ASHA American Speech-Language-Hearing Association

AJSLP

Research Article

Interplay of Semantic Plausibility and Word Order Canonicity in Sentence Processing of People With Aphasia Using a Verb-Final Language

Jee Eun Sung,[a] Gayle DeDe,[b] and Jimin Park[a]

[a] Department of Communication Disorders, Ewha Womans University, Seoul, South Korea [b] Department of Communication Sciences and Disorders, Temple University, Philadelphia, PA

ARTICLE INFO

Article History:
Received September 15, 2023
Revision received February 7, 2024
Accepted May 11, 2024

Editor-in-Chief: Michael de Riesthal
Editor: Stephen M. Wilson

https://doi.org/10.1044/2024_AJSLP-23-00353

ABSTRACT

Purpose: The Western Aphasia Battery is widely used to assess people with aphasia (PWA). Sequential Commands (SC) is one of the most challenging subtests for PWA. However, test items confound linguistic factors that make sentences difficult for PWA. The current study systematically manipulated semantic plausibility and word order in sentences like those in SC to examine how these factors affect comprehension deficits in aphasia.
Method: Fifty Korean speakers (25 PWA and 25 controls) completed a sentence–picture matching task that manipulated word order (canonical vs. noncanonical) and semantic plausibility (plausible vs. less plausible). Analyses focused on accuracy and aimed to identify sentence types that best discriminate the groups. Additionally, we explored which sentence type serves as the best predictor of aphasia severity.
Results: PWA demonstrated greater difficulties in processing less plausible sentences than plausible ones compared to the controls. Across the groups, noncanonical and less plausible sentences elicited lower accuracy than canonical and plausible sentences. Notably, the accuracy of PWA and control groups differed in noncanonical and less plausible sentences. Additionally, aphasia severity significantly correlated with less plausible sentences.
Conclusion: Even in languages with flexible word order, PWA find it challenging to process sentences with noncanonical syntactic structures and less plausible semantic roles.

The Western Aphasia Battery (WAB)–Revised is arguably the most commonly administered assessment for people with aphasia (PWA; Kertesz, 2022; Kiran et al., 2018; Wallace et al., 2019). The WAB has been translated into many languages, including Spanish (González, 2008), Kannada (Chengappa & Kumar, 2008), Korean (Kim & Na, 2001, 2012), Portuguese (Neves et al., 2014), Tagalog (Ozaeta & Kong, 2012), and Persian (Nilipour et al., 2014). The WAB is a comprehensive test whose goal is describing aphasia severity (via the Aphasia Quotient [AQ]) and diagnosing aphasia syndrome based on performance on measures of fluency, auditory comprehension, naming, and repetition. The present study focuses on how the WAB measures auditory comprehension in PWA.

The WAB measures auditory comprehension at the word (object and picture identification) and sentence levels (yes/no questions and sequential commands). Many clinicians agree that the subtest of sequential commands is the most sensitive of the auditory comprehension measures (Lazar et al., 2008; Lwi et al., 2021). This clinical observation is supported by inspection of WAB-R scores on AphasiaBank. SC is associated with the lowest average performance of the auditory comprehension subtests on the WAB-R (SC = 66.7%, word comprehension = 88.9%, and yes/no questions = 92.3%). SC also shows a greater degree of variation compared to the other comprehension

Correspondence to Jee Eun Sung: jeesung@ewha.ac.kr. **Publisher Note:** This article is part of the Special Issue: Select Papers From the 52nd Clinical Aphasiology Conference. ***Disclosure:*** *The authors have declared that no competing financial or nonfinancial interests existed at the time of publication.*

subtests (*SD* on SC = 22.1 whereas word comprehension = 9.3 and yes/no questions = 5.6; MacWhinney et al., 2011).

The present study is concerned with the sequential commands that take the form of "Point with the pen to the book" and "With the book, point to the pen." Clinically, these items were of particular interest, because they are relatively short, active voice sentences. Critically, the second item both contains noncanonical word order and is also relatively implausible, in the sense that book is a less likely instrument for pointing than pen. Thus, an item such as "With the book, point to the pen" may be difficult to interpret for both syntactic and semantic reasons.

Beginning with Caramazza and Zurif's (1976) seminal paper, it has been clear that both canonicity of word order and semantic plausibility influence sentence comprehension in PWA (Caplan et al., 2007; Thompson & Choy, 2009). A critical finding is that PWA typically have difficulty interpreting semantically reversible sentences with noncanonical word order (e.g., "It was the cat who the dog chased"). The present study examines the relative contribution of word order and semantic plausibility in understanding spoken instructions such as "With the book, point to the pen."

Although previous studies suggested that PWA struggle with sentences that impose greater processing demands (Caplan et al., 2007; Gibson et al., 2016; Grodzinsky, 2000; Sheppard et al., 2015; Thompson & Choy, 2009), there is also evidence that PWA retain the ability to use information such as linguistic regularities, semantic plausibility, and visual context to facilitate processing of spoken and written information (e.g., Gahl, 2002). In fact, even when sentences are syntactically simple, both PWA and healthy controls often demonstrate greater difficulties processing semantically less plausible sentences as compared to plausible ones (DeDe, 2013; Ferreira, 2003; Gahl, 2002). Rational inference approaches postulate that listeners rely heavily on their experience-based expectations of a sentence meaning and are sensitive to the possibility that both speakers and listeners sometimes make mistakes (Warren et al., 2017). For example, if a speaker says, "The girl was kicked by the ball," listeners may infer that the intended meaning was "The girl kicked the ball" and that either they misperceived the sentence or that the speaker made a speech error. PWA may show overreliance on semantic plausibility, because they are aware that they may have misunderstood the sentence, essentially leading them to rely on a rational inference of the speaker's most likely intended message (Gibson et al., 2016; Warren et al., 2017; also cf. Hahn et al., 2022).

Taken together, the existing work suggests that difficulty following commands such as "With the book, point to the pen" may reflect difficulty with semantic or syntactic processing, because these two factors are confounded. That is, a sentence with noncanonical word order has an implausible—but not impossible—meaning. Thus, it is difficult to determine the relative contributions of syntax and semantics to errors on these items on the WAB.

