*Top Lang Disorders*
Vol. 41, No. 1, pp. 5–26


# Measuring Change at the Discourse-Level Following Conversation Treatment

## Examples From Mild and Severe Aphasia

***Gayle DeDe and Elizabeth Hoover***

**Purpose:** This article reviews four discourse measures and examines whether they are sensitive to impairments in people with both mild and severe aphasia. We also ask whether these measures were sensitive to effects of conversation treatment in two case examples. **Method:** Two people with aphasia, one mild and fluent and the other severe and nonfluent, served as case studies. Both case studies had participated in conversation treatment, in which individualized goals were targeted in the context of naturalistic conversation-based interactions. Picture descriptions were analyzed using three discourse measures: core lexicon, words per minute, and correct information units. In addition, words per minute and conversation turns were examined in personal narratives produced by the individual with severe nonfluent aphasia in a conversational context. **Results:** For the individual with mild aphasia, both words per minute and core lexicon were sensitive to the presence of aphasia and treatment changes. For the individual with severe aphasia, all measures were sensitive to the presence of aphasia, but only words per minute and number/type of conversation turns were sensitive to effects of treatment. **Discussion/Conclusions:** Discourse measures capture relevant aspects of communication that may not be seen on standardized measures of discrete language skills. Given different aphasia profiles and individual communication goals, clinicians need to choose the most relevant, reliable, and informative measures. **Key words:** *aphasia*, *conversation*, *discourse measurement*, *outcome measurement*, *treatment*

A SIGNIFICANT CHALLENGE for both clinically and research-focused aphasiologists is identifying outcome measures that are both psychometrically sound and sensitive to meaningful changes in communication ability across a wide range of impairments. Part of this challenge relates to identifying what counts as a meaningful change in communication ability. Worrall et al. (2011) asked people with aphasia to identify their own goals for rehabilitation. Themes included returning to prestroke life, regaining communication ability, reestablishing independence, and engaging in social, leisure, and work activities. However, these types of goals are difficult to evaluate in a systematic way because their meaning will necessarily vary from person to person, as a function of individual differences in lifestyle, interests, and aphasia severity, among other factors. Nonetheless, achieving these goals is likely to require changes in production and comprehension of discourse.

***Author Affiliations:*** *Department of Communication Sciences and Disorders, Temple University, Philadelphia, Pennsylvania (Dr DeDe); and Department of Speech, Language, and Hearing Sciences, Boston University, Boston, Massachusetts (Dr Hoover).*

*This work was funded by R21DC015859 from the NIDCD.*

*Author disclosures can be found at http://links.lww.com/TLD/A73.*

***Corresponding Author:*** *Gayle DeDe, PhD, Department of Communication Sciences and Disorders, Temple University, 1701 N 13th St, Philadelphia, PA 19122 (gayle.dede@temple.edu).*

*DOI: 10.1097/TLD.0000000000000243*

The purpose of this article is to discuss examples of discourse measures that may be sensitive to both mild and severe aphasia profiles and to illustrate how they may capture treatment changes. A recent survey revealed that clinicians' use of discourse measures was limited by multiple constraints, including time (78% of respondents) and lack of certainty regarding selection of appropriate measures (71% of respondents) (Cruice et al., 2020). In our own work studying conversation treatment, we found it difficult to identify discourse measures that were reliably sensitive to group changes (DeDe et al., 2019). The present article emerged from an exercise we conducted to examine whether four different measures were sensitive to the presence of both mild and severe aphasia and to changes following conversation treatment. Thus, this article is not intended to provide a comprehensive review of discourse measures. Instead, we focus on our experiences using a set of discourse measures in our study of conversation treatment (DeDe et al., 2019).

## CONVERSATION TREATMENT

Conversation treatment focuses on communicative interaction rather than improvement in discrete linguistic skills such as word retrieval or sentence production. The theoretical motivation underlying conversation treatment is multifaceted (e.g., DeDe et al., 2019; Elman, 2007; Elman & Bernstein-Ellis, 1999a). First, because the treatment occurs in a natural conversation context, intervention can build skills and confidence in a variety of naturally occurring communication genres. Second, clinicians may introduce conversational topics, but the topic may spontaneously shift; there is a broader array of speech acts, and there is a significant impact of group dynamics (DeDe et al., 2019). In addition, people with aphasia may benefit from the diversity of language models supplied by conversation partners and spontaneity of discourse (Elman, 2007). The group dynamic is thought to foster a sense of joint purpose, which may lead to increased confidence to build interpersonal connections and engage in social activities. Conversation treatment is associated with improved communication skills and quality of life for people with aphasia (e.g., DeDe et al., 2019; Elman & Bernstein-Ellis, 1999a, 1999b; Hoover et al., 2015). A key goal of conversation treatment is to improve functional communication abilities, including discourse-level language abilities.

## DISCOURSE

Discourse is how humans are most likely to use language, and it is typically the ultimate goal of aphasia treatment. Labeling items (e.g., as in a naming task) or producing disconnected sentences may occur in some contexts, but typically language is used in interconnected, multisentence contexts (i.e., discourse). Thus, even treatments that target lexical retrieval or sentence production are intended to generalize to discourse-level tasks, even if that is not always measured (e.g., Ballard & Thompson, 1999; Boyle & Coelho, 1995). Discourse-level language processing is critical to production and comprehension of personal narratives and stories, which are fundamental to the formation and maintenance of social connections and personal identity (e.g., Shadden & Argon, 2004; Shadden & Hagstrom, 2007).

There are many different types of discourse. Common types of monologic discourse include procedural narratives (e.g., how to play a game), exposition (e.g., explaining the history of Puerto Rico), storytelling (e.g., personal narratives), and persuasive (e.g., trying to convince someone to see your perspective). Dialogic discourse involves two or more participants taking turns in an alternating pattern (Cherney et al., 1998). These interactive conversations can refer to a variety of communication situations including more structured interviews and ordering food at a restaurant to less structured workplace and home discussions (Horton, 2017). In most conversations, the communicative actions of the participants are tightly coupled; one person's contribution is shaped by what has been said previously and will also influence the subsequent turn (Clark, 1996; Garrod, 1999; Horton, 2017). Conversations may also incorporate narratives if one individual shares information (e.g., a personal narrative) with their conversation partner.

Most aphasia test batteries include a discourse task, which is typically a picture description (e.g., cookie theft picture in the Boston Diagnostic Aphasia Exam [BDAE], Goodglass et al., 2001; picnic scene in the Western Aphasia Battery [WAB], Kertesz, 2007; childcare scene in the Comprehensive Aphasia Test [CAT], Swinburn et al., 2005). In addition, the WAB and BDAE include conversational questions, such as "Have you been here before?" or "How long do you expect to be here?" Research studies also examine discourse with tasks such as picture description, story retell (e.g., Cinderella story), and procedural narratives (e.g., making a sandwich).

