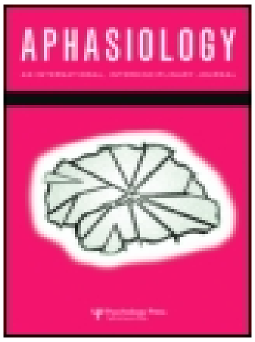

Aphasiology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/paph20

Measuring treatment outcome in severe Wernicke's aphasia

A. Lerman , M. Goral , L. A. Edmonds & L. K. Obler

APHASIOLOGY
https://doi.org/10.1080/02687038.2020.1787729

Measuring treatment outcome in severe Wernicke's aphasia

A. Lerman[a,b], M. Goral[a,c,d], L. A. Edmonds[e] and L. K. Obler[a]

[a]City University of New York, New York, NY, USA; [b]Program of Communication Disorders, Hadassah Academic College, Jerusalem, Israel; [c]Program of Speech-Language-Hearing Sciences, Lehman College, City University of New York, New York, USA; [d]MultiLing Center for Multilingualism in Society across the Lifespan, University of Oslo, Oslo, Norway; [e]Program of Communication Sciences and Disorders, Teachers College, Columbia University, New York, NY, USA

ABSTRACT

Background: Chronic severe Wernicke's aphasia has a poor prognosis and is challenging to treat. Furthermore, even when there is potential for improvement, formal assessments using accuracy scores only to measure changes in language abilities after treatment may not be sensitive enough to capture improvements. Less-constrained language tasks, such as discourse analysis, may be more sensitive to measuring change than more standard constrained tasks, such as confrontation naming and picture-based sentence construction.
Aims: In this study, we asked whether it is possible to rehabilitate language abilities in a participant with severe Wernicke's aphasia using a verb-based sentence-level treatment (Verb Network Strengthening Treatment – VNeST) that has been successful for moderate Wernicke's aphasia, as well as other types of moderate to severe aphasia. Furthermore, we investigated whether using less-constrained language tasks would be more, less or equally sensitive to measuring any treatment effects than more-constrained language tasks.
Methods and procedures: In this case study, we compared post-treatment language abilities to pre-treatment language abilities by analysing comprehension and production at the word, sentence and discourse levels, using both quantitative analyses (e.g., accuracy scores) and qualitative analyses (e.g., error analyses).
Outcomes and results: We found that discourse analysis was sensitive enough to identify improvements in quality of production concomitant with an overall reduction of output. Furthermore, in certain more-constrained tasks, a reduction in the production of neologistic jargon was observed, as well as stable comprehension requiring less repetition of stimuli, indicating improvement that was not captured by accuracy scores.
Conclusions: People with chronic severe Wernicke's aphasia may improve after treatment but formal assessments are not always sensitive enough to identify these improvements. Speech-language therapists are encouraged to include discourse analysis in their assessments as well as the analysis of more formal assessments qualitatively as well as quantitatively.

ARTICLE HISTORY
Received 30 January 2020
Accepted 15 June 2020

KEYWORDS
Wernicke's aphasia; severe aphasia; treatment; discourse; neologism

CONTACT A. Lerman alerman@gradcenter.cuny.edu Program of Speech-Language-Hearing Sciences, New York, NY 10016

Introduction

Wernicke's aphasia

Wernicke's aphasia, also known as receptive aphasia, is a relatively uncommon aphasia profile, occurring in about 16–20% of aphasia cases in the acute stages, and in about 5% of aphasia cases in the chronic stage (Pedersen et al., 2004; Robson et al., 2019). By definition, people with Wernicke's aphasia have impaired comprehension, especially auditory comprehension; if auditory and phonological abilities are significantly impaired up to 5-months post-stroke, poor language comprehension skills are predicted long term (Robson et al., 2019). Additionally, Wernicke's aphasia involves fluent speech with normal intonation, but the output is often difficult to understand because of semantic and/or phonemic paraphasias (e.g., Damasio, 1992). Speech can be empty, with generic substitutes such as "thing" or "stuff", or incomprehensible because it is full of neologisms (e.g., Damasio, 1992; Nicholas et al., 1985; Robson et al., 2019). For those patients with neologistic jargon, whose speech is fluent but often unintelligible due to inclusion of many non-words, it is generally accepted that these errors are a result of difficulties, or failure, in lexical retrieval (e.g., Marshall, 2018, 2006). This could be due either to impairment in the connections between the semantic system and the lexical system, impairment within the lexical system itself, or impairment in the connections between the lexical system and the phonological system (e.g., Marshall, 2006).

Furthermore, people with Wernicke's aphasia have a reduced sense of deficit (e.g., Damasio, 1992; Marshall, 2006) which can lead to reduced motivation to treat any language difficulties. This reduced motivation for treatment, together with the low incidence of Wernicke's aphasia at the chronic stage of recovery, has led to treatment for Wernicke's aphasia being less-studied than other aphasia types, such as Broca's aphasia, anomic aphasia, and conduction aphasia. This is especially salient for severe Wernicke's aphasia; very few cases are reported in the aphasia treatment literature.

Treatment for people with Wernicke's aphasia

Different types of treatment have been shown to be effective in improving language abilities for people with Wernicke's aphasia, and include those aimed at the single-word, sentence or discourse level (e.g., Edmonds et al., 2014, 2015; Boyle, 2004; Knoph et al., 2017; Kurland et al., 2010; Marshall, 2018; Rogalski et al., 2013). These include Semantic Feature Analysis of nouns (nSFA) (Boyle, 2004), Verb Network Strengthening Treatment (VNeST) (Edmonds et al., 2014, 2015), Semantic Feature Analysis of verbs (vSFA) (Knoph et al., 2017), discourse treatment through written home assignments (Ulatowska & Chapman, 1989), communication-based treatment (CBT) (Knoph et al., 2017), and Attentive Reading and Constrained Summarisation (ARCS) treatment (Rogalski et al., 2013).

The literature on treatment for Wernicke's aphasia is sparse and the language outcomes reported appear to vary based on the type of treatment provided. However, the relationship between treatment type (treating single-words, sentences, or discourse) and outcome (improvement to single-words, sentences, or discourse) is not necessarily linear. Rather, treatment type appears to differentially affect constrained tasks (i.e., tasks with pre-determined target responses, such as picture-naming tasks) compared with less-constrained tasks (i.e., tasks for which there are no specific target responses, such as elicited conversation) (e.g., Edmonds et al., 2014, 2015; Boyle, 2004; Knoph et al., 2017; Ulatowska & Chapman, 1989).

Measuring treatment effectiveness in Wernicke's aphasia

A variety of approaches used to measure treatment effectiveness have been identified across studies and include both measures of constrained language tasks and less-constrained language tasks. Constrained language tasks are often used in formal assessments, and for lexical-semantic skills include testing of word and sentence comprehension, confrontation naming of nouns and verbs, and picture-based sentence construction. In the literature on Wernicke's aphasia, improvements have been identified using accuracy scores on confrontation naming tests (e.g., Boyle, 2004; Kurland et al., 2010; Rogalski et al., 2013), although changes may also be observed with error analyses (Marshall, 2018). For example, for people with Wernicke's aphasia who produce neologistic jargon, an increase in empty, anomic speech or real-word errors together with a decrease in the production of non-words indicates an improvement (e.g., Marshall, 2018, 2006).

