Identifying Verbal Short-Term Memory and Working Memory Impairments in Individuals with Latent Aphasia

JoAnn P. Silkes,[a,*] Reva M. Zimmerman,[b] Wendy Greenspan,[c] Laura Reinert, [c] Diane Kendall,[d,e] and Nadine Martin[c]

[a]School of Speech, Language and Hearing Sciences, San Diego State University, CA
[b]Department of Speech and Hearing Sciences, University of Washington, Seattle
[c]Department of Communication Sciences and Disorders, Temple University, Philadelphia, PA
[d]Department of Speech and Hearing Sciences, University of Washington, Seattle
[e]University of Pretoria, South Africa

*Correspondence to JoAnn Silkes: jsilkes@sdsu.edu

Abstract

Purpose: This study was undertaken to explore whether measures of verbal short-term memory and working memory are sensitive to impairments in people with latent aphasia, who score within normal limits on typical aphasia test batteries.

Method: Seven individuals with latent aphasia and 24 neurotypical control participants completed 40 tasks from the Temple Assessment of Language and Short-term Memory in Aphasia (TALSA) that assess various aspects of verbal short-term memory, working memory, and language processing. Subtests were identified that differentiated between the two groups of participants.

Results: Twenty-one TALSA tasks were identified on which the participants with latent aphasia had significantly different performance than the typical control participants. All of these subtests engaged verbal short-term memory, and some involved working memory as well. Furthermore, the TALSA detected individual differences in linguistic profiles among participants with latent aphasia.

Conclusions: People with latent aphasia may be identified by tests that tap verbal short-term memory and working memory. In addition, the TALSA was found to be sensitive to the heterogeneity of this population. Further development of these measures will improve identification and treatment of this challenging population.

Aphasia is typically identified through a combination of formal testing and observation of conversational interactions, allowing a speech-language pathologist to provide a diagnosis, identify specific contributing impairments across language domains, create and modify treatment plans, and justify the continuation or termination of treatment. This approach is adequate for most cases of aphasia, in which linguistic impairments are significant enough that they are identified through these methods. For people with latent aphasia (PWLA; DeDe & Salis, 2020; Pichot, 1955), however, this approach is inadequate. The language skills of these people are very mildly impaired, such that they may score within normal limits on standardized tests. In addition, they may be successful enough at the level of discourse that there are no obvious errors or problems in conversational interactions. PWLA talk about difficulty with attention or focus, along with linguistic agility, but are often competent enough to modify their message so that any problems they may have with word retrieval are not evident to the listener. They may be discharged from treatment, being told that there are no further goals to address, yet be unsuccessful when they attempt to return to prestroke activities, including jobs and other complex activities (Eisenson, 1984). For instance, they report that they cannot keep up with the communication demands in their workplace, that they cannot participate in fast-paced group conversations with friends, or that they cannot discuss complex health matters with their doctors (Marshall, 1993; Raymer & LaPointe, 1986). When communication is successful, PWLA report that it requires all of their effort and they cannot maintain the levels of attention and effort that are needed to sustain their success (Armstrong et al., 2013; DeDe & Salis, 2020; Eisenson, 1984; Raymer & LaPointe, 1986). These unsuccessful outcomes can lead to depression, isolation, and loss of status within families (Lyon, 1992; Sinyor et al., 1986; Währborg & Borenstein, 1989). In addition, without the source of their communication problems accurately and adequately identified, it is difficult to determine the most appropriate intervention approaches.

The discrepancy between observed and experienced language impairment has been recognized as a limitation of standard aphasia batteries. For instance, the manual of the Western Aphasia Battery–Revised (WAB-R; Kertesz, 2006) describes a group of people who exhibit symptoms of aphasia, such as word-finding difficulties or paraphasic errors, but score within the normal range on that test. Similarly, creators of the Comprehensive Aphasia Test (CAT; Swinburn et al., 2004) suggest that other assessment tools are better at detecting mild aphasia. Although this discrepancy has been recognized, it is not well understood. PWLA often report fatigue in complex communication situations, even when they appear outwardly to be having good success. No source has been identified for this fatigue, though it could be related to an ongoing need for linguistic monitoring and modification (McNeil et al., 1991), which is not evident to the listener, as noted above. Some researchers have found that detailed discourse analysis may help to identify PWLA (DeDe & Salis, 2020; Fromm et al., 2017; Jaecks et al., 2012), but this type of analysis is complex and labor intensive to conduct, making it infeasible in standard clinical settings. As a result, there continues to be a need to find accurate, effective, efficient, functional ways to identify people with subtle language impairments following stroke.

At the same time, there are cognitive functions that are not traditionally considered linguistic but are known to support language processing and to be impaired in aphasia, including verbal short-term memory (STM) and working memory (WM; Martin & Ayala, 2004; Ween et al., 1996). Verbal STM is an inherent component of language processing that influences both production and comprehension; it enables successful word access and retrieval via the short-term maintenance of linguistic representations, which are activated by automatic spreading activation through the semantic–lexical–phonological network (Martin, 2000; Martin et al., 1996). Verbal WM, which is hypothesized to be supported by STM capacity (Cowan, 2008), reflects the ability to both activate and manipulate linguistic information (Martin, 2000; McNeil & Pratt, 2001). WM ability involves selectively activating and inhibiting relevant and irrelevant information from within the communication interaction and from the surrounding environment. Martin and colleagues (Martin & Dell, 2019; Martin & Saffran, 1997; Martin et al., 1996) have proposed that aphasia reflects a processing impairment that affects the short-term maintenance of activated linguistic representations as they support word retrieval. If this impairment is severe, it may influence retrieval of single words in general or, if mild, may affect only retrieval of abstract, low-frequency words, or processing of multiple word utterances, resulting in less severe, or no, apparent aphasia (as measured by standard aphasia batteries). Comprehension abilities may also be affected if a person with latent aphasia is having difficulty sustaining activation of incoming linguistic information long enough to parse and process it. For both production and comprehension, the skills of a person with latent aphasia may be adequate for low-demand interactions but inadequate to keep up as the level of complexity, speed, or duration of communicative interactions increases or in the presence of more complex competing environmental and cognitive demands, such as self-monitoring and selecting alternative words when word retrieval fails.