We examined the relative contributions of semantic and syntactic complexity to comprehension of spoken commands in Korean speakers with and without aphasia. Korean is an interesting test case as it allows for word order flexibility with case marking systems, enabling the scrambling of linguistic constituents in a sentence. Due to the morphosyntactic features of the case marking systems in Korean, word order can become flexible. The canonical word order of Korean is subject–object–verb (SOV), but object–subject–verb (OSV) is also allowed. Additionally, Korean is a pro-drop language that frequently omits the noun phrase (NP) in the subject position as long as the meaning of the NP can be inferred from the context. Due to this morphosyntactic flexibility, the effects of word order have not been studied as extensively as they have been in languages like English, which heavily rely on word order. Although languages with flexible word order, such as Korean, offer greater freedom to reorganize syntactic structures, one of these orders is typically preferred and regarded as the basic or canonical order. There is increasing evidence that word order affects sentence processing in Korean aging populations (Sung et al., 2017) as well as people with neurogenic communication disorders such as aphasia (Sung et al., 2018) and mild cognitive impairment (Sung et al., 2020). This phenomenon is also well documented in studies on PWA for Turkish (Duman et al., 2011), Hebrew (Friedmann & Shapiro, 2003), and Tagalog (Bondoc et al., 2018), as well as in young adults for Finnish (Kaiser & Trueswell, 2004) and older adults for Japanese (Arii et al., 2022). Thus, it becomes important to investigate whether word order contributes to syntactic complexity in languages with a relatively flexible word order.

Previous studies examining word order effects in Korean populations focused on SOV and OSV comparisons either in active or passive sentences (Sung, 2015; Sung et al., 2017, 2020). Note that, in Korean, word order (SVO vs. OSV) can be manipulated separately from voice (active vs. passive). These authors consistently found that participants demonstrated lower accuracy with noncanonical word orders, and their performance deteriorated more significantly in syntactically more complex sentence types, such as passive sentences, compared to their simpler active counterparts. However, this evidence is based on word order structures that focus on the thematic roles of agents and themes, which are typically placed in either the subject or object position of a sentence. In Korean, the

subject is marked by a nominative case marker, such as "–i" following a consonant or "–ka" following a vowel, while the object is marked by an accusative case marker "–ul" after consonants or "–lul" before vowels. For example, "The cat chases the dog" in Korean could be structured as "Cat-ka (nom.) Dog-*lul* (acc.) chase-dec," following the SOV word order. Additionally, noncanonical OSV structures such as "Dog-*lul* (acc.) Cat-ka (nom.) chase-dec" are also grammatically permissible. The nominative and accusative case markers are classified as syntactic case markers because they usually attach to the syntactic core argument structures in a sentence. Sung (2015) and Sung et al. (2017, 2018, 2020) have consistently reported canonicity effects of syntactic case markers in aging and clinical populations, showing that the nominative-first word order was easier to process than the accusative-first order. This finding is consistent with linguistic theories that posit that NPs with accusative case markers are regarded as internal arguments of verbs, forming a verb phrase (VP; Chomsky, 1986, 1995; Jun & Kim, 2007; Kempson et al., 2011; Sohn, 2001). When the internal argument is placed outside the VP, the syntactic order becomes noncanonical with an accusative-NP-first word order followed by the nominative NP, resulting in a word order pattern that requires greater cognitive resources to be processed.

In contrast, case markers denoting other thematic roles, such as instruments, goals, or locations, are categorized as semantic case markers (Lee, 1999; Sohn, 2001). Semantic case markers have received relatively little attention in Korean PWA, especially in studies on sentence processing. Thus, it is unclear whether effects of word order canonicity are also observed when the nouns are marked by a syntactic (accusative) and a semantic (instrument) case marker. The distinction between semantic and syntactic case markers somewhat resembles the English difference between verb arguments and adjuncts. This is because, in Korean, syntactic case markers are attached to the verb arguments while NPs with semantic case serve as adjuncts.

In the Korean version of the WAB, the section on serial commands contains items similar to "With the book, point to the pencil" and "Point to the pencil with the book." Because Korean is a verb-final language, the order is either instrumental-NP first (e.g., the book-instrumental case marker + the pencil-accusative case marker + verb), or accusative-NP first (e.g., the pencil-accusative case marker + the book-instrumental case marker + verb). Given that the accusative-NP case marker is viewed as an internal argument of the verb, we hypothesized that a word order with instrumental-NP + accusative-NP + verb might be considered canonical compared to the accusative-NP first structure, in which the internal argument (accusative-NP) is being placed outside the VP. However, this hypothesis has not previously been tested for PWA using the instrumental-NP.

Clinical observations seem to indicate that PWA struggle with both word order and semantic plausibility, especially when combined as in the Korean and English versions of the WAB. However, this has not been systematically examined by varying semantic plausibility and word order in these structures. The current study examined comprehension of imperatives like those on the WAB, such as "Point to the book with the pencil." Both plausibility of the instrument and word order were systematically manipulated to examine how these two factors affect sentence comprehension deficits in aphasia. We formulated the following hypotheses: If the word order effects of the NPs do not manifest themselves in imperative sentences of a verb-final language, only semantic plausibility effects affect comprehension performance. However, if word order effects do exist, we predict that performance will be differentially degraded, especially when the two linguistic features of semantic plausibility and word order complexity are combined. These systematic manipulations including additional items extended from the standardized test will aid in the clinical identification and in a deeper understanding of the relative impairments in syntactic and semantic components of sentence processing in PWA.

Method

Participants

A total of 50 Korean-speaking individuals participated in the study, with 25 PWA and 25 age- and education-matched controls. PWA had a single left hemisphere stroke and were assessed using the Korean version of the WAB (PK-WAB-R; Kim & Na, 2012). The control group took the Mini-Mental State Examination (Kang, 2006) and the Seoul Verbal Learning Test from the Seoul Neuropsychological Screening Battery–II (Kang et al., 2012), and they showed age- and education-adjusted normal range on both tests (≥ 16%ile). Additionally, they did not report any neurological or psychiatric diseases and any vision, hearing, or color perception problems on the health screening questionnaire (Christensen et al., 1991). This research was approved by the institutional review board on human subjects of Ewha Womans University (No. 2022–0140), and demographic information of participants are presented in Table 1.

Stimuli

The stimuli comprised 32 sentences that varied with respect to word order and plausibility. Examples of each sentence type with labeled conditions are provided in Table 2. We manipulated the syntactic condition based on

Table 1. Demographic information and descriptive data from neuropsychological tests.