Less-constrained tasks are usually considered more informal, and include discourse production (e.g., Dietz & Boyle, 2018a, 2018b; Thomson et al., 2018). Their analysis takes into account the way the Speech-language therapist (SLT) interacts with the patient. Also, it considers the type of stimuli used, the way they are presented, and the target response from the patient, all for the purpose of making clinical decisions (Thomson et al., 2018). In Wernicke's aphasia, for less-constrained tasks an increase in informative words together with a decrease in empty words or neologisms would indicate improvement (e.g., Rogalski et al., 2013). Improvement could also be reflected in measures of more specific lexical retrieval of nouns and verbs, such as greater type-token ratios, and/or more complete SV(O) sentences (e.g., Edmonds et al., 2014, 2015; Knoph et al., 2017).

Discourse analysis has been widely studied in the aphasia literature. On the one hand, discourse is considered an important clinical assessment tool for mild aphasia (Coelho & Flewellyn, 2003), and possibly for severe aphasia (Huber, 1990), because of the value in assessing a patient's more naturalistic linguistic performance (e.g., Armstrong, 2000; Thomson et al., 2018). On the other hand, discourse is often assumed to place a communicative load on people with aphasia, because it is considered to be more complex than other language tasks, involving pragmatic, semantic, phonological, syntactic and morphosyntactic abilities together with non-linguistic executive functions such as planning and working memory (e.g., Coelho et al., 1994; Fleming, 2007; Kavé & Goral, 2017).

However, despite an increased communicative load, some researchers suggest that the flexibility involved in producing discourse may provide an advantage to people with aphasia in lexical retrieval measures of discourse relative to constrained tasks, such as confrontation naming and picture-based sentence production (e.g., Edmonds et al., 2014, 2015; Faroqi-Shah, 2012; Ingles et al., 1996) because there are generally numerous ways to express a given idea, story, instruction or opinion (e.g., Armstrong, 2000; Bandur & Shewan, 2008). Due to individual differences among people with aphasia, together with specific language impairments that they demonstrate relative to their type of aphasia, some participants may benefit from this flexibility when producing discourse while others may benefit from tasks involving more constraint (e.g., Edmonds et al., 2014).

Furthermore, there are varying levels of flexibility concomitant with different types of discourse elicitation (Nicholas & Brookshire, 1993; Olness, 2006). Compare describing a picture scene which necessitates some use of predesignated lexical items with a request for a personal narrative such as “tell me about a family vacation” which does not require using any predesignated lexical items.

For people with Wernicke’s aphasia, difficulties in lexical retrieval can be severe compared to relatively spared syntax and morphosyntax (Marshall, 2018, 2006). These lexical retrieval difficulties may enhance the advantages of discourse flexibility despite the increased communicative load, compared with constrained tasks. However, due to comprehension difficulties, tasks with written and/or picture support may be better performed than those reliant only on verbal instructions, so picture-based stories or descriptions may be easier than producing a personal narrative or holding a conversation.

Furthermore, the type of treatment provided may differentially influence changes in discourse measures. For example, treatment for improving single-word noun retrieval (nSFA) in a participant with moderate Wernicke’s aphasia resulted in salient improvements to confrontation naming relative to minimal improvement to some aspects of discourse (Boyle, 2004). Conversely, Edmonds et al. (2014), (2015), Knoph et al. (2017), and Ulatowska and Chapman (1989) describe three participants with Wernicke’s aphasia (two with moderate-to-severe aphasia, one with moderate severity inferred from the description of the participant) who received verb-based treatments in a sentence context or discourse therapy. The treatments received included vSFA, VNeST, CBT, and discourse treatment through written home assignments. All three participants were reported to have more salient improvements in less-constrained tasks such as discourse than in more-constrained tasks such as confrontation naming of verbs or picture-based sentence construction.

However, aphasia severity may interact with the type of treatment to affect outcome measures, as observed by Rogalski et al. (2013). They found that after ARCS treatment, their participant with moderate Wernicke’s aphasia improved in both accuracy of confrontation naming and informativeness of words during the discourse. By contrast, their participant with severe Wernicke’s aphasia showed no improvement in either task (Rogalski et al., 2013). It may be that ARCS was not effective in this participant with severe Wernicke’s aphasia (at least after only 15 hours of treatment), but it may also be that the outcome measures were not sensitive enough to detect change in this participant.

While there is some evidence that discourse analysis is more sensitive to change than formal assessments (Dietz & Boyle, 2018a), conducting discourse analysis can be challenging and time-consuming. To date, discourse analysis has been utilized inconsistently both by researchers of Wernicke’s aphasia (e.g., Kurland et al., 2010) as well as clinically by SLTs assessing language impairments in aphasia (Bryant et al., 2017).

Aim of the current study

Due to the paucity of literature on severe Wernicke’s aphasia, and the inconsistency of discourse analysis in this population, in this study we aimed to identify whether using less-constrained language tasks, such as elicited discourse production, would be more, less, or equally sensitive to measuring treatment effects after a verb-based sentence-level treatment than more-constrained language tasks, such as picture naming and sentence construction, in a participant with severe Wernicke’s aphasia.

Thus, we asked the following research questions:

(1) Is it possible to rehabilitate language abilities in a participant with severe Wernicke's aphasia using a verb-based sentence-level treatment (VNeST) that has been observed to be successful for other types of severe aphasia?
(2) If language abilities improve after VNeST, are more salient improvements observed for less-constrained language tasks or for more-constrained language tasks?

We hypothesised that VNeST would improve language skills in a participant with severe Wernicke's aphasia, due to the treatment's focus on lexical-semantic strengthening, which in turn should result in better lexical retrieval (Marshall, 2018, 2006). Furthermore, due to the combination of both production and comprehension of sentences during treatment, improvement was also expected for comprehension skills, known to be specifically impaired in Wernicke's aphasia (Robson et al., 2019). Based on previous research with participants with Wernicke's aphasia who received sentence or discourse-level treatment (Knoph et al., 2017; Ulatowska & Chapman, 1989), including VNeST (Edmonds et al., 2014, 2015), we further hypothesised that more salient improvements in production would be observed for less-constrained tasks (discourse and answering WH-questions) compared to more-constrained tasks (confrontation naming of nouns and verbs, and picture-based sentence construction).

Methodology

Participant

The participant was a 78-year-old male multilingual speaker of English (native language), Modern Hebrew, and Yiddish (both acquired in early childhood), with 16 years of formal education. He had a left hemisphere ischemic cerebrovascular accident (CVA) in 2017 at age 76 years, resulting in a diagnosis of global aphasia which began resolving into Wernicke's aphasia in the months after the stroke. A CT scan at the time of the stroke showed an acute lesion in the territory of the left middle cerebral artery. The participant also reported receiving speech and language treatment after his stroke for about 1 year, at home and primarily in English. He concluded home-treatment 1-month prior to participating in the study. Fourteen months after his stroke, the participant was recruited to take part in our research study. Consent was obtained using a combination of verbal and written information – his reading comprehension was better spared than his auditory comprehension.

The participant acquired English as his first language, and it remained his most proficient language across the lifespan. He acquired Modern Hebrew and Yiddish in childhood, through formal teaching in school (both) and some use within the community (Yiddish). The participant moved from the U.S. to Israel at age 27, living in a community where English, Hebrew and Yiddish were all spoken regularly. Pre-stroke, the participant used English and Hebrew daily, and Yiddish only occasionally. He rated his pre-stroke English as native-like, and his Hebrew and Yiddish both at high proficiency for understanding and speaking (he did not report on pre-stroke literacy in Hebrew and Yiddish).