Despite the importance of discourse, standardized expressive language assessments for aphasia tend to focus on discrete linguistic tasks such as naming and repetition. Similarly, many treatment approaches focus on discrete language tasks such as naming and sentence production (e.g., semantic feature analysis, Boyle & Coelho, 1995; Verb Network Strengthening Treatment, Edmonds et al., 2009). However, discourse-level processing is much broader and more demanding than these discrete tasks. Discourse incorporates phonology, morphology, and syntax, as well as pragmatics, multimodal communication, and cognitive concepts such as theory of mind, working memory, organization, and so on. Thus, evaluation of only discrete language tasks will not necessarily reveal how well a person with aphasia is able to produce language in real-life contexts.

Indeed, there is evidence that discourse abilities cannot be inferred through performance on discrete language tasks such as word naming. For example, Fergadiotis et al. (2019) examined the relationship between confrontation naming and informativeness in discourse, measured as correct information units (CIUs). Latent structure analysis showed that naming accounted for approximately 63% of the variance in informativeness. However, that number dropped to 25% if there was only one measure of naming. Furthermore, naming does not incorporate other aspects of discourse such as coherence, syntax, and completeness. Similarly, studies have shown that language testing is not always reflective of conversation skills in people with aphasia (e.g., Beeke et al., 2007; Myrberg et al., 2018). On the contrary, there is evidence that CIUs obtained from monologic discourse positively correlate with the same measure in conversational discourse (Doyle et al., 1995). Thus, it is important that discourse abilities be directly evaluated in people with aphasia, rather than relying on discrete language tasks.

One challenge is that collecting and analyzing discourse samples is not a trivial task (e.g., Cruice et al., 2020). Many discourse measures require extensive training to ensure adequate interrater reliability. Furthermore, there may be variability between sessions, as a function of elicitation task, intraindividual variability, and so on (e.g., Boyle, 2014). It may be that the most ecologically valid measures, such as a conversation analysis (e.g., Beeke et al., 2007; Tetnowski et al., 2020), are also the most time-consuming with respect to learning and implementing the transcription and coding systems. Such measures may not be well suited for identifying changes following specific treatments or interventions, particularly in clinical settings (e.g., Cruice et al., 2020).

Another challenge is choosing measures for specific clients (cf. Boyle, 2020). Boyle (2020) posed a series of questions that clinicians might consider when choosing a discourse measure. Two of her questions are particularly relevant to the present article. The first is whether the discourse measure has been used with people with aphasia who are similar to the target client or clients. The second is whether the discourse measure is sensitive to the level of discourse being targeted in treatment. Here, we focus on how those questions relate to aphasia severity. People with different aphasia severities will have different goals. For example, someone with relatively mild aphasia may want to increase efficiency of communication whereas someone with relatively severe aphasia may want to increase use of multimodal communication, number of utterances, or number of conversational turns. No one measure is likely to be sensitive to all of these discourse functions. From a research perspective, this can be a particular challenge for group treatment studies. From a clinical perspective, it means that clinicians must choose outcome measures that reflect their participant's goals.

In this article, we examine four discourse measures and their potential sensitivity across a range of aphasia severities. We also explored whether these discourse measures were sensitive to effects of treatment using data selected from a randomized controlled trial of conversation treatment (DeDe et al., 2019).

## DISCOURSE MEASURES

There are a large number of possible discourse measures (Bryant et al., 2016; Dietz & Boyle, 2018; Kintz & Wright, 2018; Marini et al., 2011; Pritchard et al., 2017). These measures represent both structuralist and functionalist views of discourse (Marini et al., 2011). Structuralist discourse measures are more focused on microstructural linguistic variables (e.g., focus on grammaticality in Quantitative Production Analysis, Rochon et al., 2000; Saffran et al., 1989), as well as lexical diversity, syntactic diversity, cohesion, and so on. This article focuses more on functionalist measures, which are focused on how much and what information is conveyed. We explore four such measures: core lexicon (Dalton & Richardson, 2015; Kim et al., 2019), words per minute (WPM), CIUs (Nicholas & Brookshire, 1993), and number/type of utterances/conversational turns. We recognize that this is not an exhaustive list of available measures. We focused on these measures because they reflect informativeness or efficiency of discourse, which is more aligned with the goals of conversation treatment than microlinguistic measures.

### Core lexicon

Core lexicon measures the completeness of a story using word retrieval as a proxy. This type of analysis requires access to a list of "core lexical items," which are defined as lexical items that were produced by at least 50% of non-brain-damaged controls producing the same narrative (Dalton & Richardson, 2015). Thus, core lexicon is limited to standard stimuli such as picture descriptions and story retells.

Core lexicon measurement involves listening to the language sample and checking off each word on the list that the speaker produces. This approach provides information about whether people with aphasia retrieve lexical items that are typically accessed when describing a given stimulus. Core lexicon lists are available for several of the stimuli collected as part of the AphasiaBank protocol (MacWhinney et al., 2011), including the Cat Rescue scene (Nicholas & Brookshire, 1995) and the Cinderella Story (MacWhinney et al., 2010; Tanaka et al., 2016). Dalton et al. (2020) recently published a compendium of core lexicon lists. Core lexicon lists are not available for analysis of pictures associated with general aphasia batteries such as the WAB or BDAE. A benefit of core lexicon analyses is that they are relatively straightforward and require minimal training for stimuli that have a core lexicon list. Furthermore, coding is relatively time-efficient, particularly if the sample is already transcribed.

There are very limited psychometric data regarding core lexicon, and we are unaware of any treatment studies that have used it as an outcome measure. One study showed that core lexicon had adequate interrater reliability, but the tasks they used (*Good Dog, Carl* & *Picnic*) are seldom used with people who have aphasia in clinical practice (Kim

& Wright, 2020). It is unclear whether other core lexicon checklists share the same interrater reliability. That being said, a core lexicon tally based on written transcripts is quite straightforward, and there are similar procedures for all lists. Thus, it is possible that the reliability data can be extended to other core lexicon lists.

Core lexicon is not the only measure that reflects inclusion of main points of a story. Main concept analysis (MCA) is closely related to core lexicon (Dalton & Richardson, 2019; Hameister & Nickels, 2018; Nicholas & Brookshire, 1995; Richardson & Dalton, 2016). Instead of lexical items, MCA includes a list of utterances that reflect key events within a narrative (e.g., the slipper fits, the cat was in the tree). Similar to core lexicon, MCA requires that a list of main concepts is available for a given stimulus. Such lists have been generated for many of the AphasiaBank stimuli, including the Cat Rescue scene (Dalton & Richardson, 2019; Hameister & Nickels, 2018) and the Cinderella story (Richardson & Dalton, 2016). Core lexicon is highly correlated with MCAs (Dalton & Richardson, 2015), which establishes construct validity of this measure of core lexicon. We judge core lexicon to be more clinically feasible than MCA. It is comparatively more straightforward to code the presence or absence of specific lemmas than to judge whether elements of an utterance are accurate/inaccurate and complete/incomplete in a narrative.