Variables	Aphasia	Control	*t*	*p* value
Age (*SD*)	54.84 (12.56) (Range: 35–87)	55.20 (12.69) (Range: 34–85)	–.10081	.9201
Education (*SD*)	12.56 (3.08) (Range: 6–18)	13.00 (2.68) (Range: 6–16)	–.53878	.5926
K-MMSE (*SD*)	22.88 (4.98) (Range: 13–30)	28.96 (2.68) (Range: 27–30)	–5.9734	< .000***
PK-WAB-R AQ (*SD*)	68.08 (17.18) (Range: 31–92.1)	—	—	—

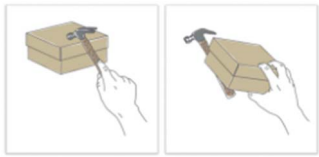
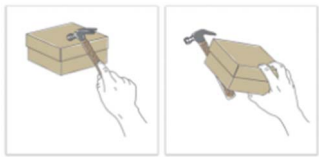
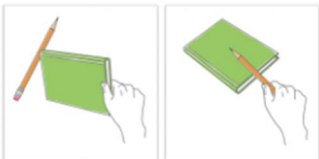
Note. *SD* = standard deviation; K-MMSE = Korean-Mini Mental State Examination (Kang, 2006); PK-WAB-R = Paradise Korea-Western Aphasia Battery–Revised (Kim & Na, 2012); AQ = Aphasia Quotient.

***$p < .001$.

the canonicity of word order (canonical vs. noncanonical) and the semantic condition based on plausibility (plausible vs. less plausible). There were eight sentences in each condition: canonical–plausible (C-P), canonical–less plausible (C-LP), noncanonical–plausible (NC-P), and noncanonical–less plausible (NC-LP).

Each sentence is structured with three components: an instrumental-NP, a theme-NP, and a verb. The sequence of the instrumental-NP determines whether a sentence is considered canonical (when it comes first) or noncanonical (when it comes second). In terms of plausibility, it relies on the semantic connection between the verb and either the

Table 2. Examples of sentence stimuli and pictures for each condition.

Syntactic conditions	Semantic conditions	Sentence stimuli	Picture example
Noncanonical	Less plausible	Point to the hammer with the box. Mangchi-lul sangca-lo twutuli-ta. Hammer-$_{Acc}$ box-$_{Ins}$ Tap-$_{verb}$	
	Plausible	Point to the box with the hammer. Sangca-lul mangchi-lo twutuli-ta. Box-$_{Acc}$ hammer-$_{Ins}$ Tap-$_{verb}$	
Canonical	Less plausible	With the book, point to the pencil. Chayk-ulo yenphil-ul kalikhi-ta. Book-$_{Ins}$ pencil-$_{Acc}$ Point-$_{verb}$	
	Plausible	With the pencil, point to the book. Yenphil-lo chayk-ul kalikhi-ta. Pencil-$_{Ins}$ book-$_{Acc}$ Point-$_{verb}$	

Note. Acc = accusative case marker; Ins = instrumental case marker.

instrumental-NP or the theme-NP. A sentence is considered plausible when the instrumental-NP is intricately tied to the verb in terms of meaning, as exemplified by instances such as "pencil (instrumental-NP)," "book (theme-NP)," and "point (verb)." In each condition, four different types of verbs (point, hit, push, and tap) appeared twice, and we reversed the NPs within each verb to manipulate syntactic or semantic conditions. The presentation sequence of 32 sentences was pseudorandomized, such that more than two items in the same condition occurred in succession. The experimental sentences have been validated across three languages (English, Korean, and Japanese; Sung et al., 2022, 2024). In the Korean database (Younger adults: N = 50, aged 18–30 years; Older adults: N = 56, aged 60–80 years), participants showed overall lower accuracy (%) on noncanonical and less plausible sentences. More specifically, the disparity in performance between the two groups was more pronounced in less plausible sentences for noncanonical (Older adults: M = 52.18, SD = 49.98; Younger adults: M = 95.84, SD = 19.98) and canonical word order (Older adults: M = 69.98, SD = 45.87; Younger adults: M = 98.82, SD = 10.82) compared to their performance on more plausible sentences for noncanonical (Older adults: M = 63.61, SD = 48.14; Younger adults: M = 96.65, SD = 17.99) and canonical word order (Older adults: M = 82.79, SD = 37.77; Younger adults: M = 98.93, SD = 10.27).

Procedure

Screening tests (i.e., PK-WAB-R for PWA and cognitive screens for controls) were conducted prior to the experimental sentence–picture matching task. Participants who met the inclusion criteria completed the main task individually in a quiet experimental room. Target sentences were recorded by a native Korean speaker of Seoul-regional dialect in a sound attenuated booth, and they were presented auditorily through a speaker or headphones to participants. At the same time, two pictures were presented on the screen that depicted the correct and reversed thematic role assignment (see Table 2). Participants were instructed to point to the matching picture as quickly as possible. Participants had three practice trials and began the main trials when the examiner confirmed that they understood the instructions.

Data Analysis

We performed a generalized linear mixed-effects model with a logit link function to analyze response accuracy, which is a categorical variable. This analysis was conducted using the "glmer" function available in the "lme4" and "lmerTest" packages within the R statistical software (R Core Team, 2022). The fixed effects were group, word order, and plausibility, with all sum-coded (−1 Controls, +1 Aphasia; −1 Canonical, +1 Noncanonical; −1 Plausible, +1 Less plausible). For the random intercepts, participants and items were included. Intercepts-only models were computed in the final analysis, since including random slopes did not improve the model fit. Furthermore, we performed a stepwise discriminant analysis using IBM's SPSS statistics 28.0, using the four sentence types (C-P, C-LP, NC-P, and NC-LP) as predictors and "group" as the dependent factor. Finally, a Spearman's correlation and a stepwise multiple regression was executed to determine which sentence type best predicts the AQ.

Results

Effects of Group, Canonicity, and Plausibility on Accuracy

Descriptive statistics of means and standard deviation was provided in Table 3 for both groups (see Figure 1). A generalized linear mixed model was fitted for group, canonicity, and plausibility and the results indicated that all three factors were significant (see Table 4). Specifically, PWA showed significantly lower accuracy than controls (β = −1.3755, SE = .1666, z = −8.256, p < .0001) and noncanonical sentences had significantly lower accuracy compared to the canonical sentences (β = −0.6606, SE = .1182, z = −5.588, p < .0001). Furthermore, less plausible sentences elicited worse performance compared to the plausible ones (β = −0.2862, SE = .1096, z = −2.612, p = .0090). The interaction effect between group and plausibility was significant, indicating PWA demonstrated greater difficulties in processing less plausible than plausible sentences compared to their controls (β = −0.3871, SE = .1096, z = −3.531, p = .0004; see Figure 2). There were no significant interaction effects observed between group and canonicity (p = .5530), canonicity and plausibility (p = .6885), or among group, canonicity and plausibility (p = .6478).