Post-stroke, the participant reported that his English remained his least impaired language, followed by Hebrew, and then Yiddish.

Based on the Western Aphasia Battery-Revised (WAB-R – Kertesz, 2006), the participant's post-stroke English language abilities were similar to his Hebrew language abilities for all subtests. Overall, his pre-treatment scores on the WAB-R (Kertesz, 2006) indicated that the participant had severe Wernicke's aphasia both in English (Aphasia Quotient (AQ) = 34.7) and in Hebrew (AQ = 32.0), with a language profile of impaired comprehension, and severe anomia evidenced by extensive use of neologisms and empty words and phrases. Although this participant was originally recruited as part of a larger study that aimed to assess both his English and Hebrew abilities, due to circumstances described in detail below we were only able to collect baseline language assessments in Hebrew, but not post-treatment data. This paper, therefore, focuses on his first-language English language abilities.

The participant reported having no difficulties with hearing and no difficulty with vision when wearing his corrective glasses. Results from the non-linguistic subtests in the Cognitive Linguistic Quick Test (CLQT – Helm-Estabrooks, 2001) indicated that the participant had a mild non-linguistic cognitive impairment in only one subtest (clock drawing), but cognition was within normal limits for attention, executive functions and visuospatial skills. See Table 1 for pre-treatment CLQT results and pre-treatment WAB-R AQ scores in English.

Design and procedure

This study takes a within-participant multiple-baseline approach with three phases: (1) pre-treatment baseline testing (3-time-points), (2) VNeST treatment block (in English), (3) post-treatment testing (2-time-points). Treatment effects were measured by monitoring performance during treatment (direct treatment effects) as well as by comparing post-treatment testing scores to pre-treatment baseline scores.

Table 1. Linguistic and non-linguistic abilities at baseline testing (pre-treatment), based on the WAB-R (Kertesz, 2006) and the CLQT (Helm-Estabrooks, 2001).

Task [max score]	Score
WAB-R (English)	
Spontaneous speech score [20]	10
Auditory Verbal Comprehension Score for AQ [10]	5.25
Repetition Score [10]	1.6
Naming and Word Finding Score [10]	0.5
TOTAL: [50]	17.35
Aphasia Quotient (TOTAL * 2) [100]	34.7
CLQT	
Symbol Cancellation [12]	11
Clock Drawing [13]	10
Symbol Trails [10]	9
Design Memory [6]	4
Mazes [8]	8
Design Generation [13]	5

Note. WAB-R = Western Aphasia Battery-Revised; CLQT = Cognitive Linguistic Quick Test.

In phase 1, the participant's language skills were assessed using an aphasia assessment battery developed by Lerman and Goral (unpublished), which includes both production and comprehension subtests, at the word, sentence and discourse levels. No writing tasks were administered in the aphasia assessment battery because the participant reported rarely writing in his daily life before the stroke. Similarly, the reading and writing section of the WAB-R was eliminated at the participant's request (due to fatigue from the long testing sessions).

We administered the following language tasks from our aphasia assessment battery:

(1) Picture-based action naming – partially based on a subset of stimuli from the Action Naming Test, which is a subtest of the Verb and Sentence Test (VAST) (Bastiaanse et al., 2002)
(2) Picture-based object naming – partially based on a subset of stimuli from the Multilingual Naming Test (Gollan et al., 2012)
(3) Picture-based sentence construction of SV and SVO sentences – partially based on a subset of stimuli from the Object and Action Naming Battery (Druks & Masterson, 2000)
(4) Answering verbally presented WH-questions
(5) Comprehension subtests – each word or sentence was presented auditorily and the participant was asked to point to the picture corresponding to what he heard, out of four options
    (a) Comprehension of nouns – largely based on a subset of stimuli from the Peabody Picture Vocabulary Test (Dunn & Dunn, 2007), where the distractors were all semantically related to the target noun
    (b) Comprehension of verbs – largely based on a subset of stimuli from the VAST (Bastiaanse et al., 2002), where the distractors were a semantically related verb and two nouns semantically related to each verb (the target and the distractor)
    (c) Comprehension of reversible sentences – also largely based on the VAST (Bastiaanse et al., 2002), but with a different subset of stimuli, where distractors were role reversal, lexical distractors or both
(6) Discourse production was elicited for procedural, expositional, and narrative discourse, as recommended by Brookshire and Nicholas (1994), with 80% of discourse production based on pictures (picture descriptions, story sequences), and 20% based on written and verbal directions (procedural and personal narratives)
(7) Non-word repetition (a control task) – a subtest from the Psycholinguistic Model of Language Processing (PALPA) (Kay, Lesser, & Coltheart, 1996), expected to be impaired in a participant with Wernicke's aphasia due to difficulty with both auditory and phonological skills (Robson et al., 2019) but not expected to improve based on the lexical-semantic treatment provided.

In phase 2, the participant received 24 hours of Verb Network Strengthening Treatment (VNeST) in English. VNeST was chosen because it is a predominantly lexical-semantic treatment, which is a likely point of impairment associated with neologistic jargon (e.g., Marshall, 2018, 2006). At the same time, it is a multi-modality treatment, using both spoken and written language (an advantage for this participant, based on his better-spared reading comprehension compared with auditory comprehension). Treatment was spread over 6 weeks for an average of three times a week, with sessions lasting for between 45 minutes to 1 hour 45 minutes at a time, depending on the participant's tiredness and motivation. Treatment was based on a published protocol (Edmonds, 2014), whereby participants are presented with a written verb, asked to read it aloud, then to write it down, and then to produce four subject-verb-object (SVO) sentences using the verb. When necessary, cues are provided for the participants: either minimal cues (such as directing the participant to a certain topic) or maximal cues (providing four written options for the participant to choose from). Following this, participants are asked to read the sentences aloud, to expand on one sentence with WH-question (where, when, why) prompts, and then to make semantic judgements on SVO sentences read aloud by the SLT, still using the same verb. Finally, participants are asked to recall the verb, and to independently produce some SVO sentences using the same verb.

The only difference between the published protocol and our treatment block involved the lack of use of minimal cues. Our participant did not benefit from minimal cueing at all, and so instead of using minimal cueing we used an assisted retrieval technique. That is, when the participant was clear as to what he intended to say but was only able to say something semantically or phonologically similar (e.g., consistently saying "my mother" while pointing to his wife, or "something we eat … /kjuke/" meaning cucumber), the SLT would verbally provide the intended response. Therefore, during our treatment block, the participant either independently retrieved subjects and objects related to the given verb, retrieved subjects and objects together with the SLT using our assisted retrieval technique, or retrieved subjects and objects only after he was provided with a maximal cue.

During the treatment block, the participant practiced 20 different verbs over 17 sessions, and each verb was trained between 2 and 3 times. Regarding treatment fidelity, the treating SLT filled out a chart for each verb in every session which tracked the different stages of each verb cycle. These charts, together with 10% of treatment sessions that were observed live by a second SLT, resulted in a very high calculation of treatment fidelity (over 99% of the stages of the VNeST protocol per verb were carried out as required).