Another reason for our focus on core lexicon is that it is sensitive to aphasia type and severity. In one study, core lexicon scores differed significantly for people with different aphasia syndromes (e.g., Anomic vs. Broca) (Dalton & Richardson, 2015). Core lexicon is also sensitive to very mild aphasia. People with latent aphasia, who did not meet the WAB criterion for aphasia, produced significantly fewer words on the Cinderella core lexicon list than non-brain-damaged controls and significantly more words than people with anomic aphasia (Fromm et al., 2017).

**Words per minute**

This measure reflects efficiency of discourse production. Words per minute, also known as speech rate, include the time taken to produce all content, including repetitions, revisions, fillers, and silent pauses. There is evidence that unimpaired speakers produce more and longer silent pauses before less frequent or semantically unpredictable words (e.g., Goldman-Eisler, 1958; 1968). This indicates that silent pauses may reflect language planning, meaning that a slow speech rate may reflect difficulty in lexicosemantic or syntactic planning (Peach & Coelho, 2016). Thus, this measure may be sensitive to relatively subtle language impairments in people with very mild aphasia. Fromm et al. (2017) and DeDe and Salis (2020) investigated timing measures in people with very mild aphasia, anomic aphasia, and non-brain-damaged adults. Fromm et al. (2017) reported that speech rate was significantly slower in people with very mild aphasia than in people with anomic aphasia. DeDe and Salis (2020) included a wider range of speech timing measures and showed that articulation rate, and not speech rate, differentiated people with very mild aphasia from those with other aphasia types. Critically, both studies reported that speech rate differentiates people with very mild aphasia (latent, or not-aphasic-by-WAB) from unimpaired individuals.

Words per minute are relatively reliable across time. Nicholas and Brookshire (1993) calculated WPM for 20 non-brain-damaged controls and 20 people with aphasia across 10 elicitation stimuli. Average cross-session correlations were .91 for controls and .98 for people with aphasia. Boyle (2014) analyzed data from five discourse samples produced by 12 individuals with anomic aphasia at three time points. Correlations for WPM measures between the three testing sessions were .99. She also calculated the minimal detectable change (MDC; 90% confidence interval), which is the smallest change in an individual's data that can be interpreted as a real change rather than measurement error, with 90% confidence. Based on her data, a change of nine WPM is unlikely to reflect measurement error. Her data suggest that WPM can be used both in group research studies and in clinical decision-making for individual clients.

It is unclear whether WPM can be used to distinguish between mild and moderate aphasia or between moderate and severe aphasia. However, WPM may distinguish people with very mild aphasia from neurotypical speakers. We do not know of any existing studies that provide definitive cutoffs. In a Cinderella retell task, Fromm et al.'s (2017) group of 27 individuals with latent aphasia produced an average of 86.9 WPM (*SD* = 24.6) compared with 134.3 WPM for their 177 controls (*SD* = 30.9). DeDe and Salis (2020), who also examined Cinderella stories, reported that their latent aphasia group (*n* = 10) produced 109.9 WPM (*SD* = 23). Their control participants produced 164.9 WPM (*SD* = 26.9). Finally, Nicholas and Brookshire (1993) tested 20 non-brain-damaged controls on several picture description tasks and found the participants produced an average of 166 WPM (*SD* = 22). Based on these data, a conservative benchmark is that fewer than 100 WPM are indicative of aphasia. However, this rule of thumb should be applied cautiously, given the variability described earlier and given that WPM could differ as a function of discourse type (personal narrative vs. story retell vs. picture description, for example).

### Correct information units

Bryant et al. (2016) suggested that CIU analysis was the most used analysis for the assessment of linguistic discourse in aphasia. Correct information units (Nicholas & Brookshire, 1993) are words that add new information to the topic and are intelligible, accurate, and relevant in context. Detailed rules and examples are provided for which words to include and exclude in the count. When counting CIUs, all CIUs are words (from the WPM count) but not all words are CIUs. Correct information units add relevant, new information or are syntactically required. Conjunctions (e.g., *and*) are never counted as CIUs, and abandoned utterances are not counted. Nonspecific words such as "thing" or "there" are only counted if they are grammatically obligatory. For example, *there* would be counted in a sentence such as "There is the cat" but not in "The mother is there." Several measures can be derived from CIUs, such as the overall number of CIUs, considered an indicator of informativeness, and the percentage of CIUs and the number of CIUs per minute, considered to be indicators of efficiency.

According to Pritchard et al. (2017), CIUs have some of the strongest psychometric data of any discourse measures. Correct information units can show good interrater reliability for well-trained coders (Brookshire & Nicholas, 1994; Leaman & Edmonds, 2019; Oelschlaeger & Thorne, 1999). Fergadiotis and Wright (2016) reported interrater reliability for CIUs (Nicholas & Brookshire, 1993) on 11 of the 98 discourse samples that they analyzed. They reported that their highly trained raters differed by less than 6% for 10 of the samples and by 8% for the 11th sample. A greater challenge is that individual performance shows a high degree of variability across days and tasks (Brookshire & Nicholas, 1994), though this can be mitigated by collapsing CIU data from at least five elicitation tasks (Boyle, 2014). Correct information units can also vary as a function of task. Significant differences between structured and conversational tasks have been observed for both the total number of words and the total number of CIUs, though CIUs were significantly correlated across tasks (e.g., Doyle et al., 1995). In another study, participants with fluent aphasia performed better than those with nonfluent aphasia during a conversational task with respect to the total number of words and the number of CIUs, but not on percent CIUs (McCullough et al., 2017). However, CIUs for participants with fluent and nonfluent aphasia did not differ on a picture description task. Thus, CIU analyses are sensitive to variables such as aphasia type and narrative task.

Many studies have reported CIU data, but there are no established norms for normal and impaired performance. Nicholas and Brookshire (1993) reported that people with aphasia ($n$ = 20) produced an average of 49 to 52 CIUs per minute across three time points, with *SD*s ranging from 24 to 27. Boyle (2014) reported CIU data for 12 people with aphasia at three time points. Boyle's (2014) participants with aphasia produced averages of 28 to 32 CIUs per minute. Her analyses suggested that CIUs per minute were a more stable measure than percent CIUs. She also reported that the MDC score (90% confidence interval) for CIUs per minute was a change of 12 CIUs per minute. Several studies have also reported CIU data for non-brain-damaged adults. Nicholas and Brookshire (1993) tested non-brain-damaged controls ($n$ = 20) at three separate time points. Across time points, average CIUs per minute ranged from 143 to 147. Capilouto et al. (2005) examined CIUs in 17 younger (mean age = 22.4 years) and 17 older (mean age = 71.4 years) adults on four narrative speech tasks. Across all four tasks, older adults produced an average of 120 CIUs per minute (*SD* = 19) and younger adults produced an average of 133 CIUs per minute (*SD* = 29). On the basis of these data, production of fewer than 100 CIUs per minute is likely to reflect some degree of language impairment. It is important to keep in mind, however, that this is a conservative estimate, which may over- or underestimate the existence of such disorders.