Verbal WM and verbal STM are typically not assessed in people with aphasia. The standardized aphasia batteries currently available, such as the WAB-R and the CAT, provide measures of language function without considering the role of these fundamental underlying mechanisms. In response to this gap, the Temple Assessment of Language and Short-term Memory in Aphasia (TALSA) has been developed to provide measures of both verbal STM and WM (Martin et al., 2018). The TALSA has been developed over the course of 20 years and has grown to include 21 subtests that tap both semantic and phonological processing at a range of stimulus lengths and complexities. Importantly, it is designed to measure the adequacy of support that verbal STM and WM provide to word processing under conditions of (a) delayed production or comprehension and (b) high memory load. Dell and colleagues (Dell et al., 1997; Schwartz et al., 2006) hypothesized two parameters of activation that together support the successful processing of words: activation transmission and activation maintenance. They proposed further that the nature of word processing impairments in aphasia can be characterized as impairment to one or both of these processing parameters, resulting in slow transmission of activation and/or too-fast decay of activation. It is difficult to know precisely how impairments of activation transmission and maintenance may relate to slowed processing as demonstrated empirically (DeDe & Salis, 2020; Fromm et al., 2017) and reported subjectively (Cavanaugh & Haley, 2020) by PWLA. Slow transmission of activation might result in a slower speaking rate, but impairment of activation maintenance will also disrupt verbal output with a resulting reduction in speaking rate. Both activation parameters work together to ensure that words are activated sufficiently and in a timely manner to complete any language task.

The TALSA uses 1-s and 5-s response delays in language measures that are typically used in assessment of aphasia (e.g., word-to-picture matching, phoneme discrimination) to reveal whether one of these two activation parameters is disproportionately impaired. If activation transmission is disproportionately impaired, more time is needed to achieve activation levels that support successful word processing. In this situation, better performance is observed after a response delay. If activation maintenance is disproportionately impaired, linguistic representations of a word are activated fully but decay too quickly and before the word can be retrieved. In this case, performance accuracy declines after a response delay. Martin and Dell (2019) have recently provided behavioral and computational evidence supporting these two patterns of activation impairment.

The TALSA also examines the effect of increasing the WM load of language processing in several ways. For example, similarity judgments (rhyming and synonymy) are varied for the number of items that need to be compared to make the judgments, and sentence repetition is varied by adding modifiers to nouns in active transitive sentences (e.g., The boy watered the plants with a hose ➔ The helpful boy watered the plants with a leaky hose). The TALSA also includes several verbal span tasks that vary language characteristics that are sensitive to semantic or phonological processing of words (e.g., high- and low-imageability word spans, word and nonword spans) and to input and output processing of words (pointing span and repetition span). While full details of the TALSA subtests and norms from the first version of the test battery can be found in Martin et al. (2018), brief descriptions of the different subtests are provided in the Appendix, and the research version of the TALSA (which runs using E-Prime 2.0, Psychology Software Tools, Inc., 2012) will be provided upon request to the last author (N. Martin).

In the research version of the TALSA, there are a total of 56 tasks that can be administered, across 21 subtests. Many subtests comprise several tasks that are different versions of the same thing, manipulating factors as described above, such as the time between stimulus and response (1 s vs. 5 s), WM load, and the distance between target stimuli in span tasks. Although the TALSA is not intended to be administered in its entirety, the research version of the TALSA is too large to be clinically viable; therefore, a study is currently underway to create and standardize a shorter version of the TALSA that will meet clinical needs. This ongoing study includes people with a range of aphasia severity from very mild to very severe, allowing a unique opportunity to assess verbal STM and the effects of STM and WM load on word and sentence processing in PWLA.

Accordingly, this article presents TALSA outcomes from the subset of individuals with aphasia who scored within normal limits on standard aphasia test batteries. The research question addressed in this article is whether there are specific subtests of the TALSA that may be sensitive to identifying latent aphasia that may not otherwise be captured by standard aphasia batteries. Identifying such subtests would provide insight into the status of verbal STM and WM in this challenging population and, importantly, provide a clinically useful tool that opens new avenues for treatment of latent aphasia, improving outcomes for this population.

Method

This study was undertaken with approval of the Temple University and University of Washington Institutional Review Boards. Informed consent was obtained from all participants through a multimodal conversation that ensured participant comprehension.

Participants

Seven PWLA and 24 typical adults provided data for this report (see Table 1). Inclusion criteria for all participants included (a) age between 21 and 80 years; (b) a minimum of a high school education; (c) normal hearing at 25 dB HL at 1000, 2000, and 4000 Hz in at least one ear on an audiometric pure-tone air conduction screening; (d) at least 20/40 vision as measured by the “Tumbling E” char; and (e) English as their first language. The two groups did not differ significantly in age (PWLA: M = 64.19, SD = 9.82; typical controls: M = 57.73, SD = 14.09; $t = -1.36$, p = .19), or level of education (PWLA: M = 15.57, SD = 2.64; typical controls: M = 14.21, SD = 2.57; $t = -1.21$, p = .25). Further criteria for the typical participants included (a) no history of neurological injury or illness and (b) no history of cognitive impairment as determined by the Montreal Cognitive Assessment (Nasreddine et al., 2005).

Table 1. Characteristics of the participants with latent aphasia.

Participant ID	Age (years;months)	MPO	Sex	Education (years)	Etiology	WAB-R AQ	CAT overall severity *t* score
SX3	56;5	188	M	14	L CVA	97.0	—
EL5	49;0	181	F	13	L CVA	94.3	—
FD26	71;8	15	M	18	L CVA	95.0	—
KG47	60;4	159	F	13	L CVA	94.3	—
UW1	78;3	21	M	18	L CVA	—	71.1
UW6	65;1	137	M	19	L CVA	—	73.0
UW7	68;7	73	F	14	L CVA	—	71.3

Note. Em dashes indicate data not available. MPO = months postonset; WAB-R AQ = Western Aphasia Battery–Revised Aphasia Quotient; CAT = Comprehensive Aphasia Test; M = male; L CVA = left cerebrovascular accident; F = female.