Following the treatment block, phase 3 included reassessment of the participant's language skills using the aphasia assessment battery only. All testing sessions were audio-taped and then later transcribed and scored with inter-rater reliability for both transcribing (on 25% of the data) and scoring (on 33.3% of the data). Transcribing reliability was found to be high (over 99% of words were transcribed correctly). Blind scoring was conducted on all the data at the conclusion of data collection, and reliability was found to be high with point-by-point rating (over 95% agreement) and with Krippendorff's alpha (at $\alpha > 0.9884$). Outcome measures were task-specific and were either quantitative, qualitative, or both. See Table 2 for a summary of outcome measures and relevant statistical tests employed.

At baseline, the participant was assessed in both English and Hebrew, with testing sessions conducted by different testers, one in English and one in Hebrew (both experienced SLTs). However, throughout this study, the participant displayed low motivation, although he was mostly compliant during baseline testing. His compliance and motivation improved during the treatment block until the last session. Due to personal and medical reasons, the participant's compliance dropped drastically during post-treatment testing, particularly during Hebrew testing, to the point where he requested to drop out of the study. We thus terminated the study and were therefore unable to adequately assess his Hebrew language skills post-treatment; we report here on his English language skills only, pre- and post-treatment in English.

Table 2. Language tasks and outcome measures for pre- and post-treatment assessment.

Task	Quantitative Measure	Statistical analyses	Qualitative measure
Direct treatment effects	Independent retrieval of subject/ object	Correlation	
	Assisted retrieval of subject/object	Correlation	
	Retrieval of subject/object after a maximal cue	Correlation	
	Independent SVO sentence production	Correlation	
Action naming (pictures)	Accuracy	McNemar, ES, NAP	Error analyses
Object naming (pictures)	Accuracy	McNemar, ES, NAP	Error analyses
Sentence construction (pictures)	Accuracy of subject retrieval	McNemar, ES, NAP	
	Accuracy of verb retrieval	McNemar, ES, NAP	Error analysis
	Accuracy of object retrieval	McNemar, ES, NAP	
	Relevant SVO sentence production	McNemar, ES, NAP	
Answering WH-questions	Relevant verb retrieval	ES, NAP	
	Relevant SVO sentence production	ES, NAP	
Noun comprehension	Accuracy	McNemar	No. of stimulus repetitions before eliciting a response
Verb comprehension	Accuracy	McNemar	No. of stimulus repetitions before eliciting a response
Sentence comprehension	Accuracy	McNemar	No. of stimulus repetitions before eliciting a response
Discourse[a]	TVUs[b]	ES, NAP	
	No. of nouns	ES, NAP	
	No. of verbs	ES, NAP	
	CIUs[c]	ES, NAP	
	CUs[d]	ES, NAP	
	Noun and verb type-token ratios		
	%CIUs		
	Rate = CIUs/min	ES, NAP	
Non-word repetition	Accuracy	n/a[e]	

Note. SVO = Subject-Verb-Object; ES = effect size; NAP = Non-overlap of All Pairs; TVUs = Total Verbal Units; CIUs = Correct Information Units; CUs = Complete Utterances; %CIUs = CIUs/TVUs.

[a]Discourse was analysed for overall production and for relevant and informative production. Measures of quality of production are especially important in Wernicke's aphasia, since an increase in production does not necessarily indicate improvement. Rather, an increase in informativeness (together with any or no change in the amount of production) will indicate improvement (Rogalski et al., 2013).

[b]Total Verbal Units included whole and part words, word produced in the non-target language, and extended filled pauses. (Although this participant produced a lot of neologistic jargon, there were very few instances where it was difficult to separate the neologisms into words based on syllable stress and inter-word pauses. In the face of uncertainty, the whole neologism was counted as one TVU.)

[c]Correct Information Unit counts were based on Nicholas and Brookshire (1993).

[d]Complete Utterances are relevant information provided in an SVO sentence structure, as described in Edmonds et al. (2009).

[e]Non-word repetition was at floor-level, so statistical testing was not conducted.

Results

Direct treatment effects

Across the treatment sessions, there was a significant increase in the participant's ability to retrieve a subject or object independently. At the same time, significant decreases were observed in the participant's need for assisted retrieval or

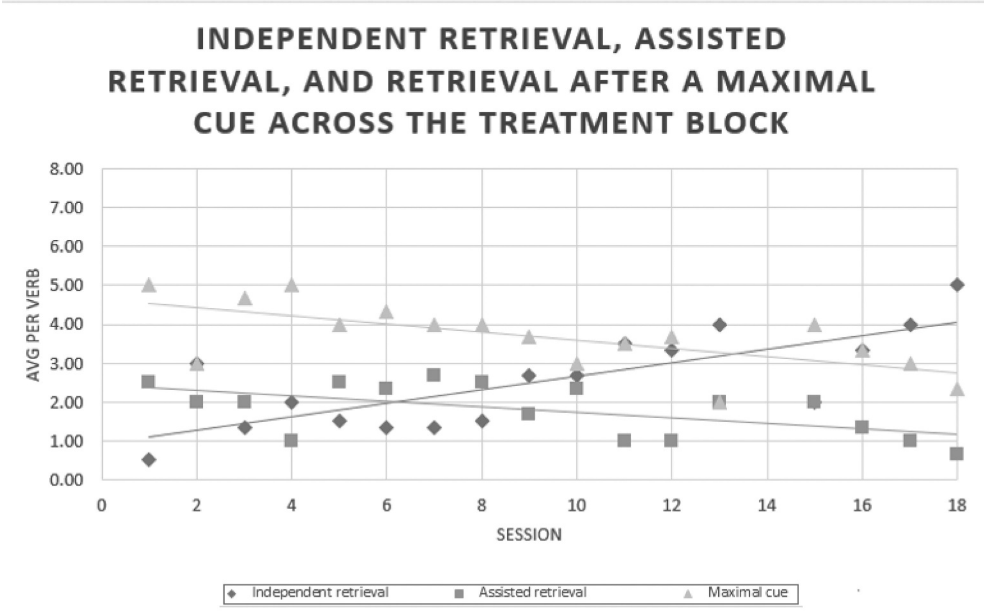

Figure 1. Independent retrieval, assisted retrieval, and retrieval after a maximal cue across the treatment block.

a maximal cue in order to retrieve a subject or object (see Figure 1). The correlation between session and independent retrieval was positive and significant, r(16) =.773, *p* = <.001. The regression coefficient *(b)* = .189, *p* = <.001 indicates that for every additional treatment session there was a significant increase of .189 in the average independent retrieval of subjects and objects. The correlation between session and assisted retrieval was negative and significant, r(16) = −.563, *p* = .019. The regression coefficient *(b)* = −.074, *p* = .019 indicates that for every additional treatment session there was a significant decrease of .074 in the average assisted retrieval of subjects and objects. The correlation between session and retrieval after a maximal cue was also negative and significant, r(16) = −.677, *p* = .003. The regression coefficient *(b)* = −.114, *p* = .003 indicates that for every additional treatment session there was a significant decrease of .114 in the average retrieval of subjects and objects after receiving a maximal cue.

Furthermore, during the final stages of each verb cycle, there was a significant increase in independent production of relevant SVO sentences. The average number of independent SVO sentences produced per verb was calculated (maximum of 4 per verb, see Figure 2) and the correlation between session and independent SVO sentence production was positive and significant: r(16) = .919, *p* < .0001. The regression coefficient *(b)* = .107, *p* < .0001, indicates that for every additional treatment session there was a significant increase of .107 in the average independent production of SVO sentences in the final stages of each verb cycle.