## Conversational turns

The final measure we consider is the number of conversational turns. Turn-taking refers to a type of organization in conversation and discourse in which participants speak one at a time in alternation. Following a given turn, the speaker has the option to identify the next speaker (speaker selection) or another participant of the conversation can begin speaking at the completion of the turn (self-selection). Self-selection is subject to competition, as multiple participants in the conversation may want to speak at once (e.g., Bernstein-Ellis & Elman, 2007; Simmons-Mackie et al., 2007). Thus, the number of turns taken during a conversation at least partly reflects confidence in communication and processing speed, as a long pause may result in another participant taking the conversational turn.

There are different types of conversational turns (e.g., Herbert et al., 2014). A broad distinction can be made between substantive, maintenance, and minimal turns. Substantive turns add new information to the conversation with at least one content word, including repetition of another speaker and recognizable paraphasic errors. Maintenance turns do not add new information but keep the conversation moving. These may include passing the turn to a specific speaker (e.g., *Liz, did you want to say something?*), repetition of a previous turn, or greetings/closings. Minimal turns do not contribute new content and only serve to return the conversation to the original speaker. Examples of minimal turns include "*mhm, okay, yeah*."

Seminal work on turn-taking in conversational dyads shows that people with aphasia retain turn-taking skills (e.g., Holland, 1982; Schienberg & Holland, 1980) and are also able to formulate requests using verbal and nonverbal communication (e.g., Prinz, 1980) and use repair strategies to address conversational breakdowns (e.g., Linebaugh et al., 1985; Newhoff et al., 1982). Previous studies have found acceptable interrater reliability for ratings of conversational turns (e.g., McCarney & Johnson, 2001; Ramsberger & Rende, 2002), though test–retest reliability has not been examined. One study used the number of conversational turns as an outcome measure following a treatment of anomia (Best et al., 2011). They did not find changes in the number of conversational turns in their group of 11 people with aphasia, though some individuals within the group did show increased turn-taking. Nonetheless, we include conversational turns because this measure may be more relevant to the dialogic nature of conversation treatment and other interventions. Thus, counting the number and types of conversational turns may provide valuable information about communicative participation and engagement. This measure is analogous to counting the number of utterances in a monologic narrative.

## PURPOSE OF THE ARTICLE

We recently completed a study investigating the effects of group size in conversation group treatment in aphasia (DeDe et al., 2019). We hypothesized that discourse production should change as a result of a discourse-based intervention and selected percent CIUs (%CIU; Nicholas & Brookshire, 1993) as the primary outcome measure. Although %CIU did not change for any group, both treatment groups showed significant improvements on several standardized tests of language. In contrast, the control group did not show any significant changes from pre- to posttreatment. This finding prompted discussion about the potentially complex relationship between profile of aphasia, conversation goals, and the best measures to capture treatment effects in a diverse cohort of people with aphasia. To this end, we selected pre- and posttreatment discourse samples from two cases from this larger group to explore whether several different discourse measures (1) were sensitive to the presence of aphasia and (2) reflected behavioral changes for different aphasia profiles following conversation treatment. The data presented here are secondary analyses of group data in DeDe et al. (2019). We selected discourse samples from two participants from the larger study, one who presented with mild, fluent aphasia and the other with severe, nonfluent aphasia.

## CASE STUDIES

Both case studies participated in a larger study of conversation treatment (DeDe et al., 2019). Participants in the larger study were randomly assigned to one of two treatment conditions (large group or dyad) or a delay-treatment control group. Both participants reported here were assigned to the large group condition, in which they received 10 weeks of conversation treatment, twice per week, with 60 min per session. For full details, see DeDe et al. (2019).

Trained graduate students administered the treatment under supervision of a licensed speech–language pathologist. Treatment sessions were structured around a predetermined set of topics in order to maximize consistency across groups. However, conversations were allowed to flow naturally after topics were introduced. During all sessions, multimodal supports were available for participants, including PowerPoint slides with relevant visuals, tablet or laptop computers, paper/whiteboards, and paper printouts of communication supports. Sample topics included family history, favorite restaurants, and current events. Clinicians modeled the use of multimodal supports and supplemented speech with key words. The overall goal of the treatment was for participants to successfully communicate their thoughts and ideas in a conversational setting. In addition, all participants had individualized goals, which were determined on the basis of a comprehensive diagnostic evaluation and participant concerns.

All participants completed a comprehensive battery of tests. Here, we report a subset of standardized tests as well as discourse analyses of connected language samples (monologic narrative in response to a picture). For the picture description task, no external supports (e.g., paper or tablet computer) were provided. In addition, the severely impaired participant engaged in conversations with a novel partner. This task is described in greater detail later. Data are reported from immediately before and after treatment. Trained graduate students collected pre- and posttreatment data under the supervision of a licensed speech–language pathologist.

## Mild aphasia: A.B.

### *Background and standardized testing*

A.B. was a 60-year-old African American woman with anomic aphasia secondary to a stroke 5 years before testing. When asked for her perspective on her aphasia, A.B. said, "It's gotten a lot better. I can speak clearly on some things. Other things, I have to work at it. But I'm speaking better." From a clinician's perspective, A.B. is very communicative with familiar conversation partners in a supportive environment, such as an aphasia support group. However, A.B. reports that when she is in less supportive environments, such as social gatherings of friends and family, she has great difficulty participating in conversations. She reports that she does not initiate conversation outside of the aphasia support group.

Based on pretreatment standardized testing (see Table 1), A.B. has mild aphasia. On the short form of the Philadelphia Naming Test (Walker & Schwartz, 2012), she correctly named 28 of 30 items. On the Northwestern Assessment of Verbs and Sentences (NAVS; Cho-Reyes & Thompson, 2012) Verb Naming Test, she correctly named 20 of 22 verbs. She earned 30/30 points on the NAVS Sentence Production Priming Test, indicating that she is able to correctly produce simple active and complex object relative sentences. A.B. completed several sections of the CAT (Swinburn et al., 2005). The auditory comprehension section of the CAT includes word picture matching, sentence picture matching with simple and complex sentences, and comprehension of verbally presented paragraphs. A.B. earned a T-score of 52 on the auditory comprehension section, which is below the normal cutoff score of 56. The naming section of the CAT includes verbal fluency, object confrontation naming, and a small number of action confrontation naming items. A.B. earned a T-score of 61, just below the cutoff of 62 for that section. Stimuli in the repetition section of the CAT include words, nonwords, sentences, and number sequences. On this section, A.B. earned a T-score of 66, which is above the normal cutoff of 59. On the basis of standardized testing, we concluded that A.B. presented with mild deficits in word retrieval and auditory comprehension.