The presence of aphasia in PWLA was based on diagnosis by a speech-language pathologist or physician. Additional inclusion criteria for the PWLA included (a) a positive history of left-hemisphere stroke(s), (b) a negative history of a right-hemisphere stroke, and (c) being at least 6 months poststroke. PWLA were also screened for visual neglect and moderate or greater dysarthria or apraxia of speech (Wertz et al., 1984), which may have precluded their completion of verbal tasks, and completed either the WAB-R or CAT, based on standard assessment protocols at the two data collection sites (the WAB-R is used at Temple University, and the CAT is used at the University of Washington). To be identified as having latent aphasia, participants who completed the WAB-R must have scored an Aphasia Quotient of 93.8 or higher, and participants who completed the CAT must have scored an Overall Severity *t* score of 68.2 or higher. Specific information regarding the characteristics of the PWLA is included in Table 1.

Equipment and Setting

The TALSA protocol was computer administered, using E-Prime 2.0 software on a PC (laptop or desktop, depending on the setting). Sessions were conducted in a quiet room either in the Temple University Aphasia Research Lab, in the University of Washington Aphasia Research Lab, or at the participant's home. All tasks that involved verbal responses were audio-recorded to allow reliability assessment of data.

Experimental Protocol

Participants were seen 1–3 times/week for 1–2 hr per session, depending on their tolerance and scheduling availability. Approximately 25–30 hr were required to complete the entire TALSA protocol. The first session involved the informed consent process, hearing and vision screening, and background speech, language, and cognitive testing.

All tasks of the TALSA were computer administered, ensuring reliable administration across experimenters and settings. Test tasks were administered in different sequences for each participant, following specific rules for proximity of related subtests (e.g., at least 2 weeks between tasks within the same subtest). Response modes differed by subtest, including keypresses, pointing responses, and verbal responses. For subtests that involved keypress responses, response data were automatically logged by E-Prime and responses were also recorded by the experimenter on paper response forms. The data logged by E-Prime for tasks that used keypress responses were used for verification of hand-recorded responses if necessary. Verbal responses were scored by hand during the session and were audio-recorded for later verification.

Data Analysis

Of the TALSA's 56 tasks, complete data from both PWLA and controls were available for 40 of the tasks. The scores obtained on each of the 40 tasks were nonnormally distributed so, to minimize the influence of extreme scores in the PWLA sample, outliers in that data set were identified using the interquartile range and removed from the analysis. Then, the median performance on each task for PWLA and typical controls was compared using Wilcoxon rank sum tests. Initially, *p* values were interpreted using the Dunn–Sidak-corrected α level of .0012; however, because this correction is very stringent and was deemed likely to obscure possible differences of interest in this exploratory study, we also interpreted the results using the Benjamini–Hochberg correction for multiple comparisons (Benjamini & Hochberg, 1995). This adjustment integrates the rank of the obtained *p* value, the total number of tests conducted, and the acceptable false discovery rate (Acharya, 2014). For the purposes of this exploratory study, the false discovery rate was set at 0.05.

Results

Using the Dunn–Sidak correction, three tasks were identified as significantly different between groups (see Table 2 for a summary of results from all tasks administered to both groups and Table 3 for scores on each task from each person with latent aphasia). These included two repetition tasks and one span task. Using the Benjamini–Hochberg correction, 22 tasks emerged on which person with latent aphasia showed significantly different median accuracy from the typical control participants. These subtests were distributed across task types, including two naming tasks, eight span tasks, six repetition tasks, and six auditory processing and discrimination tasks

Discussion

This study was undertaken to determine whether assessment of verbal STM and WM skills might be useful to identify aphasia in people with language processing symptoms that are mild enough that they are undetectable on standard aphasia test batteries. A conservative interpretation, using the Dunn–Sidak correction, suggests that three repetition tasks may be sensitive to differences between PWLA and people who do not have aphasia. A less conservative, more exploratory interpretation, using the Benjamini–Hochberg correction, provides insight into other tasks that may be sensitive to PWLA, at least for some people. All of the tasks that were identified using the less conservative approach rely on the ability to retain and manipulate various types of linguistic information.

Using the more conservative interpretation, the three tasks that were identified as significantly different between PWLA and the typical control participants were all repetition tasks. Two of them involved repeating sentences after a 1-s delay, with significant differences found regardless of sentence complexity (i.e., the condition without adjectives added as padding and the condition with adjectives included). The third task that emerged in this analysis was word repetition span, in which participants repeat strings of unrelated words that they hear. Both of these tasks involve activation and maintenance of linguistic information, at least at a phonological level. There are a few reasons to suggest that these tasks are tapping impairments at the phonological level rather than at the semantic level. First, the sentence repetition task provides semantic cues to assist with retrieval, but the word span task does not; as the words within each string are not related to each other, they provide no easy semantic structure to assist in recall. This is consistent with the finding that repetition span is

Table 2. Summary of Temple Assessment of Language and Short-term Memory in Aphasia (TALSA) scores for people with latent aphasia (PWLA) and control participants.