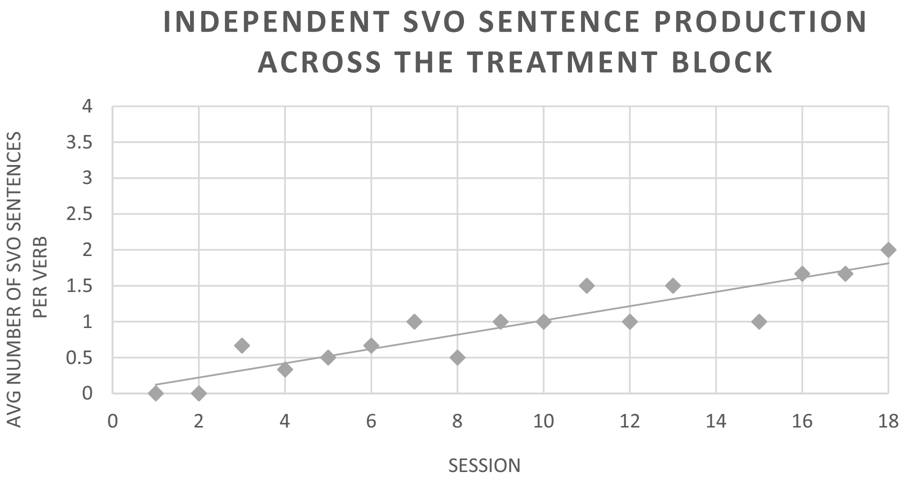

Figure 2. Independent SVO sentence production across the treatment block.

Comparing post-treatment language skills to pre-treatment language skills

For the comprehension of nouns, verbs and sentences, no significant changes were observed, as measured by the McNemar test of equal change (see Table 3). However, for all three comprehension tasks, the need to repeat the stimuli before the participant responded decreased. Specifically, for noun comprehension, 50.0% of the stimuli were repeated pre-treatment, 18.75% post-treatment; for verb comprehension, 52.38% of the stimuli were repeated pre-treatment, 28.57% post-treatment; for sentence comprehension, 55.56% of the stimuli were repeated pre-treatment, 33.33% post-treatment. In other words, the participant needed fewer repetitions of stimuli post-treatment as compared to pre-treatment to maintain stable comprehension abilities. A closer look at the pre- and post-treatment comprehension data shows that errors for verb comprehension were either distractor verbs or a noun semantically related to the target, but never a noun semantically related to the distractor verb. By contrast, errors for sentence comprehension were spread across all options (role reversal, lexical distractors or both).

For the production tasks, raw scores were relatively low for all measures of noun and verb retrieval and SVO sentence production in the picture-based single-word and sentence tasks (object retrieval, action retrieval, and sentence construction). The McNemar test of equal change indicated that no change was observed for any measure. Based on

Table 3. Noun, verb and sentence comprehension, pre- and post-treatment.

		Accuracy (%)		
		Pre-English treatment	Post-English treatment	McNemar[a]
Comprehension	Nouns[b]	70.83	75.0	0
	Verbs[b]	71.43	71.43	0.2
	Sentences[b]	50.0	41.67	0.67

[a]Significance: > 3.84 for $p < .05$; trend towards significance: > 2.71 for $p < .1$.
[b]Pre-treatment, all data was collected (Nouns n = 24; Verbs n = 21; Sentences n = 18). Post-treatment, 2/3 of the data was collected (Nouns n = 18; Verbs n = 14; Sentences n = 12). The McNemar was calculated on the 2/3 of the stimuli that were collected both pre- and post-treatment.

effect sizes together with Non-overlap of All Pairs (NAP) ratings, a decrease was observed for SVO sentence production in the sentence construction task, as well as verb use when answering WH-questions. See Table 4 for single-word naming and sentence measures pre- and post-treatment.

An error analysis showed that for action naming, the percentage of incorrect responses that included non-words (neologisms) decreased from 48.78% to 37.04% pre- to post-treatment. By contrast, the percentage of incorrect responses that included only real words (e.g., semantic paraphasias, general verbs, nouns, descriptions, empty speech without jargon) increased from 51.22% to 62.96% pre- to post-treatment. For object naming, an error analysis showed that the percentage of incorrect responses that included non-words remained fairly stable at 63.41% pre-treatment and 65.38% post-treatment, and the percentage of incorrect responses that included only real words (e.g., semantic paraphasias, descriptions, empty speech without jargon) also remained fairly stable at 36.58% pre-treatment, and 34.62% post-treatment. Non-word repetition (our control task) was at floor-level, both pre-treatment (1/20 non-words repeated correctly) and post-treatment (0/20 non-words repeated correctly).

For discourse, less output was produced overall post-treatment compared to pre-treatment, as indicated by the TVU count of 600 verbal units pre-treatment compared with 368 verbal units post-treatment. Furthermore, all raw scores of the other measures were lower post-treatment than pre-treatment, except for rate of relevant production (CIUs/min). Measures of the TVU count, number of CUs, number of nouns, number of different nouns, number of verbs, and number of different verbs all decreased based on negative small to large effect sizes and negative medium to high NAP ratings (See Table 5). When discourse production was examined further, the quality of production was observed to increase based on an increase in noun type-token ratios (from 0.70 to 0.78), in verb type-token ratios (from 0.67 to 0.79) and in the %CIUs (from 28.67% to 42.39%). Furthermore,

Table 4. Single-word noun and verb retrieval, sentence construction measures, and verb use in answering WH-questions in English, pre- and post-English treatment.

		Accuracy (%)				
		Pre-treatment	Post-treatment	McNemar[a]	Effect size[b]	NAP rating[c]
Object naming[d]		8.89	13.33	2.0	0.58	0.17
Action naming[d]		8.89	10.00	0	0.29	0.17
Sentence construction[d]	SVO sentence production	7.41	2.78	0.33	−1.44	−0.67
	Subject retrieval	31.48	30.56	0	−0.29	0
	Verb retrieval	16.67	16.67	0.11	0	0
	Object retrieval	7.41	13.89	0.11	0.76	0.5
WH-questions[d]	Relevant verb use	50.00	30.00	n/a	−2.02	−0.83
	SVO sentence production	25.00	0	n/a	−1.15	−0.67

Note. NAP = Non-overlap of All Pairs.
[a]Significance: > 3.84 for p <.05; trend towards significance: > 2.71 for p <.1.
[b]Effect size: Small > 1.2, Medium > 1.7, and Large > 3.3.
[c]Using zero chance level: 0-.31 = weak, 32-.84 = medium, 85–1.0 = strong. A negative score indicates higher scores pre-treatment than post-treatment.
[d]Pre-treatment, all data was collected (Action naming n = 45; object naming n = 45; sentence construction n = 54; WH-questions n = 16). Post-treatment, 2/3 of the data was collected (Action naming n = 30; object naming n = 30; sentence construction n = 36; WH-questions n = 10). The McNemar was calculated on the 2/3 of the stimuli that were collected both pre- and post-treatment. Effect sizes and NAP scores were calculated for all the data collected.

Table 5. Discourse measures, pre- and post-treatment.