A.B. also completed the adaptive version of the Aphasia Communication Outcome Measure (ACOM; Hula et al., 2015). The Adaptive ACOM is a 12-item version of the test in which an algorithm adaptively selects 12 items from the test bank based on

**Table 1.** Standardized test scores pre- and posttreatment

| | Case 1: A.B. | | Case 2: C.D. | |
|---|---|---|---|---|
| **Measure** | **Pre-Tx** | **Post-Tx** | **Pre-Tx** | **Post-Tx** |
| Philadelphia Naming Test | 28 | 26 | 5 | 6 |
| Northwestern Assessment of Verbs and Sentences | | | | |
| Verb Naming | 20 | 22 | 3 | 6 |
| Sentence Production Priming | 30 | 28 | 0 | 1 |
| Comprehensive Aphasia Test | | | | |
| Auditory comprehension | 52 | 59 | 34 | 43 |
| Written comprehension | 50 | 57 | 35 | 41 |
| Repetition | 59 | 66 | 44 | 43 |
| Naming section | 61 | 65 | 43 | 45 |
| Aphasia communication outcome measure | 53.01 (2.7) | 52.11 (2.6) | 44.4 (2.77) | 45.06 (2.47) |

*Note*. Tx = treatment.

the participant's responses. On the ACOM, participants are presented with communication tasks and must indicate how effectively they complete each task on a 4-point scale ("not very," "somewhat," "mostly," and "completely"). Participants can also indicate that they do not have the opportunity to complete certain tasks. A.B.'s estimated T-score was 53.01, which is above the average of 50 (for people with aphasia). On the ACOM, A.B. reported that she (1) follows simple spoken requests "completely" effectively, (2) requests information from store employees "mostly" effectively, (3) tells a story "somewhat" effectively, and (4) has conversations with strangers "somewhat" effectively. These ratings were largely in line with her own description of her impairment, which related to speaking in more challenging environments. These results did not reflect the difficulty with auditory comprehension revealed by the standardized testing, possibly because the comprehension tasks sampled by the ACOM (e.g., following simple commands) were not particularly demanding.

On almost every measure, A.B. would be classified as having mild to no aphasia. From her own perspective, however, aphasia continued to significantly limit her ability to participate in life activities. In the next section, we explore how different discourse measures capture her language impairments, including any changes following conversation treatment.

### *Discourse measures*

As part of her evaluation, A.B. also produced several connected language samples for discourse analysis. Here, we consider three different methods of analyzing her description of Nicholas and Brookshire's (1993) Cat Rescue scene: core lexicon, WPM, and CIUs. Her pre- and posttreatment samples are presented in Appendix A. Results are summarized in Table 2.

A.B.'s description of the Cat Rescue scene was analyzed using the core lexicon list developed by Tanaka et al. (2016). This list contains 34 words such as *bark*, *dog*, *cat*, and *ladder.* Pretreatment, A.B. mentioned

**Table 2.** Discourse measures pre- and posttreatment: Cat Rescue scene and conversation with a naive partner

| | Case 1: A.B. | | Case 2: C.D. | |
|---|---|---|---|---|
| | **Pre-Tx** | **Post-Tx** | **Pre-Tx** | **Post-Tx** |
| Cat Rescue scene | | | | |
| Core lexicon (max = 34) | 22 | 25 | 6 | 6 |
| Number of words | 74 | 153 | 25 | 37 |
| Duration of sample (s) | 56 | 103 | 137 | 70 |
| WPM | 79.6 | 89.0 | 10.96 | 31.62 |
| CIUs | 70 | 125 | 3 | 4 |
| CIUs per minute | 73.7 | 72.7 | 2 | 3.43 |
| % CIUs | 95 | 82 | 12 | 11 |
| Personal narrative | | | | |
| Duration (s) | NA | NA | 345 | 675 |
| Words | NA | NA | 64 | 160 |
| WPM | NA | NA | 11.1 | 14.22 |
| # Conversational turns | NA | NA | 3 | 2 |
| Types of utterance within turn | | | | |
| Substantive | NA | NA | 11/18 | 31/73 |
| Maintenance | NA | NA | 1/18 | 40/73 |

*Note*. CIU = correct information unit; NA = data not available; Tx = treatment; WPM = words per minute.

22 of the 34 items, or 64.7%. In contrast, non-brain-damaged individuals ($n$ = 92) produced an average of 26 items ($SD$ = 3.39) (Tanaka et al., 2016). Thus, A.B. falls 1 *SD* below the normative sample's mean. Posttreatment, A.B. mentioned 25 of the 34 items, or 73.5%, which is within 1 *SD* of the non-brain-damaged individuals. Thus, there is some evidence of change using the core lexicon method for this picture description task, though it is unclear whether this would be considered a significant change.

We next calculated how many WPM A.B. produced using the rules established by Nicholas and Brookshire (1993). This method excludes filled pauses (*uh*, *um*, etc.) and comments that introduce, conclude, or comment on the task. Pretreatment, A.B. produced 74 words in 56 s, or 79.6 WPM, which is below the 100 WPM cutoff suggested earlier for aphasia. Posttreatment, A.B. produced 153 words in 103 s, for a total of 89.0 WPM. This change of 9.4 WPM exceeds the MDC score (90% confidence interval) proposed by Boyle (2014).

The next analysis used Nicholas and Brookshire's (1993) method for calculating CIUs. Pretreatment, A.B. produced 70 CIUs. From an efficiency perspective, A.B. produced 95% CIUs and 73.7 CIUs per minute. Her percent CIU score does not seem to reflect impaired production, but she produced fewer than 100 CIUs per minute, which was the cutoff proposed earlier for typical language production. Posttreatment, she produced 153 words and 125 CIUs, which correspond to 81.7% CIUs and 72.7 CIUs per minute, respectively. The apparent decline in efficiency, as measured by percent CIUs and CIUs per minute, falls within the range that Boyle (2014) identified as reflecting intraindividual variability across trials (cf. Boyle, 2014). Thus, this change is unlikely to reflect a true decline in performance.

### *A.B. summary*

A.B.'s case report demonstrates how standardized tests, especially those normed on people with aphasia, may underestimate mild aphasia. Critically, A.B.'s standardized tests identified auditory comprehension as her greatest area of impairment whereas she reported spoken language production as the greatest concern. Discourse analyses showed that core lexicon, WPM, and CIUs per minute were sensitive to her impairments, placing her at least 1 *SD* below the mean for non-brain-damaged controls. Thus, these types of discourse analyses may better reflect the effects of mild aphasia than standardized tests. Of these, core lexicon and WPM showed sensitivity to effects of conversation treatment.

## Case 2: Severe aphasia

### *Background and standardized testing*

The second case is C.D., a 73-year-old White man with severe Broca's aphasia secondary to a stroke 16 years before testing. When asked for his perspective on his aphasia, C.D. said, "better ... but ah ... not great—oh boy." C.D. engages in many hobbies with his family; he enjoys traveling, politics, and boating. His verbal expression is severely impaired. His speech is nonfluent, marked by anomia and agrammatism. Utterances are often limited to short social phrases, key nouns, and isolated verbs. Comprehension is good for contextually supported sentences. He reports difficulty participating in conversations due to challenges understanding in large groups and effectively sharing thoughts and ideas.

