sentence deficits in agrammatic aphasia: The Complexity Account of Treatment Efficacy (CATE). *Journal of Speech, Language, and Hearing Research, 46,* 591–607.

Thompson, C. K., & Weintraub, S. (2014). *Northwestern Naming Battery (NNB)* [Measurement instrument]. Evanston, IL: Northwestern University.

Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension: The integration of habits and rules*. Cambridge, MA: MIT Press.

Turkeltaub, P. E., Messing, S., Norise, C., & Hamilton, R. H. (2011). Are networks for residual language function and

recovery consistent across aphasic patients? *Neurology, 76,* 1726–1734.

Van Petten, C., & Luka, B. J. (2012). Prediction during language comprehension: Benefits, costs, and ERP components. *International Journal of Psychophysiology, 83,* 176–190.

Wierenga, C. E., Maher, L. M., Moore, A. B., White, K. D., McGregor, K., Soltysik, D. A., . . . Crosson, B. (2006). Neural substrates of syntactic mapping treatment: An fMRI study of two cases. *Journal of the International Neuropsychological Society, 12,* 132–146.