	Pre-treatment (*2/3)[a]	Post-treatment	Effect size[b]	NAP rating[c]
No. of total verbal units	900 (600)	368	−2.39	−1.0
No. of CUs	26 (17.3)	15	−1.30	−0.33
No. of nouns	106 (70.67)	37	−1.32	−1.0
No. of different nouns	74 (49.3)	29	−1.28	−0.83
No. of verbs	144 (96)	61	−3.69	−1.0
No. of different verbs	96 (64)	48	−2.08	−0.67
No. of CIUs	258 (172)	156	−0.59	−0.33
%CIUs	28.67	42.39	n/a	n/a
CIUs/min (rate)	31.59	41.05	0.40	0.67

Note. CU = Complete Utterance; CIU = Correct Information Units; %CIU = Correct information units out of the total verbal units; NAP = Non-overlap of All Pairs.

[a]Pre-treatment, discourse data were collected from three time-points; post-treatment, discourse data were collected from only two time-points.

[b]Effect size: Small > 1.2, Medium > 1.7, and Large > 3.3. Measured using all three time-points pre-treatment for the variance, but only the two time-points pre-treatment where the stimuli overlapped with the two time-points post-treatment to calculate the mean.

[c]Using zero chance level: 0-.31 = weak, 32-.84 = medium, 85–1.0 = strong. A negative score indicates higher scores pre-treatment than post-treatment. Measured using three time-points pre-treatment, and two time-points post-treatment.

there was an increase in the number of relevant utterances from 43/110 (39.1%) pre-treatment to 27/53 (50.9%) post-treatment.

Finally, it is important to note that the participant's wife provided verbal feedback on her husband's language and communication. She reported that since taking part in the study, the participant is better able to communicate with English-speaking, Hebrew-speaking or Yiddish-speaking friends who come to visit him, although he switches languages without being aware of it more often between Hebrew and English, and sometimes Yiddish, than he did before the study.

To summarise the results, significant improvement was observed during the treatment block. For the aphasia assessment battery (comparing post-treatment measures with a pre-treatment baseline), while either no changes were observed or a decline was observed for most measures of single-word naming, sentence production and raw scores of discourse measures, a number of specific changes were observed that indicate improvement from pre- to post-treatment. First, the participant required less repetition of stimuli to maintain his comprehension abilities. Second, fewer neologisms and proportionately more real-word output were observed for the action naming task. Third, for the discourse measures, less output was produced overall together with an increase in the quality of production (as measured by increased noun and verb type-token ratios, relevant utterances, %CIUs, and CIUs/min). These changes are consistent with the participant's wife's report of an improvement in his ability to communicate with friends.

Discussion

This study investigated whether it was possible to rehabilitate the language abilities of a participant with severe Wernicke's aphasia with a verb-based sentence-level treatment (VNeST), and, if so, whether improvements were more salient for less-constrained language tasks or for more-constrained language tasks. We anticipated that, overall, the participant's language abilities would be difficult to rehabilitate, as Robson et al. (2019) suggest, because his diagnosis of severe Wernicke's aphasia (including poor auditory comprehension and phonological skills) was still applicable more than a year after his stroke. Thus, we did not expect to observe substantial improvements in his language skills. However, we hypothesised that VNeST would result in some improvement to production and comprehension skills because the VNeST protocol focuses on strengthening the connections between the lexicon and the semantic system, particularly between verbs and their associated nouns, within the context of both sentence comprehension and production (Marshall, 2018, 2006; Robson et al., 2019). We found support for this hypothesis, with significant positive direct treatment effects observed during 24 hours of VNeST administered over 6 weeks, both for independent subject and object retrieval, and for independent relevant SVO sentence retrieval across sessions. Furthermore, limited generalisation was also observed to other tasks, as measured by post-treatment language abilities compared to pre-treatment language abilities.

We also expected this treatment generalisation to be more salient for less-constrained tasks compared to more-constrained tasks, based on a limited number of previous studies on participants with Wernicke's aphasia who received sentence-level or discourse-level treatment and showed improvement on less-constrained tasks such as discourse production (Edmonds et al., 2014, 2015; Knoph et al., 2017; Ulatowska & Chapman, 1989). We found partial support for this hypothesis because the most salient of the observed generalisation findings was observed for the discourse task, where the quality of production increased together with decreased overall output. This pattern of change represents an improvement for people with Wernicke's aphasia, since their output becomes more focused and informative and there is a reduction in empty, uninterpretable sentences (Rogalski et al., 2013).

In addition to the observed improvement in discourse, we also observed an increase in the proportion of real-word output rather than neologisms for action naming. This switch from neologisms to real-words indicates an improvement, as the lexical-semantic connections are presumably strengthened (Marshall, 2018, 2006), and may have been observed only for action naming due to the focus placed on verbs during treatment. Furthermore, we observed a potential improvement in comprehension in that the participant succeeded at the same accuracy level while being exposed to the stimuli a reduced number of times. This improvement to comprehension may also have been a direct result of VNeST, where semantic judgements are repeatedly made for auditorily presented sentences (Edmonds, 2014).

Conversely, another relatively less-constrained task – answering WH-questions – did not show any improvements pre- to post-treatment; rather we observed a decline in verb and SVO production. This decline may be related to the participant's difficulty with auditory comprehension, because the WH-questions were all presented auditorily (compared with the discourse measures that were mostly picture-based). Furthermore, the participant frequently responded to WH-questions by first asking questions such as "why are you asking me these questions?" or "what is the point of these questions?", and only then attempting a response. It is likely that the participant's comprehension difficulties of the questions contributed to the production difficulties in answering those questions and, in turn, may have been related to his resistance to this task. Comparing less-constrained tasks presented auditorily, via writing, or via pictures, would be an interesting direction for future research on participants with Wernicke's aphasia.

We did not expect our participant to improve on all language tasks, based on previous observations of individual differences among participants who received VNeST and their specific impairments relative to each task (e.g., Edmonds et al., 2009; Edmonds & Babb, 2011). It is encouraging, therefore, that some improvements to the participant's language were observed even with a diagnosis of severe Wernicke's aphasia 14 months post-stroke. While we acknowledge that the syndrome approach to diagnosing aphasia is not always useful, in this case, our participant showed the classic signs of Wernicke's aphasia, and, most crucially, his improvements were directly related to specific impairments expected in Wernicke's aphasia (better comprehension skills, reduced neologism production, reduced amount of output, and increased informativeness of output). These improvements are in contrast to other tasks that were not expected to improve after VNeST, and which indeed did not, such as in the non-lexical phonological control task (non-word repetition), which was at floor-level pre-treatment and showed no change post-treatment.

Indeed, our results, are comparable to those of Edmonds et al. (2014, 2015), who also treated a participant with Wernicke's aphasia (moderate-to-severe) with 35 hours of VNeST across 10 weeks, and found that the most salient results were for discourse (including for increased informativeness relative to overall output). However, single-word naming of nouns and verbs also improved, with minimal or no improvement to sentence construction or sentence comprehension (as measured by accuracy). Our results showed increased informativeness relative to overall output in discourse, as well as improvement to single-word action naming as measured by reduced neologisms. We too found no improvement in sentence construction or sentence comprehension as measured by accuracy. Rather, we observed improved comprehension of nouns, verbs and sentences when considering the number of repetitions required in the auditory comprehension tasks, together with stable accuracy scores.