Based on pretreatment standardized testing (see Table 1), C.D. has a severe aphasia profile. On the short form of the Philadelphia Naming Test (Walker & Schwartz, 2012), he correctly named five of 30 items. On the NAVS (Cho-Reyes & Thompson, 2012) Verb Naming Test, he correctly named 3 of 22 verbs. He did not produce any complete sentences on the NAVS Sentence Production Priming Test, earning 0 of 30 points. On the Comprehensive Aphasia Test (Swinburn et al., 2005), he earned a T-score of 34 on the auditory comprehension section based on a strength understanding 11 canonical sentences. On the naming section of the

CAT, he earned a T-score of 35; he named two animals given a categorical prompt, one word given an alphabetical prompt ("s"), and one pictured noun. On the repetition section, C.D. earned a T-score of 44, correctly repeating nine simple real words, two nonwords, and a three-digit string. He did not accurately repeat any complex words or sentences. Overall, C.D. presents with significant deficits in all domains, with relative strengths in comprehension of canonical sentences and repetition of simple words. On the adaptive version ACOM (Hula et al., 2015), C.D.'s estimated T-score was 44.40, which is below the average of 50 for people with aphasia. He reported that he could effectively read signs and ask for help in stores but could not effectively tell stories or communicate with strangers. Within conversation treatment, his individual goals focused on initiating conversation turns and communicative effectiveness using multimodal communication of key words/phrases.

### *Discourse measures*

C.D. produced several connected language samples for discourse analysis. Consistent with the previous case, we report three analyses of his description of the Cat Rescue picture. Transcripts of his pre- and posttreatment samples are presented in Appendix A, and results are summarized in Table 2. In addition, we collected dialogic conversation samples before and after treatment. Each conversation was conducted with a novel conversation partner who had no prior training or experience with aphasia. The partner was instructed to meet and converse with C.D. for about 5 min and to start by asking about his weekend. No other instructions were provided. After the initial prompt, the conversation was intended to proceed naturally. However, C.D. did not ask any questions and the naive listener only produced minimal turns such as "*mmhmm*" or "yeah." Thus, the conversations unfolded as personal narratives about C.D.'s weekend. For this reason, these conversations are referred to as personal narratives later. Critically, use of multimodal communication was allowed during this personal narrative but not during the picture description task. A tablet computer and pen and paper were provided to the participant. The samples are transcribed in Appendix A.

C.D.'s description of the Cat Rescue picture was analyzed using the core lexicon list developed by Tanaka et al. (2016). Pretreatment and posttreatment, C.D. mentioned six of the 34 items, or 17.6%, well below the mean of 26 ($SD = 3.39$) for healthy controls.

Using Nicholas and Brookshire's (1993) rules, we tabulated the total number of words produced during the sample. Pretreatment, C.D. produced 25 words in 137 seconds, or 10.96 WPM, which is significantly below the 100 WPM cutoff suggested earlier for aphasia. Posttreatment, C.D. produced 37 words in 70 seconds, or 31.62 WPM.

Next, we computed CIUs based on Nicholas and Brookshire's (1993) method. Pretreatment, C.D. produced a total of three CIUs, 12% CIUs (three CIUs, 25 total words), and two CIUs per minute, reflecting a sparse narrative with little informative value. Posttreatment, he produced four CIUs, which corresponded to 11% CIUs (four CIUs, 36 words), and 3.43 CIUs per minute.

The personal narratives were analyzed using the methods described by Ulatowska et al. (1992). We used this method because the sample was derived in a conversational setting. However, as noted previously, these conversations more closely resembled personal narratives. Each sample was orthographically transcribed and parsed into utterances using c-units (Miller & Iglesias, 2008). All words within each utterance were counted with the exception of repetitions (without emphasis) and false starts. Utterances were then coded as follows: (1) substantive utterances, which add new information to the dialogue; and (2) maintenance utterances, which added no new information but served to maintain the conversation or hold the conversational turn. Utterances were (3) "not codable" when they were unintelligible in context, meaning that they could not reliably be judged as substantive or maintenance utterances.

The pretreatment personal narrative lasted 345 s, and the posttreatment personal narrative lasted 675 s. The number of words produced in each utterance was counted and summed for each conversation. C.D. produced 64 words pretreatment and 160 words posttreatment. To control for duration, WPM were calculated revealing a change from 11.1 WPM pretreatment to 14.22 WPM posttreatment. In terms of type of conversational utterances, pretreatment C.D. produced 11 substantive utterances (61%), one maintenance utterance (6%), and six not codable utterances (33%). Posttreatment, he produced 31 substantive utterances (42%), 40 maintenance utterances (55%), and two not codable utterances (3%). One important observation was his use of multimodal conversational turns in the posttreatment conversation. C.D. used four distinct multimodal strategies to support his conversational turns, such as locating his hometown on a map on his tablet, drawing the pond near his home where he took guests, writing numbers, and circling groups of numbers to indicate numbers of guests. These multimodal substantive turns were critical to his success in conveying his message.

### *C.D. summary*

Narrative analyses of a picture description task showed that core lexicon and CIUs per minute were sensitive to his impairments, but they did not show change following treatment. In contrast, WPM in monologic narratives were both sensitive to his impairments and showed change following treatment. The data from the personal narrative provided meaningful information about C.D.'s overall communicative effectiveness, as evidenced by the increase in the number of substantive turns. Furthermore, he showed greater perseverance during the narrative, as demonstrated by the increase in maintenance turns, which were used to indicate that his story was continuing. That is, instead of abandoning his turn or cutting his narrative short, he continued to speak. Thus, the use of WPM and examination of a personal narrative provided valuable information in this case profile of severe aphasia.

## CONCLUSIONS

This article had two main purposes. The first was to determine which, if any, of the discourse measures we examined were sensitive to both mild and severe aphasia. The second goal was to explore the effects of conversation treatment on commonly used discourse measures of informativeness and efficiency in people with varying degrees of severity. We were initially interested in these questions because our randomized controlled trial of conversation treatment included the full range of aphasia severities (DeDe et al., 2019). These questions are also clinically important because speech–language pathologists see a wide range of aphasia severities in practice.

Perhaps, not surprisingly, all of the discourse measures that we explored were sensitive to the presence of severe aphasia. Standardized tests, which are often less time-consuming than discourse measures, were also sensitive to the presence of severe aphasia. In contrast, only core lexicon, WPM, and CIUs per minute were sensitive to our case example of mild aphasia. Some of the standardized measures were sensitive to mild aphasia, but they did not capture the individual's self-reported areas of difficulty. Across the spectrum of aphasia severity, discourse measures can provide important insights into the individual's communication strengths and weaknesses and point to appropriate treatment targets. However, to document the presence of aphasia (e.g., for third party payers in the American health care system), discourse analyses are probably more necessary in cases of mild aphasia than severe aphasia.