Stepwise Discriminant Analysis

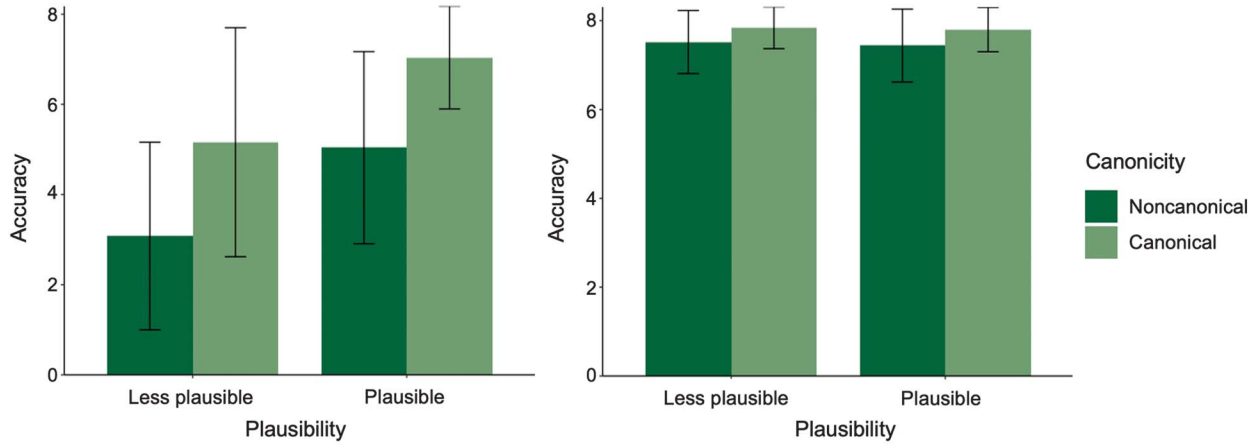
Four sentence types (C-P, C-LP, NC-P, and NC-LP) were included as predictors for stepwise discriminant

Table 3. Accuracy of sentences by canonicity and plausibility between groups.

Canonicity	Plausibility	PWA	Controls
Noncanonical	Less plausible (*SD*)	3.08 (2.08)	7.52 (0.71)
	Plausible (*SD*)	5.04 (2.13)	7.44 (0.82)
Canonical	Less plausible (*SD*)	5.16 (2.54)	7.84 (0.47)
	Plausible (*SD*)	7.04 (1.14)	7.80 (0.50)

Note. PWA = people with aphasia; *SD* = standard deviation.

Figure 1. Accuracy by canonicity and plausibility of each group. PWA = people with aphasia.



analysis with group as the dependent measure. Results revealed that NC-LP was the significant determinant that differentiated PWA from the control group (see Table 5). Controls were 100% successfully classified as a control group, and 88% of PWA were successfully classified as the aphasic group based on NC-LP alone. When all four sentence types were included, 94% of participants were correctly classified (χ^2 = 54.090, Wilks' Lambda = .320, $p < .001$).

Correlational Analyses and Stepwise Multiple Regression Analyses

Spearman's correlation coefficients were computed to explore the association between AQ and each of four sentence types (C-P, C-LP, NC-P, and NC-LP). Before conducting Spearman's correlation analysis, diagnostic tests revealed no concerns regarding multicollinearity (C-P VIF = 1.412; C-LP VIF = 1.309; NC-P VIF = 1.666; NC-LP VIF = 1.309). The results showed that AQ was significantly positively correlated to C-LP (r = .787, $p <$ .001) and NC-LP (r = .648, $p < .001$).

Additionally, stepwise multiple regression analysis was conducted to examine which sentence type best predicted aphasia severity, as indicated by the AQ from the K-WAB. All four sentence types (C-P, C-LP, NC-P, and NC-LP) were included as predictors. The final model revealed that C-LP and NC-LP were significant predictors for AQ, $F(2, 22)$ = 20.293, $p < .001$, R^2 = .648, accounting for 64.8% of the total variance (see Table 6 and Figure 3).

Discussion

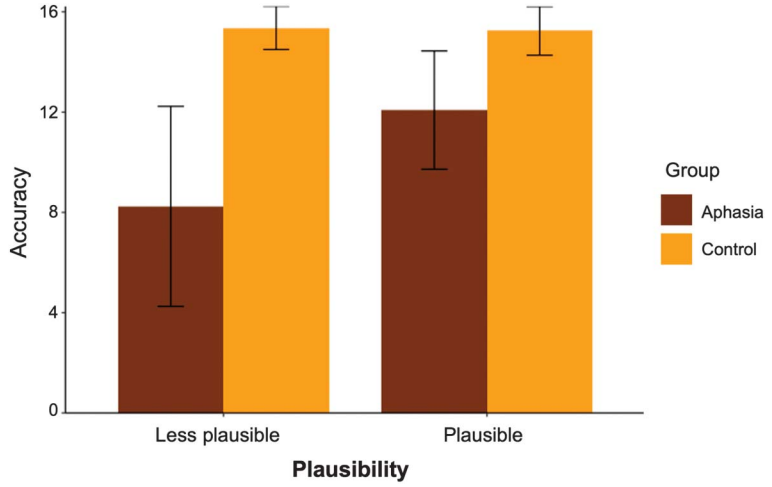
The current study investigated how semantic plausibility and syntactic canonicity contribute to sentence comprehension deficits in aphasia. Not surprisingly, people without aphasia generally performed at ceiling across all conditions. In contrast, PWA demonstrated sensitivity to

Table 4. Summary of the generalized linear mixed-effects models for accuracy.

Predictors	β	*SE*	*z*	*p* value
(Intercept)	2.1477	.1775	12.100	.0000***
Group	–1.3755	.1666	–8.256	.0000***
Canonicity	–0.6606	.1182	–5.588	.0000***
Plausibility	–0.2862	.1096	–2.612	.0090**
Group × Canonicity	–0.0649	.1094	–0.593	.5530
Group × Plausibility	–0.3871	.1096	–3.531	.0004***
Canonicity × Plausibility	0.0455	.1134	0.401	.6885
Group × Canonicity × Plausibility	0.0499	.1093	0.457	.6478

Note. R model equation for accuracy: Accuracy ~ Group × Canonicity × Plausibility + (1 | Subject) + (1 | Item). *SE* = standard error.
**$p < .01$. ***$p < .001$.