We interpret these results within the framework of VNeST, which is systematic in that it follows a very specific protocol, and is relatively flexible in that it allows for the retrieval of a variety of lexical items (rather than one specific target item) per turn. When Edmonds and colleagues first developed VNeST, it was for Broca's aphasia, conduction aphasia, and transcortical motor aphasia (Edmonds et al., 2009; Edmonds & Babb, 2011). The researchers built into the VNeST protocol repetition of specific language structures conducted on a rotating list of verbs, in order to strengthen the semantic verb network within an SVO framework (including building arguments around a verb and retrieving thematic roles relative to those arguments). For people with Wernicke's aphasia generally, and our participant specifically, building an SVO sentence structure is not a likely place of impairment (Robson et al., 2019). Instead, strengthening of the semantic network and its connections to the lexicon resulted in better quality of production rather than an improvement of specific words. This was observed both in increased informativeness of discourse and reduced neologisms in action naming. Finally, we interpret the participant's improved comprehension scores as a direct result of the process of VNeST training, because the treatment protocol includes responding to auditorily presented WH-questions and making semantic feasibility judgements on auditorily presented sentences.

It is possible that our participant did not show quantitative improvements to single-word object and action naming, whereas Edmonds et al.'s (2014, 2015) participant did, due to the discrepancy in treatment hours. Our participant received 24 hours of treatment over 6 weeks before his motivation began to decline and it was decided, together with the participant, to end the first treatment block earlier than planned and to re-assess his language skills (during which he requested to terminate his participation in the study completely). Edmonds et al.'s participant with Wernicke's aphasia received 35 hours of treatment (Edmonds et al., 2014, 2015) and it may be that this difference contributed to the discrepancy in the results. This explanation is supported when we compare our results to the study by Rogalski et al. (2013) who provided less treatment to their participant with Wernicke's aphasia (15 hours) and observed no improvement to either confrontation naming or discourse. Furthermore, we can gain insight from comparing the participants in two studies by Edmonds and colleagues with different aphasia severities who received a different number of treatment hours. In one study (Edmonds & Babb, 2011) the researchers provided VNeST to each of two participants with severe (Broca's) aphasia (one participant received 45 hours of treatment over 15 weeks, the other 37.5 hours of treatment over 12 weeks). The authors found less widespread improvements to these participants' language skills than that observed for participants with moderate (transcortical-motor or conduction) aphasia who received just 10–18 hours of VNeST each (Edmonds et al., 2009), suggesting that number of treatment hours for severe aphasia may need to pass a threshold in order to show robust treatment effects.

A strong limitation to our study was our participant's low motivation towards the end of the treatment block and his low compliance post-treatment. On the one hand, we expect that patients with severe Wernicke's aphasia may have low or fluctuating motivation levels due to their reduced understanding of their language deficit (Marshall, 2018, 2006). On the other hand, we were unable to provide more than 24 hours of treatment, and unable to collect a full set of post-treatment data, and the results that we obtained post-treatment are likely not fully reflective of our participant's best language abilities at that time-point. It is all the more encouraging, then, that direct treatment effects were observed and that we observed evidence of treatment generalisation, especially to the quality of the participant's discourse, although we are cautious in interpreting our results within these limitations. As discussed above, it may be that 24 hours of treatment in severe Wernicke's aphasia is not enough to result in more widespread improvements across different language tasks. It is possible that with more hours of treatment, we too would have found more robust improvements in our participant not only in the relatively less-constrained task of discourse but also in other constrained and less-constrained language tasks.

Furthermore, the fact that this is a case-study, and that the results are not extensive in their scope and strength limits the implications that can be drawn from them. However, people with severe Wernicke's aphasia are challenging to work with, especially under controlled research conditions, due to factors relating to consent, compliance and motivation to receive treatment, and therefore to date, there is very limited knowledge regarding treatment outcomes in this population. Our case-study provides another piece of the converging evidence that in people with severe Wernicke's aphasia, lexical-semantic treatments at the sentence or discourse levels are more likely to result in generalisation to less-constrained discourse tasks before generalising to more-constrained language tasks (Edmonds et al., 2014, 2015; Knoph et al., 2017; Ulatowska & Chapman, 1989), and that this generalisation is a result of treatment, rather than general improvement (Edmonds et al., 2014, 2015). We strongly encourage researchers to recruit, treat, and publish on participants in this population, to avoid underrepresentation in the literature and to support evidence-based practice in this population.

Conclusion

In our study, we observed improvement to some language skills in a participant with severe Wernicke's aphasia after a verb-based sentence-level treatment (VNeST). However, the commonly used measure of comparing pre- and post-treatment accuracy scores in constrained language tasks was not sensitive enough to detect change. Rather, improvement was observed through monitoring performance during treatment, and through pre- and post-treatment comparisons including discourse analysis, error analysis, and other qualitative measures such as the number of repetitions required in an auditory comprehension task, together with accuracy. We conclude that treatment studies and clinical work with people who have Wernicke's aphasia should always include the assessment of less-constrained language production tasks, such as discourse, because these tasks may be more sensitive to treatment effects than constrained picture-naming or sentence construction tasks, in this population. We recommend more studies on participants with severe Wernicke's aphasia generally, due to the dearth of research on this population, and strongly advocate the use of both quantitative and qualitative methods to analyse language data. Furthermore, the next step in research on participants with Wernicke's aphasia should not only include general discourse analysis but consider which types of discourse analysis may be most sensitive to change, taking into account the flexibility of the task together with the need to rely on auditory comprehension in order to complete the task. For example, future research could compare more natural and open discourse tasks that rely heavily on auditory comprehension (e.g., conversational analysis), or potentially rely less heavily on auditory comprehension (e.g., narrating a personal story), with more constrained and picture-based discourse tasks that rely minimally on auditory comprehension (e.g., picture descriptions or story sequences).

Acknowledgments

We thank the participant and his family for their cooperation in this study. We also thank Dorit Mais for help with recruitment and Einat Galezer for collecting all the language data in Hebrew. Thanks also to Dr. Heather Harris Wright in her role as editor and to an anonymous reviewer for their helpful suggestions and comments on a previous draft of this paper.

Disclosure statement

The authors report no conflict of interest.

Funding

This work was supported by the Graduate Center, CUNY under the Graduate Center Dissertation Year Award.

References

Armstrong, E. (2000). Aphasic discourse analysis: The story so far. *Aphasiology*, *14*(9), 875–892. https://doi.org/10.1080/02687030050127685

Bandur, D. L., & Shewan, C. M. (2008). Language-oriented treatment: A psycholinguistic approach to aphasia. In R. Chapey (Ed.), *Language intervention strategies in aphasia and related neurogenic communication disorders* (5th ed., pp. 756–799). Lippincott, Williams and Wilkins.

Bastiaanse, R., Edwards, S., & Rispens, J. (2002). *Verb and Sentence test (VAST)*. Thames Valley Test Company.

Boyle, M. (2004). Semantic Feature Analysis treatment for anomia in two fluent aphasia syndromes. *American Journal of Speech-Language Pathology*, *13*(3), 236–249. https://doi.org/10.1044/1058-0360(2004/025)

Brookshire, R. H., & Nicholas, L. E. (1994). Test-retest stability of measures of connected speech in aphasia. *Clinical Aphasiology*, *22*, 119–133.

Bryant, L., Spencer, E., & Ferguson, A. (2017). Clinical use of linguistic discourse analysis for the assessment of language in aphasia. *Aphasiology*, *31*(10), 1105–1126. https://doi.org/10.1080/02687038.2016.1239013

Coelho, C. A., & Flewellyn, L. (2003). Longitudinal assessment of coherence in an adult with fluent aphasia: A follow-up study. *Aphasiology*, *17*(2), 173–182. https://doi.org/10.1080/729255216

Coelho, C. A., Liles, B. Z., Duffy, R. J., Clarkson, J. V., & Elia, D. (1994). Longitudinal assessment of narrative discourse in a mildly aphasic adult. *Clinical Aphasiology*, *22*, 145–155.