With respect to the second question, our case studies suggest that WPM may be sensitive to treatment changes in both mild and severe aphasia. In A.B.'s case, one measure of informativeness—core lexicon—also increased following treatment. However, other measures of efficiency (e.g., CIUs per minute or percent CIUs) did not change. For the individual with severe aphasia (C.D.), analysis of a personal narrative showed an increase in the number and type of utterances following treatment. The lack of change in measures of efficiency may reflect the nature of our conversation treatment, which focused more on expressing one's thoughts than on doing so within a set time frame. It is thus interesting that the overall length of the samples also increased from pre- to posttreatment.

The change in duration of our samples reflects a potential limitation of this project. We did not attempt to match samples for length, and both case studies produced longer samples posttreatment than pretreatment. As a result, it is difficult to determine whether increases in core lexicon scores (A.B.) or number of substantive utterances (C.D.) reflect an improved ability to produce content or just more time spent talking. We did not match samples for length in part because we were interested in whether people with aphasia would talk more after treatment. Many of our clients with aphasia—including mildly impaired individuals such as A.B.—report that they talk less than they used to even with familiar conversation partners. Thus, we viewed an increase in output as potentially meaningful, even as it introduced duration as a potential confound. We also recognize that the data presented here do not constitute evidence of the efficacy of conversation treatment, nor do these results necessarily generalize to all individuals with mild and severe aphasia. Instead, the purpose of this article is to illustrate how the discourse measures that we used can be applied to individual case studies.

Choosing one discourse measure for multiple aphasia profiles is a challenge. Choosing one discourse measure to detect meaningful change in a group study following a multifaceted intervention adds another layer of complexity. For example, individuals enrolled in conversation group treatment may have different communication goals. Communication goals might target lexical retrieval in discourse, grammatical completeness, or communicative gist, depending on the individual's aphasia profile and personal goals. For a client with a severe profile of aphasia, communication goals may focus on retrieval of key words or the use of multimodal communication for communicative effectiveness, and, conceivably, all these aspects of language may have changed as a result of a discourse-focused intervention. Possible metrics may include core lexicon or CIUs to capture lexical changes, and CIUs or conversation analysis to measure change in communicative effectiveness and completing all these analyses on multiple samples is often unrealistic. Therefore, it is critical to choose the most relevant, salient discourse measure from all the possible options.

One clear implication of this work is that further psychometric studies are needed to establish the reliability of discourse measures, in both controlled research settings and ecologically valid clinical settings. In addition, normative data for discourse measures are necessary to establish impaired performance and to establish MDC scores in order to identify whether variations in an individual's performance are due to chance (see Bryant et al., 2016).

Despite these challenges, discourse measures offer a unique and valuable perspective on an important aspect of communication. This in-depth look at two different cases of aphasia illustrates the importance of discourse measurement in general and the potential benefits of particular measures, given differences in the profiles and goals. Our results show that in milder cases of aphasia, discourse measures have the potential to capture relevant communication aspects, which are not seen on traditional, standardized measures of discrete areas of language, such as naming or sentence production. Indeed, WPM and core lexicon were the only measure from our larger testing battery to support A.B.'s report of significant challenges in her communicative ability. As such, this discourse measure provides data to justify the need for services and to use as a baseline for measuring progress. For C.D., analyses of monologic discourse (i.e., picture description) revealed a communication deficit across all levels of the linguistic hierarchy, but these analyses did not capture the nuanced multimodal communicative improvements evidenced in the analysis of the personal narrative.

The information detailed in this article offers support for the use of discourse measurement in aphasia measurement as it has the potential to add valuable information to communication profiles and provide a set of tools for measuring change.

## REFERENCES


Ballard, K. J., & Thompson, C. K. (1999). Treatment and generalization of complex sentence production in agrammatism. *Journal of Speech, Language, and Hearing Research*, *42*(3), 690–707. https://doi.10.1044/jslhr.4203.690

Beeke, S., Maxim, J., & Wilkinson, R. (2007). Using conversation analysis to assess and treat people with aphasia. *Seminars in Speech and Language*, *28*(2), 136–147. https://doi.10.1055/s-2007-970571

Bernstein-Ellis, E., & Elman, R. (2007). Aphasia group communication treatment: The Aphasia Center of California approach. In E. Bernstein-Ellis & R. Elman (Eds.), *Group treatment of neurogenic communication disorders: The expert clinician's approach* (pp. 71–94). Plural Publishing.

Best, W., Grassly, J., Greenwood, A., Herbert, R., Hickin, J., & Howard, D. (2011). A controlled study of changes in conversation following aphasia therapy for anomia. *Disability and Rehabilitation*, *33*(3), 229–242. https://doi.10.3109/09638288.2010.534230

Boyle, M. (2014). Test–retest stability of word retrieval in aphasic discourse. *Journal of Speech, Language, and Hearing Research*, *57*(3), 966–978. https://doi.10.1044/2014_JSLHR-L-13-0171

Boyle, M. (2020). Choosing discourse outcome measures to assess clinical change. *Seminars in Speech and Language*, *41*(1), 1–9. https://doi.10.1055/s-0039-3401029

Boyle, M., & Coelho, C. A. (1995). Application of semantic feature analysis as a treatment for aphasic dysnomia. *American Journal of Speech-Language Pathology*, *4*(4), 94–98. https://doi.org/10.1044/1058-0360.0404.94

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

Bryant, L., Ferguson, A., & Spencer, E. (2016). Linguistic analysis of discourse in aphasia: A review of the literature. *Clinical Linguistics & Phonetics*, *30*(7), 489–518. https://doi.10.3109/02699206.2016.1145740

Capilouto, G., Wright, H. H., & Wagovich, S. A. (2005). CIU and main event analyses of the structured discourse of older and younger adults. *Journal of Communication Disorders*, *38*(6), 431–444. https://doi.10.1016/j.jcomdis.2005.03.005

Cherney, L., Shadden, B., & Coelho, C. (1998). *Analyzing discourse in communicatively impaired adults—Rehabilitation Institute of Chicago*. Aspen Publications.

Cho-Reyes, S., & Thompson, C. K. (2012). Verb and sentence production in aphasia: Northwestern Assessment of Verbs and Sentences (NAVS). *Aphasiology*, *26*(10), 1250–1277. https://doi.10.1080/02687038.2012.693584

Clark, H. H. (1996) *Using language*. Cambridge University Press.