Task type	Subtests shared across versions (max. score)	PWLA median score (range)	Control median score (range)	Uncorrected *p* value	Benjamini–Hochberg corrected *p* value
Span tasks	Category Coordinate Probe Span (7.00)	4.64 (2.89–7.00)	6.87 (4.64–7.00)	.007	.020[a]
	Digit Span – Pointing Response (7.00)	5.20 (3.40–7.00)	7.00 (4.60–7.00)	.002	.010[a]
	Digit Span – Verbal Response (7.00)	5.40 (3.80–7.00)	7.00 (4.60–7.00)	.091	.140
	Word Span – Pointing Response (7.00)	4.20 (2.40–6.20)	5.60 (4.20–7.00)	.003	.011[a]
	Word Span – Verbal Response (7.00)	4.40 (2.20–6.40)	6.00 (3.40–7.00)	.017	.035[a]
	Rhyming Probe Span (7.00)	7.00 (3.60–7.00)	7.00 (3.73–7.00)	.462	.543
	Repetition Span – Words (5.00)	2.40 (1.40–3.20)	3.10 (1.60–3.80)	.001[b]	.040[a]
	Repetition Span – Nonwords (5.00)	4.00 (2.60–5.00)	5.00 (4.20–5.00)	.021	.040[a]
	High–Frequency Words Span (7.00)	4.20 (2.40–4.25)	4.80 (4.10–6.00)	.007	.020[a]
	Low–Frequency Words Span (7.00)	3.60 (2.80–4.36)	4.90 (3.60–5.70)	.017	.035[a]
Naming tasks	TALSA Naming Test – 1 -s delay (1.00)	0.97 (0.92–1.00)	1.00 (0.93–1.00)	.024	.043[a]
	TALSA Naming Test – 5 -s delay (1.00)	0.95 (0.87–0.98)	0.99 (0.93–1.00)	.003	.011[a]
Auditory processing and discrimination tasks	Category Judgment: Picture Stimuli – 1 -s Delay (1.00)	1.00 (0.95–1.00)	1.00 (0.75–1.00)	.407	.522
	Category Judgment: Written Word Stimuli – 1 -s Delay (1.00)	1.00 (0.95–1.00)	1.00 (0.95–1.00)	.791	.850
	Category Judgment: Picture Stimuli – 5 -s Delay (1.00)	1.00 (.85–1.00)	1.00 (0.90–1.00)	.935	.959
	Category Judgment: Written Word Stimuli – 5 -s Delay (1.00)	1.00 (0.75–1.00)	1.00 (0.85–1.00)	.379	.522
	Lexical Comprehension – 1 -s delay (1.00)	1.00 (0.94–1.00)	1.00 (0.98–1.00)	.053	.085
	Lexical Comprehension – 5 -s delay (1.00)	1.00 (0.96–1.00)	1.00 (1.00–1.00)	.001	.010[a]
	Noun/Verb–Concrete/Abstract Synonymy Judgment – 2-word choice (1.00)	1.00 (0.93–1.00)	1.00 (0.88–1.00)	.711	.790
	Noun/Verb–Concrete/Abstract Synonymy Judgment – 3-word choice (1.00)	0.93 (0.88–1.00)	0.98 (0.85–1.00)	.390	.522
	Phoneme Discrimination for Words – 1 -s delay (1.00)	1.00 (0.98–1.00)	1.00 (0.95–1.00)	.808	.850
	Phoneme Discrimination for Nonwords – 1 -s delay (1.00)	1.00 (0.98–1.00)	1.00 (0.95–1.00)	1.000	1.000
	Phoneme Discrimination for Words – 5 -s delay (1.00)	0.98 (0.95–1.00)	1.00 (0.95–1.00)	.016	.035[a]
	Phoneme Discrimination for Nonwords – 5 -s delay (1.00)	1.00 (0.90–1.00)	1.00 (0.90–1.00)	.429	.522
	Rhyming Judgment for Nonwords – 1 -s delay (1.00)	0.98 (0.87–1.00)	0.98 (0.85–1.00)	.647	.739
	Rhyming Judgment for Words – 1 -s delay (1.00)	0.98 (0.47–1.00)	1.00 (0.85–1.00)	.431	.522
	Rhyming Judgment for Nonwords – 5 -s delay (1.00)	0.90 (0.88–1.00)	0.95 (0.85–1.00)	.336	.498
	Rhyming Judgment for Words – 5 -s delay (1.00)	0.98 (0.87–1.00)	1.00 (0.90–1.00)	.415	.522
	Rhyming Triplets – 2-word choice (1.00)	1.00 (0.93–1.00)	1.00 (0.97–1.00)	.012	.030[a]
	Rhyming Triplets – 3-word choice (1.00)	0.90 (0.87–1.00)	1.00 (0.80–1.00)	.003	.011[a]
	Sentence Comprehension – 1 -s delay (1.00)	0.88 (0.74–0.98)	1.00 (0.83–1.00)	.002	.010[a]
	Sentence Comprehension – 5 -s delay (1.00)	0.90 (0.79–0.98)	1.00 (0.90–1.00)	.002	.016[a]

Task type	Subtests shared across versions (max. score)	PWLA median score (range)	Control median score (range)	Uncorrected *p* value	Benjamini–Hochberg corrected *p* value
Repetition tasks	Nonword Repetition – 1 -s delay (1.00)	0.73 (0.33–0.80)	0.87 (0.49–1.00)	.014	.032[a]
	Nonword Repetition – 5 -s delay (1.00)	0.73 (0.07–0.93)	0.87 (0.33–1.00)	.051	.085
	Word Repetition – 1 -s delay (1.00)	0.93 (0.80–1.00)	1.00 (0.93–1.00)	.002	.010[a]
	Word Repetition – 5 -s delay (1.00)	0.98 (0.84–1.00)	1.00 (0.78–1.00)	.030	.052
	Sentence Repetition w/o Adjectives – 1 -s delay (1.00)	0.80 (0.40–1.00)	1.00 (0.80–1.00)	.001[b]	.010[a]
	Sentence Repetition w/ Adjectives – 1 -s delay (1.00)	0.71 (0.00–0.90)	0.97 (0.60–1.00)	.001[b]	.010[a]
	Sentence Repetition w/o Adjectives – 5 -s delay (1.00)	0.86 (0.20–1.00)	1.00 (0.80–1.00)	.002	.010[a]
	Sentence Repetition w/ Adjectives – 5 -s delay (1.00)	0.44 (0.00–0.96)	0.90 (0.30–1.00)	.005	.009[a]

Note. Numbers in parentheses for each task represent the total possible number of points available for that task.

[a]Task was significantly different between groups using Benjamini–Hochberg correction with a false discovery rate of 0.05. [b]Task was significantly different between groups using Dunn–Sidak corrected α level of .0012.

Table 3. Summary of individual Temple Assessment of Language and Short-term Memory in Aphasia (TALSA) scores for people with latent aphasia (PWLA).