Figure 2. Accuracy by group and plausibility.



effects of both semantic plausibility and syntactic canonicity. For both canonical and noncanonical word orders, PWA demonstrated worse performance on semantically less plausible sentences compared to plausible ones, relative to age-matched controls.

The present results demonstrate the effects of semantic plausibility and syntactic canonicity in a new structure (imperatives) and in the case where semantic roles are not impossible but less likely. Recall that the nouns differed with respect to the relative plausibility of being the instrument, but the sentences were reversible in the sense that both nouns could serve as instrument. Thus, the results indicate that PWA are sensitive to this relatively fine-grained, top-down semantic knowledge.

The PWA also responded more accurately to sentences with canonical than noncanonical word order, demonstrating that word order exerts an influence on comprehension even in a language with relatively flexible word order. Although the group by canonicity interaction was not significant, the descriptive statistics suggest that healthy controls demonstrated minimal differences depending on the word order. In contrast to many previous studies on Korean sentence processing, which primarily focused on manipulating two obligatory arguments of the verb (e.g., agent vs. theme), the current study employed the instrumental NP to vary semantic plausibility. In the Korean case-marking system, case markers that denote the agent and theme are referred to as syntactic case markers, while those attached to optional arguments (e.g., instrumental or locative) are categorized as semantic case markers (Sohn, 2001). This is the first study to demonstrate the effects of canonicity in Korean PWA when both a syntactic and a semantic case marker are present.

The canonicity effects are consistent with the idea that the instrumental-NP first condition can be considered syntactically canonical, because the accusative-NP is an internal verb argument. The Isomorphic Mapping Hypothesis (O'Grady & Lee, 2005) provides an alternative explanation of the canonicity effect. This hypothesis posits that syntactic processing difficulties arise when the order of NPs does not align with the corresponding event sequence. For instance, in a canonical Korean sentence from our study (e.g., "The Pencil-Instrumental case marker" + "The Book-accusative case marker" + "point to"), the agent first takes

Table 5. Summary of the stepwise discriminant analysis.

Predictor	Standardized canonical discriminant function coefficients	Structure matrix	Univariated *F* ratio	Rank
C-P	—	.420	9.377	3
C-LP	—	.452	26.813	2
NC-P	—	.350	27.621	4
NC-LP	1.000	1.000	101.897	1
χ^2 = 54.090, Wilks' Lambda = .320, p < .001				

Note. C-P = canonical–plausible; C-LP = canonical–less plausible; NC-P = noncanonical–plausible; NC-LP = noncanonical–less plausible.

Table 6. Model summary and coefficient results of stepwise multiple regression analysis on sentence types.

Model		Sum of squares	*df*	Mean square	*F*	*p*	R^2
1	Regression	3990.520	1	3990.520	29.700	< .001***	.564
	Residual	3090.340	23	134.363			
	Total	7080.860	24	—			
2	Regression	4591.838	2	2295.919	20.293	< .001***	.648
	Residual	2489.022	22	113.137			
	Total	7080.860	24	—			

Model		Unstandardized coefficient		Standardized coefficient β	*t*	*p*	Collinearity diagnosis	
		B	*SE*				Tolerance	VIF
1	(Constant)	41.929	5.329	—	7.868	< .001***	—	—
	C-LP	5.068	0.930	.751	5.450	< .001***	1.000	1.000
2	(Constant)	39.089	5.043	—	7.751	< .001***	—	—
	C-LP	3.975	0.976	.589	4.072	< .001***	.764	1.309
	NC-LP	2.753	1.194	.333	2.305	.031*	.764	1.309

Note. *df* = degree of freedom; VIF = variance inflation factor; C-LP = canonical–less plausible; NC-LP = noncanonical–less plausible.
*$p < .05$. ***$p < .001$.

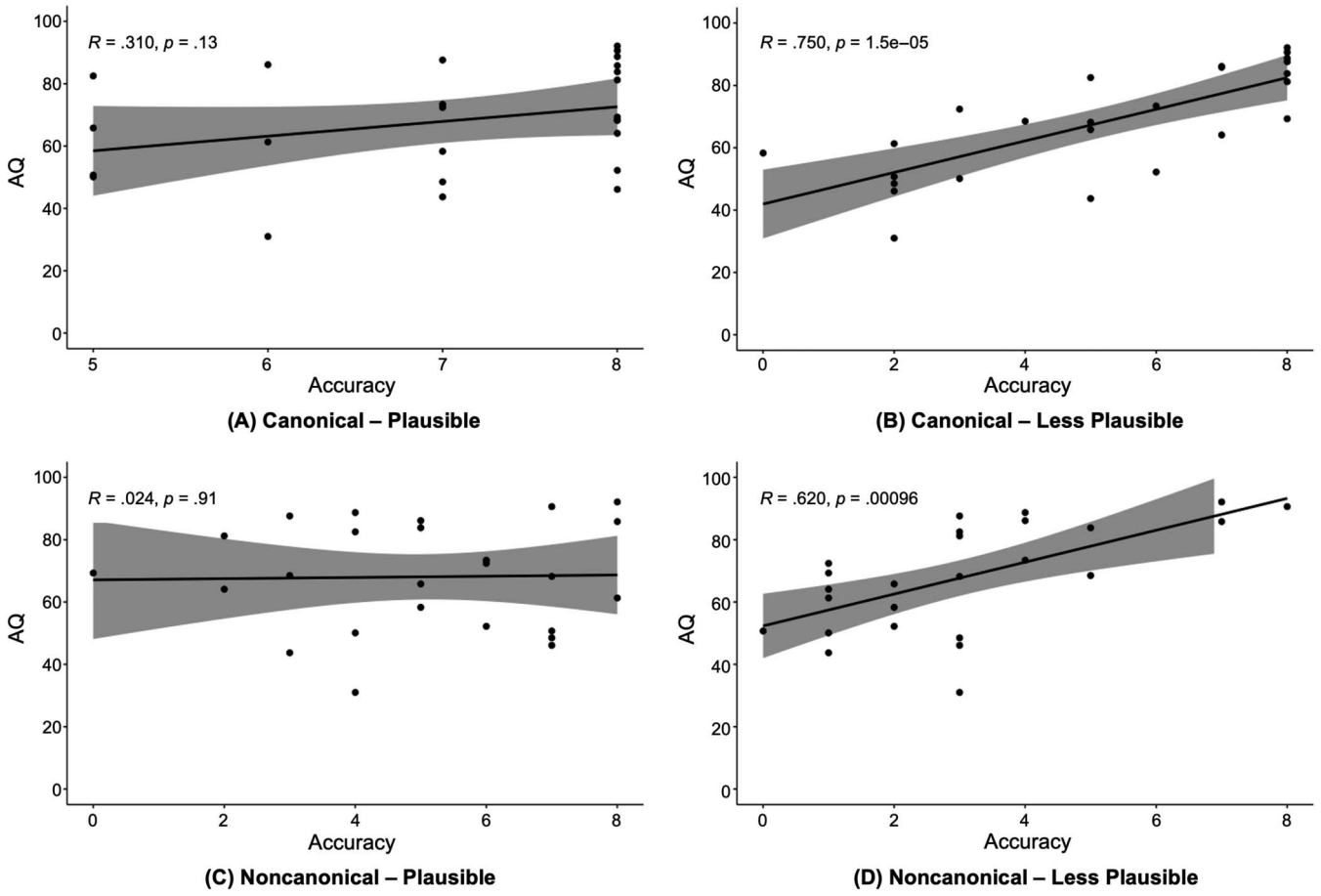
the pencil (1st NP) to point to the book (2nd NP), such that the syntactic structure aligns with the event order. This isomorphic mapping in canonical word order may contribute to better performance in instrument-first sentence types for PWA, as suggested by O'Grady and Lee (2005). Further research is necessary to adjudicate between these interpretations of the canonicity effect.