Damasio, A. (1992). Aphasia. *The New England Journal of Medicine*, *326*(8), 531–539. https://doi.org/10.1056/NEJM199202203260806

Dietz, A., & Boyle, M. (2018a). Discourse measurement in aphasia: Consensus and caveats. *Aphasiology*, *32*(4), 487–492. https://doi.org/10.1080/02687038.2017.1398814

Dietz, A., & Boyle, M. (2018b). Discourse measurement in aphasia research: Have we reached the tipping point? *Aphasiology*, *32*(4), 459–464. https://doi.org/10.1080/02687038.2017.1398803

Druks, J., & Masterson, J. (2000). *An object and action naming battery*. Psychology Press.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody picture vocabulary test (4th ed.)*. Pearson.

Edmonds, L. A. (2014). Tutorial for verb network strengthening treatment (VNeST): Detailed description of the treatment protocol with corresponding theoretical rationale. *Perspectives on Neurophysiology and Neurogenic Speech and Language Disorders*, *24*(3), 78. https://doi.org/10.1044/nnsld24.3.78

Edmonds, L. A., & Babb, M. (2011). Effect of verb network strengthening treatment in moderate-to-severe aphasia. *American Journal of Speech-Language Pathology*, *20*(2), 131. https://doi.org/10.1044/1058-0360(2011/10-0036)

Edmonds, L. A., Mammino, K., & Ojeda, J. (2014). Effect of Verb Network Strengthening Treatment (VNeST) in persons with aphasia: Extension and replication of previous findings. *American Journal of Speech-Language Pathology*, *23*(2), S312. https://doi.org/10.1044/2014_AJSLP-13-0098

Edmonds, L. A., Nadeau, S. E., & Kiran, S. (2009). Effect of Verb Network Strengthening Treatment (VNeST) on lexical retrieval of content words in sentences in persons with aphasia. *Aphasiology*, *23*(3), 402–424. https://doi.org/10.1080/02687030802291339

Edmonds, L. A., Obermeyer, J., & Kernan, B. (2015). Investigation of pretreatment sentence production impairments in individuals with aphasia: Towards understanding the linguistic variables that impact generalisation in Verb Network Strengthening Treatment. *Aphasiology*, *29*(11), 1312–1344. https://doi.org/10.1080/02687038.2014.975180

Faroqi-Shah, Y. (2012). Grammatical category deficits in bilingual aphasia. In M. R. Gitterman, M. Goral, & L. K. Obler (Eds.), *Aspects of multilingual aphasia* (pp. 158–170). Multilingual Matters.

Fleming, V. B. (2007). *Cognitive flexibility and spoken discourse in younger and older adults*. [Doctoral dissertation], University of Texas.

Gollan, T. H., Weissberger, G. H., Runnqvist, E., Montoya, R. I., & Cera, C. M. (2012). Self-ratings of spoken language dominance: A Multilingual Naming Test (MINT) and preliminary norms for young and aging Spanish–English bilinguals. *Bilingualism: Language and Cognition*, *15*(3), 594–615. https://doi.org/10.1017/S1366728911000332

Helm-Estabrooks, N. (2001). *Cognitive linguistic quick test*. The Psychological Corporation.
Huber, W. (1990). Text comprehension and production in aphasia: Analysis in terms of micro- and macro-structure. In Y. Joanette & H. Brownell (Eds.), *Discourse ability and brain damage: Theoretical and empirical perspective* (pp. 154–179). Springer-Verlag.
Ingles, J. L., Mate-Kole, C. C., & Connolly, J. F. (1996). Evidence for multiple routes of speech production in a case of fluent aphasia. *Cortex*, *32*(2), 199–219. https://doi.org/10.1016/S0010-9452(96)80047-5
Kavé, G., & Goral, M. (2017). Do age-related word retrieval difficulties appear (or disappear) in connected speech? *Aging, Neuropsychology, and Cognition*, *24*(5), 508–527. https://doi.org/10.1080/13825585.2016.1226249
Kay, J., Lesser, R., & Coltheart, M. (1996). Psycholinguistic assessments of language processing in aphasia (PALPA): An introduction. *Aphasiology*, *10*(2), 159–180. https://doi.org/10.1080/02687039608248403
Kertesz, A. (2006). *Western Aphasia Battery—Revised*. Pro-Ed.
Knoph, M. I. N., Simonsen, H. G., & Lind, M. (2017). Cross-linguistic transfer effects of verb-production therapy in two cases of multilingual aphasia. *Aphasiology*, *31*(12), 1482–1509. https://doi.org/10.1080/02687038.2017.1358447
Kurland, J., Baldwin, K., & Tauer, C. (2010). Treatment-induced neuroplasticity following intensive naming therapy in a case of chronic Wernicke's aphasia. *Aphasiology*, *24*(6–8), 737–751. https://doi.org/10.1080/02687030903524711
Marshall, J. (2006). Jargon aphasia: What have we learned? *Aphasiology*, *20*(5), 387–410. https://doi.org/10.1080/02687030500489946
Marshall, J. (2018). Therapy for people with jargon aphasia. In P. Coppens & J. Patterson (Eds.), *Aphasia Rehabilitation: Clinical Challenges* (pp. 73–99). Jones & Bartlett Learning.
Nicholas, L. E., & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *Journal of Speech, Language, and Hearing Research*, *36*(2), 338–350. https://doi.org/10.1044/jshr.3602.338
Nicholas, M., Obler, L. K., Albert, M. L., & Helm-Estabrooks, N. (1985). Empty speech in Alzheimer's disease and fluent aphasia. *Journal of Speech Language and Hearing Research*, *28*(3), 405. https://doi.org/10.1044/jshr.2803.405
Olness, G. S. (2006). Genre, verb, and coherence in picture-elicited discourse of adults with aphasia. *Aphasiology*, *20*(2–4), 175–187. https://doi.org/10.1080/02687030500472710
Pedersen, P. M., Vinter, K., & Olsen, T. S. (2004). Aphasia after stroke: Type, severity and prognosis. *Cerebrovascular Diseases*, *17*(1), 35–43. https://doi.org/10.1159/000073896
Robson, H., Griffiths, T. D., Grube, M., & Woollams, A. M. (2019). Auditory, phonological, and semantic factors in the recovery from Wernicke's aphasia poststroke: Predictive value and implications for rehabilitation. *Neurorehabilitation and Neural Repair*, *33*(10), 800–812. https://doi.org/10.1177/1545968319868709
Rogalski, Y., Edmonds, L. A., Daly, V. R., & Gardner, M. J. (2013). Attentive Reading and Constrained Summarisation (ARCS) discourse treatment for chronic Wernicke's aphasia. *Aphasiology*, *27*(10), 1232–1251. https://doi.org/10.1080/02687038.2013.810327
Thomson, J., Gee, M., Sage, K., & Walker, T. (2018). What 'form' does informal assessment take? A scoping review of the informal assessment literature for aphasia. *International Journal of Language & Communication Disorders*, *53*(4), 659–674. https://doi.org/10.1111/1460-6984.12382
Ulatowska, H. K., & Chapman, S. B. (1989). Discourse considerations for aphasia management. *Seminars in Speech and Language*, *10*(4), 298–314. https://doi.org/10.1055/s-2008-1064270