Task type	Task	Max. score	Mean control score	Score 2 *SDs* below control performance	Score 2 *SDs* above control performance	Participants						
						TUSX3	TUEL5	TUFD26	TUKG47	UW1	UW6	UW7
Span tasks	Category Coordinate Probe Span	7.00	6.10	4.46	7.74	**3.80**	**2.89**	4.64	7.00	4.64	4.91	4.80
	Digit Span – Pointing Response	7.00	6.69	5.37	8.01	**4.80**	**3.40**	**5.20**	6.80	7.00	6.00	**5.00**
	Digit Span – Verbal Response	7.00	6.69	5.45	7.93	**5.40**	**3.80**	**5.40**	7.00	7.00	7.00	**4.80**
	Word Span – Pointing Response	7.00	5.60	3.98	7.22	4.20	**2.40**	**3.80**	4.60	6.20	4.80	**3.80**
	Word Span – Verbal Response	7.00	5.60	3.24	7.96	4.40	**2.20**	3.80	5.20	6.40	5.00	3.40
	Rhyming Probe Span	7.00	6.70	5.02	8.38	**3.60**	6.99	7.00	7.00	7.00	5.98	7.00
	Repetition Span – Words	5.00	4.84	4.36	5.32	**4.00**	**2.60**	**3.60**	4.40	5.00	**4.20**	**3.00**
	Repetition Span – Nonwords	5.00	2.92	1.56	4.28	2.40	**1.40**	2.40	2.00	3.20	2.40	2.40
	High-Frequency Words Span	7.00	4.87	3.22	6.52	3.45	3.40	4.25	4.20	4.20	4.20	**2.40**
	Low-Frequency Words Span	7.00	4.65	3.13	6.17	4.35	**2.92**	4.36	3.60	4.20	3.20	**2.80**
Naming tasks	TALSA Naming Test – 1-s delay	1.00	0.98	0.92	1.04	1.00	0.97	0.97	0.92	0.93	0.98	0.96
	TALSA Naming Test – 5-s delay	1.00	0.98	0.94	1.02	**0.93**	**0.87**	0.97	**0.92**	0.97	0.98	0.95
Auditory processing and discrimination tasks	Category Judgment: Picture Stimuli – 1-s delay	1.00	0.98	0.86	1.10	1.00	0.95	1.00	0.97	1.00	1.00	1.00
	Category Judgment: Written Word Stimuli – 1-s delay	1.00	0.99	0.95	1.03	1.00	0.95	1.00	0.98	1.00	0.97	1.00
	Category Judgment: Picture Stimuli – 5-s delay	1.00	0.98	0.94	1.02	0.95	**0.75**	1.00	1.00	0.98	1.00	1.00
	Category Judgment: Written Word Stimuli – 5-s delay	1.00	0.98	0.90	1.06	0.95	1.00	**0.85**	1.00	0.98	1.00	1.00
	Lexical Comprehension – 1-s delay	1.00	1.00	1.00	1.00	**0.94**	1.00	1.00	**0.94**	1.00	1.00	1.00
	Lexical Comprehension – 5-s delay	1.00	1.00	1.00	1.00	1.00	1.00	1.00	**0.96**	**0.98**	**0.98**	1.00
	Noun/Verb–Concrete/Abstract Synonymy Judgment – 2-word choice	1.00	0.98	0.90	1.06	1.00	1.00	0.95	0.93	0.98	1.00	1.00
	Noun/Verb–Concrete/Abstract Synonymy Judgment – 3-word choice	1.00	0.94	0.82	1.06	0.90	1.00	0.95	0.88	0.93	0.90	1.00
	Phoneme Discrimination for Words – 1-s delay	1.00	0.99	0.95	1.03	1.00	1.00	1.00	1.00	0.98	0.98	1.00
	Phoneme Discrimination for Nonwords – 1-s delay	1.00	1.00	0.98	1.02	1.00	1.00	1.00	0.98	1.00	1.00	1.00
	Phoneme Discrimination for Words – 5-s delay	1.00	1.00	0.98	1.02	1.00	**0.95**	1.00	1.00	**0.95**	**0.95**	0.98
	Phoneme Discrimination for Nonwords – 5-s delay	1.00	0.98	0.92	1.04	1.00	**0.90**	1.00	1.00	0.93	1.00	0.91
	Rhyming Judgment for Nonwords – 1-s delay	1.00	0.96	0.86	1.06	0.90	1.00	1.00	1.00	**0.47**	0.98	0.98
	Rhyming Judgment for Words – 1-s delay	1.00	0.98	0.90	1.06	0.90	1.00	1.00	0.98	**0.87**	0.98	0.97
	Rhyming Judgment for Nonwords – 5-s delay	1.00	0.95	0.85	1.05	0.95	1.00	1.00	0.98	0.87	0.97	0.98
	Rhyming Judgment for Words – 5-s delay	1.00	0.98	0.90	1.06	0.90	1.00	0.90	0.98	**0.88**	0.98	0.90
	Rhyming Triplets – 2-word choice	1.00	1.00	0.98	1.02	1.00	**0.97**	**0.97**	1.00	**0.93**	1.00	1.00
	Rhyming Triplets – 3-word choice	1.00	0.96	0.82	1.10	0.90	0.90	0.87	0.87	0.93	0.98	0.90
	Sentence Comprehension – 1-s delay	1.00	0.96	0.83	1.09	0.84	**0.74**	0.89	0.88	0.98	0.93	0.85
	Sentence Comprehension – 5-s delay	1.00	0.98	0.90	1.06	**0.89**	**0.79**	0.95	**0.88**	0.95	0.98	0.90

Task type	Task	Max. score	Mean control score	Score 2 *SDs* below control performance	Score 2 *SDs* above control performance	Participants						
						TUSX3	TUEL5	TUFD26	TUKG47	UW1	UW6	UW7
Repetition tasks	Nonword Repetition – 1-s delay	1.00	0.88	0.64	1.12	**0.60**	**0.33**	0.80	0.80	**0.49**	0.80	0.73
	Nonword Repetition – 5-s delay	1.00	0.83	0.47	1.19	0.73	0.07	0.93	0.76	0.49	0.69	0.80
	Word Repetition – 1-s delay	1.00	1.00	0.96	1.04	**0.93**	0.80	**0.93**	1.00	**0.84**	0.98	1.00
	Word Repetition – 5-s delay	1.00	0.99	0.93	1.05	0.93	**0.87**	1.00	1.00	**0.84**	0.98	0.98
	Sentence Repetition w/o Adjectives – 1-s delay	1.00	1.00	0.98	1.02	**0.94**	**0.78**	1.00	1.00	**0.80**	**0.40**	**0.40**
	Sentence Repetition w/ Adjectives – 1-s delay	1.00	0.94	0.84	1.04	0.86	**0.71**	0.90	**0.80**	**0.50**	**0.00**	**0.10**
	Sentence Repetition w/o Adjectives – 5-s delay	1.00	0.99	0.95	1.03	**0.90**	**0.86**	1.00	**0.80**	**0.90**	**0.20**	**0.40**
	Sentence Repetition w/ Adjectives – 5-s delay	1.00	0.89	0.69	1.09	0.80	**0.44**	0.96	**0.60**	**0.40**	**0.10**	**0.00**

Note. Bold numbers indicate scores > 2 *SDs* below the mean control participant scores. Max. = maximum.

associated with phonological abilities but not lexical–semantic abilities of people with aphasia (Martin & Ayala, 2004). Finding that both tasks are impaired suggests a common element at work, which is most readily explained as phonological processing and retention. Second, the sentence repetition task was sensitive to group differences only in the 1-s delay condition but not in the 5-s delay condition. This could indicate that the PWLA needed more time to fully process the incoming stimulus, consistent with an activation transmission deficit as hypothesized by Martin and Dell (2019). Looking at the individual data, however, there is not a consistent distinction between the 1- and 5-s conditions for this task; it is possible that a larger sample size would reveal the 5-s condition to be sensitive for many individuals, as well.