Mixed-effects models did not reveal a significant interaction between effects of word order and plausibility on comprehension. However, there was evidence that these variables interacted in interesting ways. Most notably, discriminant analysis showed that the predictor that most reliably differentiated PWA from their controls was the condition combining noncanonical word order with less semantic plausibility. These results suggest that noncanonical sentences with semantically less plausible roles (i.e., the NC-LP condition) serve as the most effective items for making a differential diagnosis of aphasia. Stepwise regression showed that difficulty comprehending semantically less plausible sentences—regardless of whether the word order was canonical or noncanonical—was a significant predictor of overall aphasia severity as measured by WAB-AQ. Taken together, the results indicate that the most significant predictor that differentiates individuals with aphasic symptoms from their controls is a combination of two types of linguistic complexity: syntactic canonicity and semantic plausibility. However, within the group of PWA, semantic plausibility is a significant factor correlated with overall severity, irrespective of the canonicity conditions.

It is important to note that in general, the WAB-AQ primarily reflects semantic rather than syntactic processing. The WAB has relatively few items that assess performance based on syntactic manipulations, unlike our test which systematically varied the complexity of syntactic features through canonicity. This might be why the AQ scores correlate more closely with performance on semantic plausibility. The finding that the combined factors of both syntactic canonicity and semantic plausibility identify aphasic symptoms has important clinical implications. Despite the simplicity of our paradigm, it proved to be a significant discriminator with only a small number of items ($n = 8$) for the noncanonical and less plausible condition. These results suggest that this straightforward paradigm could serve as a quick screening tool to determine the presence of aphasia. An interesting next step would be to examine whether people who score above the cut-off for aphasia on the WAB show difficulty on these sentence types.

These results are consistent with the extensive body of literature showing that both semantic plausibility and word order contribute to sentence comprehension in PWA (Caramazza & Zurif, 1976), even in syntactically simple sentences (DeDe, 2013; Gahl, 2002). They are also consistent with the rational inference approach (Gibson et al., 2016). Warren et al. (2017) examined whether PWA's comprehension of active and passive sentences is affected by both the possibility ("The girl kicks the ball" vs. "The ball kicks the girl") and implausibility (e.g., "The cat licked the girl" vs. "The girl licked the cat") of semantic roles. Warren et al. reported an overall effect of coherence, but it is not clear whether the semantically plausible sentences differed significantly from both gradations of semantic plausibility (i.e., possibility or implausibility). Our work builds on Warren et al.'s by demonstrating a

Figure 3. Scatter plots depicting the relationship between four sentence types and the Aphasia Quotient (AQ).



significant difference in accuracy for plausible and less plausible semantic role assignments.

Furthermore, Warren et al. (2017) reported no significant effect of sentence structure (active vs. passive) in their materials, whereas the present study reported effects of word order in active-voice, imperative sentences. The differences in findings may reflect the use of different sentence structures. Alternatively, this divergence may reflect cross-linguistic differences in how word order and semantic plausibility can be manipulated. For instance, in English, word order and plausibility are inherently confounded; altering one inevitably changes the other (i.e., you cannot change the plausibility without also changing the word order). Conversely, in Korean, case markers can be adjusted to change the plausibility of the thematic roles without altering the word order, or vice versa. This flexibility allows for independent variations of word order and semantic plausibility. Intriguingly, an independent effect of word order was observed in Korean, a language known for its relatively free word order. Inspection of the descriptive statistics and results of the discriminant analysis suggest that this was especially true for individuals who struggle with sentence processing due to brain damage. These findings illuminate the complex interplay between semantic and syntactic cues in the processing of verb-final languages for PWA.

We employed a sentence–picture paradigm in the current study. However, it is relatively very simple and easy to modify this paradigm to object manipulation versions as used in the WAB under the section of comprehending the serial commands. The serial commands section in the Korean version of the WAB, mirroring the English-WAB, comprises an equal number of sentences and three items for object manipulations. However, as previously mentioned, the limited number of sentences makes it challenging to identify the locus of processing difficulties within this subsection of the WAB. The current paradigm with more items and systematic control of the linguistic features could provide clinicians with a more thorough examination on the source of sentence comprehension difficulties for PWA. Furthermore, the current paradigm can be applied to various populations with neurogenic communication disorders

such as people with mild cognitive impairment to further investigate whether the two combinations of the linguistic complexity levels serve as a significant factor to differentiate people at risk for dementia.

Another advantage of using this paradigm is that it can easily be translated into various languages given its simplicity of syntactic and semantic structures. Considering these features, the current paradigm may be a very good tool to be used for cross-linguistic comparisons for diverse populations from normal aging to people with neurogenic communication disorders. This is particularly crucial for comparing the processing of linguistic complexities across various languages and could aid in identifying which linguistic structures are more or less susceptible to the effects of neurogenic communication disorders. Overall, the current paradigm holds promise for both expanding our understanding of neurogenic communication disorders and enhancing diagnostic and treatment approaches.