The disconnect between statistically significant findings with the more conservative approach and the patterns of performance seen in the individual data motivated us to also consider a less conservative interpretation of the data to avoid obscuring possible differences of interest in this exploratory investigation. Indeed, using this less conservative approach identified considerably more subtests that show potential for differentiating between PWLA and people who do not have aphasia; these are discussed here.

Two of the subtests that have discriminatory potential involve picture naming, using the TALSA Naming Test. Interestingly, however, the condition in which participants needed to wait 5 s between the stimulus and response was more sensitive to impairment than the 1-s condition. This suggests that these individuals with latent aphasia have difficulty with maintaining activation of target representations over time, even when they may have correctly retrieved the target initially. This finding is consistent with the general idea that people with aphasia may have deficits in accessing and retrieving word representations due to weak spreading of activation and/or in the ability to maintain activation long enough to make use of the activated target (Dell et al., 1997; Kasselimis et al., 2013; Martin & Saffran, 1997; McNeil & Pratt, 2001; Nadeau, 2001).

Eight of the subtests with discriminatory potential are span tasks. Most of these tasks involve participants repeating progressively longer sequences of words or numbers, requiring WM to actively inhibit previous sequences. One span task that has a different structure and was identified as having discriminatory value was the Category Coordinate Probe Span task. In this task, the participant did not repeat the list of words but, instead, heard a list of words and then a probe word was presented at the end. Participants were asked to identify if the probe word belonged to the same category as any word in the list just presented. This task also required WM skills to maintain all words in the list and compare the probe word presented at the end. Thus, the span tasks present a heavy WM load, which is a process not typically tapped by standard aphasia test batteries.

Other tasks that showed high levels of discrimination between groups were those that involved auditory discrimination and comprehension. Despite their commonality in relying on auditory processing skills, however, there are a few subcategories that should be identified. First, the Phoneme Discrimination for Words task distinguished between groups only in the 5-s delay condition, indicating that PWLA are capable of discriminating between phonemes but that their abilities falter when they are required to maintain representations across a delay of several seconds. This is consistent with difficulty in maintaining representations, as discussed earlier regarding performance on the naming task in the 5-s delay condition. Similarly, the lexical comprehension task showed differential performance between delay lengths, with a longer delay being more sensitive to the presence of aphasia.

There was no difference seen between the 1- and 5-s conditions for sentence comprehension. However, the finding that both of these conditions differentiated PWLA from the neurotypical controls suggests that these tasks have a significant maintenance component and WM load to support the ability to retain and interpret lengthy sentences. Finally, performance on both of the rhyming triplets tasks showed statistically significant differences between the two groups when interpreted with the less conservative criterion. In this subtest, the participant hears three words and sees three pictures corresponding to those words and is asked to identify which two of the three words rhyme. In the two-word comparison task, a box is placed around one word to identify it as the target and the participant is to point to which of the remaining two options rhymes with it. In the three-word comparison task, the participant is given no guidance to determine which of the three words are part of the rhyming pair. While both engage phonological activation, maintenance, and manipulation to make the required decision, the three-word version is more difficult, even for typical adults, in that it requires that three possible word pairs be compared, requiring that all three representations remain activated and engaged throughout the decision-making process (Martin et al., 2012). This is reflected in there being a greater numerical difference between groups for the three-word version than the two-word version (although both were statistically significantly different). If a difference between these conditions were to be borne out in a larger sample, this would further support the idea that verbal WM is particularly sensitive to latent aphasia.

Finally, there were six repetition tasks that were sensitive to differences in performance by PWLA and neurotypical adults. These tasks ranged from single words and nonwords to sentences with varying degrees of complexity, suggesting that repetition is a particularly sensitive indicator of latent aphasia. This may be due to the WM requirements for both maintaining and planning output at the same time. Given that there are perception, motor planning, and production requirements for repetition tasks, along with whatever linguistic processing may occur for comprehension, it is likely that it is more difficult to adequately allocate cognitive resources, a function of verbal WM. This finding is consistent with prior research that has shown repetition to be sensitive to impairments when other aspects of aphasia test batteries are not (Raymer & LaPointe, 1986).

Importantly, there was variability in which subtests of the TALSA were sensitive to impairment in different individuals with aphasia. Thus, in addition to performance on tasks with greater memory load revealing impairment in mild aphasia, the variability across tasks tapping into semantic or phonological abilities suggests the potential for identifying patterns of performance across these tasks that provide insight into impairments at specific levels of linguistic representation (i.e., semantic or phonological) and modalities. For instance, in the span tasks, it is possible to discern weaknesses in input versus output processing and at phonological and semantic levels. As an example, TUEL5's category span (a semantic task) is 2.89, but her rhyming span (a phonological task) is 6.99 items. These two tasks are sensitive to input processing as they do not require a verbal response. TUEL5's repetition spans for words and nonwords are relatively low, that is, 2.6 and 1.4, respectively. These tasks require input and output processing. From these four span scores and some of the other tests shown in Table 3, preliminary hypotheses about TUEL5's language profile could be proposed as follows: On the input pathways, input phonological activation is strong and enduring enough to support matching of two similar phonological representations (the rhyme) but is not sufficient to support output phonological encoding in repetition of nonwords. At the lexical–semantic level, input activation may not be strong or enduring enough to support access to and maintenance of lexical–semantic representations, resulting in the lower category probe span score. The decline in accuracy on the category judgments (picture stimuli) from .95 to .75 after a 5-s interval suggests an activation–maintenance problem rather than weak activation. Additionally, the significant decline in naming accuracy after a 5-s delay could reflect poor maintenance of activation spreading from semantics to lexical and/or lexical–phonological representations. The type of errors made in naming would also provide insight into which representations (i.e., phonological or semantic) are affected by the activation maintenance deficit.

As another example of the ability to use specific task performance patterns to understand underlying mechanisms and impairments, TUKG47 shows a different pattern from TUEL5. Rather than a gap between the phonologically based rhyming probe span and the semantically based category probe span, TUKG47 showed equally high levels of performance on both of these tasks. This suggests strong and enduring input phonological and semantic activation processes. On the repetition tasks, however, which engage both input and output processes, word repetition span (4.00) is greater than nonword repetition span (2.00) and repetition of single words (1.00) is greater than repetition of single nonwords (0.76 correct). Repetition of nonwords relies heavily on phonological processing with minimal lexical–semantic support. Additionally, he shows lower performance on the rhyming triplets (a phonological WM task) in the higher memory load condition (a three-word comparison). Therefore, a likely hypothesis about TUKG47's profile is that input phonological and lexical–semantic processing is strong and enduring enough to support rhyme and category recognitions in the probe span tasks but begins to falter if the word representations must be maintained in WM (as in the rhyming triplet judgment task). The lower score on repetition of nonwords indicates difficulty with strength and/or maintenance of output phonological representations. This difficulty may extend to picture naming and sentence repetition, both of which are lower as compared to the other participants, which may be due to a problem with phonological output processing. As in the case of TUEL5, it would be important to observe the types of errors made in naming and sentence repetition to help determine the involvement of output processing of semantic or phonological representations.

TUKG47's profile does not suggest a transmission or activation maintenance deficit in the way that parts of TUEL5's profile did. Instead, evidence for the presence of one processing difficulty or the other comes from changes in performance following a response delay (5-s delay in the TALSA). For TUKG47, performance on the various tasks after a 5-s delay did not change much except for the sentence repetition tasks. If there is a maintenance deficit in TUKG47's case, it may not be apparent until the stimuli that need to be maintained exceed a certain memory load.

The example profiles of TUEL5's and TUKG47's test results could provide some insight into the language impairments that are otherwise not evident in standard aphasia batteries such as the WAB-R. In a clinical setting, this level of diagnostic precision provided by the TALSA would enable clinicians to establish individualized treatment goals that target the specific deficits of PWLA. In the case of TUEL5, for instance, a relevant therapy goal could focus on improving the maintenance of lexical–semantic information for input tasks, whereas TUKG47 might benefit most from treating phonological level processing.

The outcome of this analysis suggests some important considerations in the quest to develop sensitive methods of identifying the presence and nature of latent aphasia. First, there are ways to identify subtle language impairments in the heterogeneous population of people with high-level aphasia who score above diagnostic cutoffs on standard aphasia batteries. This finding acknowledges that these people do, indeed, have impairments despite their apparent communicative success and demonstrates that the significant individual differences that occur can be captured by formal evaluation. Second, regardless of whether a more or less conservative approach is taken to analyze these data, the tasks that identify this population appear to capture impairments in activation transmission and maintenance, and in WM, none of which are specifically addressed or assessed in standard aphasia test batteries.

Limitations and Future Directions

This was a preliminary investigation, with a small sample of PWLA. As such, the suggested interpretations of performance across the various subtests should be regarded as preliminary and supplemented by further data, such as naming error profiles. Nonetheless, the profiles discussed here provide a useful starting point for considering how the TALSA might be used to evaluate the language abilities of someone with latent aphasia. Moreover, these two examples make clear that language profiles across PWLA are variable but can be determined with span tasks and high–memory load language tasks that are sensitive to input and output lexical–semantic and phonological processing. We are currently investigating the interpretation of various patterns of performance on the subtests of the TALSA.

While statistical analysis of the age and education ranges of the two groups (PWLA and typical adults) did not show significant differences in these factors, this may have been an artifact of the small group size for the PWLA paired with the heterogeneity noted for this group. Future investigation should consider the potential impact of age and education on the patterns of performance that are observed, particularly because STM and WM may be affected by age. Additional future directions for research include working to understand the functional, practical effects that latent aphasia has on people's lives and continuing to develop methods for identifying this population and ways to improve treatment and maximize their communication function and success.

Acknowledgments

Research reported in this article was supported by the National Institute on Deafness and Other Communication Disorders under Award Number R01DC016094 (awarded to N. Martin). The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

The authors would like to thank Kevin McCaffery and Sónia Vieira for their contributions to this project; Yeban Lee and Yusheng Wang for their assistance in data management at the University of Washington; Margaret Driscoll, Jade Uffelman, and Gabriella Langan for their assistance in data management at Temple University; and the participants at both sites for their time and efforts.

References

Acharya, A. (2014). *A complete review of controlling the FDR in a multiple comparison problem framework—The Benjamini–Hochberg Algorithm*. arXiv preprint arXiv:1406.7117

Armstrong, E., Fox, S., & Wilkinson, R. (2013). Mild aphasia: Is this the place for an argument? *American Journal of Speech-Language Pathology*, 22(2), S268–S278. https://doi.org/10.1044/1058-0360(2012/12-0084)

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological),* 57(1), 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

Cavanaugh, R., & Haley, K. L. (2020). Subjective communication difficulties in very mild aphasia. *American Journal of Speech-Language Pathology*, 29(1S), 437–448. https://doi.org/10.1044/2019_AJSLP-CAC48-18-0222

Cowan, N. (2008). What are the differences between long-term, short-term, and working memory? *Progress in Brain Research*, 169, 323–338. https://doi.org/10.1016/S0079-6123(07)00020-9

DeDe, G., & Salis, C. (2020). Temporal and episodic analyses of the story of Cinderella in latent aphasia. *American Journal of Speech-Language Pathology*, 29(1S), 449–462. https://doi.org/10.1044/2019_AJSLP-CAC48-18-0210

Dell, G. S., Schwartz, M. F., Martin, N., Saffran, E. M., & Gagnon, D. A. (1997). Lexical access in aphasic and nonaphasic speakers. *Psychological Review*, 104(4), 801–838. https://doi.org/10.1037/0033-295X.104.4.801

Eisenson, J. (1984). The recovered aphasic: Residual problems and vocational implications. In J. Eisenson (Ed.), *Adult aphasia* (2nd ed., pp. 240–251). Prentice-Hall.