There are also some limitations regarding the design of the task. First, there was a small number of exemplars of each sentence type, and materials were not independently normed. The former concern may have reduced power to detect a significant interaction between sentence structure and plausibility. The latter concern is mitigated by fact that significant effects of plausibility were detected in both the PWA and control participants. Another limitation is that we report only accuracy data. Online measures such as eye tracking might reveal more subtle differences in the time course of processing these sentence types. Future work might also explore cross-linguistic variation in how these types of structures are processed by speakers of subject–verb–object languages such as English and other SOV languages such as Japanese.

To conclude, the present study began with an observation about the measurement of auditory comprehension in a commonly used aphasia test, the Western Aphasia Battery. Specifically, we were interested in sequential commands that were in active voice but varied with respect to canonicity of word order and semantic plausibility. The results indicate that difficulty understanding two-step commands with noncanonical word order and implausible semantic roles is informative with respect to discriminating PWA from controls. Within the group of PWA, the results indicated that commands with implausible semantic roles provide information about aphasia severity, regardless of word order.

Data Availability Statement

The data sets generated and/or analyzed during the current study are available from the corresponding author on reasonable request.

Acknowledgments

This research was partly supported by the National Research Council of Science & Technology Grant CAP21053-000 funded by the Korea government (Ministry of Science and ICT [MSIT]), the National Research Foundation of Korea (NRF) Grant 2022R1A2C2005062 funded by the Korea government (MSIT), and Basic Science Research Program through the NRF Grant NRF2022R1I1A4063209 funded by the Ministry of Education (awarded to Jee Eun Sung).

References

Arii, T., Lee, S. J., Yoshida, T., & Sung, J. E. (2022). Effects of age, word order, and sentence types on Japanese sentence comprehension: A replication study of Sung et al. (2017) and Sung (2015) on Korean. *Educational Studies, 64,* 37–49. https://doi.org/10.34577/00005101

Bondoc, I. P., O'Grady, W., Deen, K., & Tanaka, N. (2018). Agrammatism in Tagalog: Voice and relativisation. *Aphasiology, 32*(5), 598–617. https://doi.org/10.1080/02687038.2017.1366417

Caplan, D., Waters, G., DeDe, G., Michaud, J., & Reddy, A. (2007). A study of syntactic processing in aphasia I: Behavioral (psycholinguistic) aspects. *Brain and Language, 101*(2), 103–150. https://doi.org/10.1016/j.bandl.2006.06.225

Caramazza, A., & Zurif, E. B. (1976). Dissociation of algorithmic and heuristic processes in language comprehension: Evidence from aphasia. *Brain and Language, 3*(4), 572–582. https://doi.org/10.1016/0093-934X(76)90048-1

Chengappa, S. K., & Kumar, R. (2008). Normative & clinical data on the Kannada version of Western Aphasia Battery (WAB-K). *Language in India, 8*(6), 1–5.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. Greenwood Publishing Group.

Chomsky, N. (1995). *The minimalist program*. MIT Press.

Christensen, K. J., Multhaup, K. S., Nordstrom, S., & Voss, K. (1991). A cognitive battery for dementia: Development and measurement characteristics. *Psychological Assessment, 3*(2), 168–174. https://doi.org/10.1037/1040-3590.3.2.168

DeDe, G. (2013). Verb transitivity bias affects on-line sentence reading in people with aphasia. *Aphasiology, 27*(3), 326–343. https://doi.org/10.1080/02687038.2012.725243

Duman, T. Y., Altınok, N., Özgirgin, N., & Bastiaanse, R. (2011). Sentence comprehension in Turkish Broca's aphasia: An integration problem. *Aphasiology, 25*(8), 908–926. https://doi.org/10.1080/02687038.2010.550629

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47*(2), 164–203. https://doi.org/10.1016/S0010-0285(03)00005-7

Friedmann, N., & Shapiro, L. P. (2003). Agrammatic comprehension of simple active sentences with moved constituents: Hebrew OSV and OVS structures. *Journal of Speech, Language, and Hearing Research, 46*(2), 288–297. https://doi.org/10.1044/1092-4388(2003/023)

Gahl, S. (2002). Lexical biases in aphasic sentence comprehension: An experimental and corpus linguistic study. *Aphasiology, 16*(12), 1173–1198. https://doi.org/10.1080/02687030244000428

Gibson, E., Sandberg, C., Fedorenko, E., Bergen, L., & Kiran, S. (2016). A rational inference approach to aphasic language

comprehension. *Aphasiology, 30*(11), 1341–1360. https://doi.org/10.1080/02687038.2015.1111994

González, R. (2008). *Batería de Afasias Western Revisada. Versión traducida y adaptada al Español Chileno* [Western Aphasia Battery Revised translated and adapted to the Chilean Spanish]. PsychCorp.

Grodzinsky, Y. (2000). The trace deletion hypothesis and the tree-pruning hypothesis: Still valid characterizations of Broca's aphasia. *Behavioral and Brain Sciences, 23*(1), 55–64. https://doi.org/10.1017/S0140525X00582399

Hahn, M., Futrell, R., Levy, R., & Gibson, E. (2022). A resource-rational model of human processing of recursive linguistic structure. *Proceedings of the National Academy of Sciences, 119*(43), Article e2122602119. https://doi.org/10.1073/pnas.2122602119

Jun, S. A., & Kim, H. S. (2007). VP focus and narrow focus in Korean. In *Proceedings of the XVIth international congress of phonetic sciences* (pp. 1277–1280).

Kaiser, E., & Trueswell, J. C. (2004). The role of discourse context in the processing of a flexible word-order language. *Cognition, 94*(2), 113–147. https://doi.org/10.1016/j.cognition.2004.01.002

Kang, Y. (2006). A normative study of the Korean-mini mental state examination (K-MMSE) in the elderly. *Korean Journal of Psychology, 25*(2), 1–12.

Kang, Y., Jahng, S., & Na, D. L. (2012). *Seoul Neuropsychological Screening Battery (SNSB-II)*. Human Brain Research & Consulting.

Kempson, R. M., Gregoromichelaki, E., & Howes, C. (2011). *The dynamics of lexical interfaces*. Center for the Study of Language and Information.