E-Prime 2.0. (2012). [Computer software]. *Psychology Software Tools, Inc*.

Fromm, D., Forbes, M., Holland, A., Dalton, S. G., Richardson, J., & MacWhinney, B. (2017). Discourse characteristics in aphasia beyond the Western Aphasia Battery cutoff. *American Journal of Speech-Language Pathology,* 26(3), 762–768. https://doi.org/10.1044/2016_AJSLP-16-0071

Jaecks, P., Hielscher-Fastabend, M., & Stenneken, P. (2012). Diagnosing residual aphasia using spontaneous speech analysis. *Aphasiology*, 26(7), 953–970. https://doi.org/10.1080/02687038.2012.663075

Kasselimis, D. S., Simos, P. G., Economou, A., Peppas, C., Evdokimidis, I., & Potagas, C. (2013). Are memory deficits dependent on the presence of aphasia in left brain damaged patients? *Neuropsychologia*, 51(9), 1773–1776. https://doi.org/10.1016/j.neuropsychologia.2013.06.003

Kertesz, A. (2006). *Western Aphasia Battery—Revised*. Pearson Clinical. https://doi.org/10.1037/t15168-000

Lyon, J. G. (1992). Communication use and participation in life for adults with aphasia in natural settings: The scope of the problem. *American Journal of Speech-Language Pathology*, 1(3), 7–14. https://doi.org/10.1037/t15168-000

Marshall, R. C. (1993). Problem-focused group treatment for clients with mild aphasia. *American Journal of Speech-Language Pathology*, 2(2), 31–37. https://doi.org/10.1044/1058-0360.0202.31

Martin, N. (2000). Word processing and verbal short-term memory: How are they connected and why do we want to know? *Brain and Language*, 71(1), 149–153. https://doi.org/10.1006/brln.1999.2237

Martin, N., & Ayala, J. (2004). Measurements of auditory–verbal STM span in aphasia: Effects of item, task, and lexical impairment. *Brain and Language*, 89(3), 464–483. https://doi.org/10.1016/j.bandl.2003.12.004

Martin, N., & Dell, G. S. (2019). Maintenance versus transmission deficits: The effect of delay on naming performance in aphasia. *Frontiers in Human Neuroscience,* 13, 406. https://doi.org/10.3389/fnhum.2019.00406

Martin, N., Kohen, F., Kalinyak-Fliszar, M., Soveri, A., & Laine, M. (2012). Effects of working memory load on processing of sounds and meanings of words in aphasia. *Aphasiology*, 26(3–4), 462–493. https://doi.org/10.1080/02687038.2011.619516

Martin, N., Minkina, I., Kohen, F. P., & Kalinyak-Fliszar, M. (2018). Assessment of linguistic and verbal short-term memory components of language abilities in aphasia. *Journal of Neurolinguistics*, 48, 199–225. https://doi.org/10.1016/j.jneuroling.2018.02.006

Martin, N., & Saffran, E. M. (1997). Language and auditory–verbal short-term memory impairments: Evidence for common underlying processes. **Cognitive Neuropsychology**, 14(5), 641–682. https://doi.org/10.1080/026432997381402

Martin, N., Saffran, E. M., & Dell, G. S. (1996). Recovery in deep dysphasia: Evidence for a relation between auditory–verbal STM capacity and lexical errors in repetition. *Brain and Language,* 52(1), 83–113. https://doi.org/10.1006/brln.1996.0005

McNeil, M. R., Odell, K., & Tseng, C. H. (1991). Toward the integration of resource allocation into a general theory of aphasia. In T. Prescott (Ed.), *Clinical aphasiology* (Vol. 20, pp. 21–39). Pro-Ed.

McNeil, M. R., & Pratt, S. R. (2001). Defining aphasia: Some theoretical and clinical implications of operating from a formal definition. *Aphasiology*, 15(10–11), 901–911. https://doi.org/10.1080/02687040143000276

Nadeau, S. E. (2001). Phonology: A review and proposals from a connectionist perspective. Brain and Language, 79(3), 511–579. https://doi.org/10.1006/brln.2001.2566

Nasreddine, Z. S., Phillips, N. A., Bedirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal cognitive assessment, MoCA: A brief screening tool for mild cognitive impairment. *Journal of the American Geriatric Society*, 53(4), 695–699. https://doi.org/10.1111/j.1532-5415.2005.53221.x

Pichot, P. (1955). Language disturbances in cerebral disease: Concept of latent aphasia. *Archives of Neurology & Psychiatry*, 74(1), 92–96. https://doi.org/10.1001/archneurpsyc.1955.02330130094011

Raymer, A. M., & LaPointe, L. L. (1986). The nature and assessment of the mildly-impaired aphasic person. *Seminars in Speech and Language*, 7(2), 207–221. https://doi.org/10.1055/s-0028-1085230

Schwartz, M. F., Dell, G. S., Martin, N., Gahl, S., & Sobel, P. (2006). A case-series test of the interactive two-step model of lexical access: Evidence from picture naming. *Journal of Memory and Language*, 54(2), 228–264. https://doi.org/10.1016/j.jml.2005.10.001

Sinyor, D., Amato, P., Kaloupek, D. G., Becker, R., Goldenberg, M., & Coopersmith, H. (1986). Post-stroke depression: Relationships to functional impairment, coping strategies, and rehabilitation outcome. *Stroke*, 17(6), 1102–1107. https://doi.org/10.1161/01.STR.17.6.1102

Swinburn, K., Porter, G., & Howard, D. (2004). *Comprehensive Aphasia Test*. Taylor & Francis Group. https://doi.org/10.1037/t13733-000

Währborg, P., & Borenstein, P. (1989). Family therapy in families with an aphasic member. Aphasiology, 3(1), 93–98. https://doi.org/10.1080/02687038908248978

Ween, J. E., Verfaellie, M., & Alexander, M. P. (1996). Verbal memory function in mild aphasia. *Neurology,* 47(3), 795–801. https://doi.org/10.1212/WNL.47.3.795

Wertz, R. T., LaPointe, L. L., & Rosenbek, J. C. (1984). *Apraxia of speech in adults: The disorder and its management*. Grune &Stratton.