Kertesz, A. (2022). The Western Aphasia Battery: A systematic review of research and clinical applications. *Aphasiology, 36*(1), 21–50. https://doi.org/10.1080/02687038.2020.1852002

Kim, H. H., & Na, D. L. (2001). *Paradise Korean version the Western Aphasia Battery (PARADISE·K-WAB)*. Paradise Welfare Foundation.

Kim, H., & Na, D. L. (2012). *Paradise Korean Version of the Western Aphasia Battery Revised (PK-WAB-R)*. Paradise Welfare Foundation.

Kiran, S., Cherney, L. R., Kagan, A., Haley, K. L., Antonucci, S. M., Schwartz, M., Holland, A. L., & Simmons-Mackie, N. (2018). Aphasia assessments: A survey of clinical and research settings. *Aphasiology, 32*(Suppl. 1), 47–49. https://doi.org/10.1080/02687038.2018.1487923

Lazar, R. M., Speizer, A. E., Festa, J. R., Krakauer, J. W., & Marshall, R. S. (2008). Variability in language recovery after first-time stroke. *Journal of Neurology, Neurosurgery, and Psychiatry, 79*(5), 530–534. https://doi.org/10.1136/jnnp.2007.122457

Lee, H. (1999). The domain of grammatical case in lexical-functional grammar. *Proceedings of the LFG99 Conference. Stanford, CA: CSLI Publications.* https://web.stanford.edu/group/cslipublications/cslipublications/LFG/4/lfg99lee.pdf [PDF]

Lwi, S. J., Herron, T. J., Curran, B. C., Ivanova, M. V., Schendel, K., Dronkers, N. F., & Baldo, J. V. (2021). Auditory comprehension deficits in post-stroke aphasia: Neurologic and demographic correlates of outcome and recovery. *Frontiers in Neurology, 12,* Article 680248. https://doi.org/10.3389/fneur.2021.680248

MacWhinney, B., Fromm, D., Forbes, M., & Holland, A. (2011). AphasiaBank: Methods for studying discourse. *Aphasiology, 25,* 1286–1307. https://doi.org/10.1080/02687038.2011.589893

Neves, M. D. B., Borsel, J. V., Pereira, M. M. D. B., & Paradela, E. M. P. (2014). Cross-cultural adaptation of the Western Aphasia Battery - Revised screening test to Brazilian Portuguese: A preliminary study. *CoDAS, 26*(1), 38–45. https://doi.org/10.1590/s2317-17822014000100006

Nilipour, R., Pourshahbaz, A., & Ghoreyshi, Z. S. (2014). Reliability and validity of bedside version of Persian WAB (P-WAB-1). *Basic and Clinical Neuroscience, 5*(4), 253–258.

O'Grady, W., & Lee, M. (2005). A mapping theory of agrammatic comprehension deficits. *Brain and Language, 92*(1), 91–100. https://doi.org/10.1016/j.bandl.2004.05.009

Ozaeta, C., & Kong, A. (2012). Development of the Tagalog version of the Western Aphasia Battery-revised: A preliminary report. *Procedia - Social and Behavioral Sciences, 61,* 174–176. https://doi.org/10.1016/j.sbspro.2012.10.134

R Core Team. (2022). *R: A language and environment for statistical computing* [Computer software]. R Foundation for Statistical Computing. http://www.r-project.org

Sheppard, S. M., Walenski, M., Love, T., & Shapiro, L. P. (2015). The auditory comprehension of *wh*-questions in aphasia: Support for the intervener hypothesis. *Journal of Speech, Language, and Hearing Research, 58*(3), 781–797. https://doi.org/10.1044/2015_JSLHR-L-14-0099

Sohn, H. M. (2001). *The Korean language*. Cambridge University Press.

Sung, J. E. (2015). Effects of syntactic structure on sentence comprehension ability as a function of the canonicity of word-order and its relation to working memory capacity in Korean-speaking elderly adults. *Communication Sciences & Disorders, 20*(1), 24–33. https://doi.org/10.12963/csd.15229

Sung, J. E., Choi, S., Eom, B., Yoo, J. K., & Jeong, J. H. (2020). Syntactic complexity as a linguistic marker to differentiate mild cognitive impairment from normal aging. *Journal of Speech, Language, and Hearing Research, 63*(5), 1416–1429. https://doi.org/10.1044/2020_JSLHR-19-00335

Sung, J. E., DeDe, G., & Lee, S. J. (2022). *Cross-linguistic validations of semantic and syntactic interface for aging-related changes in sentence processing* [Poster presentation]. 60th Annual Meeting of the Academy of Aphasia, Philadelphia, PA, United States.

Sung, J. E., DeDe, G., Lee, S. J., & Park, J. (2024). *Cross-linguistic differences and similarities in age-related sentence processing changes: Effects of word order and semantic plausibility* [Manuscript in preparation]. Department of Communication Disorders, Ewha Womans University.

Sung, J. E., Eom, B., & Lee, S. E. (2018). Effects of working memory demands on sentence production in aphasia. *Journal of Neurolinguistics, 48,* 64–75. https://doi.org/10.1016/j.jneuroling.2018.03.006

Sung, J. E., Yoo, J. K., Lee, S. E., & Eom, B. (2017). Effects of age, working memory, and word order on passive-sentence comprehension: Evidence from a verb-final language. *International Psychogeriatrics, 29*(6), 939–948. https://doi.org/10.1017/S1041610217000047

Thompson, C. K., & Choy, J. J. (2009). Pronominal resolution and gap filling in agrammatic aphasia: Evidence from eye movements. *Journal of Psycholinguistic Research, 38*(3), 255–283. https://doi.org/10.1007/s10936-009-9105-7

Wallace, S. J., Worrall, L., Rose, T., Le Dorze, G., Breitenstein, C., Hilari, K., Babbitt, E., Bose, A., Brady, M., Cherney, L. R., Copland, D., Cruice, M., Enderby, P., Hersh, D., Howe, T., Kelly, H., Kiran, S., Laska, A.-C., Marshall, J., . . . Webster, J. (2019). A core outcome set for aphasia treatment research: The ROMA consensus statement. *International Journal of Stroke, 14*(2), 180–185. https://doi.org/10.1177/1747493018806200

Warren, T., Dickey, M. W., & Liburd, T. L. (2017). A rational inference approach to group and individual-level sentence comprehension performance in aphasia. *Cortex, 92,* 19–31. https://doi.org/10.1016/j.cortex.2017.02.015