Cruice, M., Botting, N., Marshall, J., Boyle, M., Hersh, D., Pritchard, M., & Dipper, L. (2020). UK speech and language therapists' views and reported practices of discourse analysis in aphasia rehabilitation. *International Journal of Language & Communication Disorders*, *55*(3), 417–442. https://doi.10.1111/1460-6984.12528

Dalton, S. G. H., Kim, H., Richardson, J. D., & Wright, H. H. (2020). A compendium of core lexicon checklists. *Seminars in Speech and Language*, *41*(1), 45–60. https://doi.10.1055/s-0039-3400972

Dalton, S. G. H., & Richardson, J. D. (2015). Core-lexicon and main-concept production during picture-sequence description in adults without brain damage and adults with aphasia. *American Journal of Speech-Language Pathology*, *24*(4), S923–S938. https://doi.10.1044/2015_AJSLP-14-0161

Dalton, S. G. H., & Richardson, J. D. (2019). A large-scale comparison of main concept production between persons with aphasia and persons without brain injury. *American Journal of Speech-Language Pathology*, *28*(1S), 293–320. https://doi.10.1044/2018_AJSLP-17-0166

DeDe, G., Hoover, E., & Maas, E. (2019). Two to tango or the more the merrier? A randomized controlled trial of the effects of group size in aphasia conversation treatment on standardized tests. *Journal of Speech, Language, and Hearing Research*, *62*(5), 1437–1451. https://doi.10.1044/2019_JSLHR-L-18-0404

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.10.1044/2019_AJSLP-CAC48-18-0210

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

Doyle, P. J., Goda, A. J., & Spencer, K. A. (1995). The communicative informativeness and efficiency of connected discourse by adults with aphasia under structured and conversational sampling conditions. *American Journal of Speech-Language Pathology*, *4*(4), 130–134.

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.10.1080/02687030802291339

Elman, R. J. (2007). The importance of aphasia group treatment for rebuilding community and health. *Topics in Language Disorders*, *27*(4), 300–308. https://doi.10.1097/01.TLD.0000299884.31864.99

Elman, R. J., & Bernstein-Ellis, E. (1999a). The efficacy of group communication treatment in adults with chronic aphasia. *Journal of Speech, Language, and Hearing Research*, *42*(2), 411–419. https://doi.10.1044/jslhr.4202.411

Elman, R. J., & Bernstein-Ellis, E. (1999b). Psychosocial aspects of group communication treatment. Preliminary findings. *Seminars in Speech and Language*, *20*(1), 65–71; quiz 71–72. https://doi.10.1055/s-2008-1064009

Fergadiotis, G., Kapantzoglou, M., Kintz, S., & Wright, H. H. (2019). Modeling confrontation naming and discourse informativeness using structural equation modeling. *Aphasiology*, *33*(5), 544–560. https://doi.org/10.1080/02687038.2018.1482404

Fergadiotis, G., & Wright, H. H. (2016). Modelling confrontation naming and discourse performance in aphasia. *Aphasiology*, *30*(4), 364–380. https://doi.org/10.1080/02687038.2015.1067288

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.10.1044/2016_AJSLP-16-0071

Garrod, S. (1999). The challenge of dialogue for theories of language processing. In S. Garrod & M. Pickering (Eds.), *Language processing* (pp. 389–415). Psychology Press.

Goldman-Eisler, F. (1958). The predictability of words in context and the length of pauses in speech. *Language and Speech*, *1*, 226–231.

Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech*. Academic Press.

Goodglass, H., Kaplan, E., & Barresi, B. (2001). *BDAE-3: Boston Diagnostic Aphasia Examination—Third Edition*. Lippincott Williams & Wilkins.

Hameister, I., & Nickels, L. (2018). The cat in the tree—Using picture descriptions to inform our understanding of conceptualisation in aphasia. *Language, Cognition and Neuroscience*, *33*(10), 1296–1314.

Herbert, R., Gregory, E., & Best, W. (2014). Syntactic versus lexical therapy for anomia in acquired aphasia: Differential effects on narrative and conversation. *International journal of Language & Communication Disorders*, *49*(2), 162–173. https://doi.10.1111/1460-6984.12054

Holland, A. L. (1982). Observing functional communication of aphasic adults. *Journal of Speech and Hearing Disorders*, *47*(1), 50–56. https://doi.10.1044/jshd.4701.50

Horton, W. (2017). Theories and approaches to the study of conversation and interactive discourse. In M. F. Schober, D. N. Rapp, & A. M. Britt (Eds.), *The Routledge handbook of discourse processes* (2nd ed., pp. 22–68). Routledge Press.

Hoover, E. L., Caplan, D., Waters, G., & Budson, A. (2015). Effects of impairment-based individual and socially oriented group therapies on verb production in aphasia. *Aphasiology*, *29*, 781–798.

Hula, W. D., Doyle, P. J., Stone, C. A., Austermann Hula, S. N., Kellough, S., Wambaugh, J. L., Ross, K. B., Schumacher, J. G., & St Jacque, A. (2015). The Aphasia Communication Outcome Measure (ACOM): Dimensionality, item bank calibration, and initial validation. *Journal of Speech, Language, and Hearing Research*, *58*(3), 906–919. https://doi.10.1044/2015_JSLHR-L-14-0235

Kertesz, A. (2007). *WAB-R: Western Aphasia Battery–Revised*. PsychCorp.

Kim, H., Kintz, S., Zelnosky, K., & Wright, H. H. (2019). Measuring word retrieval in narrative discourse: Core lexicon in aphasia. *International Journal of Language & Communication Disorders*, *54*(1), 62–78. https://doi.10.1111/1460-6984.12432

Kim, H., & Wright, H. H. (2020). Concurrent validity and reliability of the core lexicon measure as a measure of word retrieval ability in aphasia narratives. *American Journal of Speech-Language Pathology*, *29*(1), 101–110. https://doi.10.1044/2019_AJSLP-19-0063

Kintz, S., & Wright, H. H. (2018). Discourse measurement in aphasia research. *Aphasiology*, *32*(4), 472–474.

Leaman, M. C., & Edmonds, L. A. (2019). Revisiting the correct information unit: Measuring informativeness in unstructured conversations in people with aphasia. *American Journal of Speech-Language Pathology*, *28*(3), 1099–1114. https://doi.10.1044/2019_AJSLP-18-0268

Linebaugh, C. W., Person Margulies, C., & Mackisack, E. L. (1985). Contingent queries and revisions used by aphasic individuals and their most frequent communication partners. *Clinical Aphasiology*, *15*, 229–236.

MacWhinney, B., Fromm, D., Forbes, M., & Holland, A. (2011). AphasiaBank: Methods for studying discourse. *Aphasiology*, *25*(11), 1286–1307. https://doi.10.1080/02687038.2011.589893

MacWhinney, B., Fromm, D., Holland, A., Forbes, M., & Wright, H. (2010). Automated analysis of the

Cinderella story. *Aphasiology*, *24*(6–8), 856–868. https://doi.10.1080/02687030903452632

Marini, A., Andreetta, S., Del Tin, S., & Carlomagno, S. (2011). A multi-level approach to the analysis of narrative language in aphasia. *Aphasiology*, *25*(11), 1372–1392.

McCarney, C. T., & Johnson, A. F. (2001). Examining conversational characteristics of persons with chronic aphasia during three group conversations. *Aphasiology*, *15*(10–11), 1017–1028.

McCullough, K. C., Lance, D., & Beverly, B. L. (2017). Comparison of structured and unstructured discourse tasks in persons with aphasia. *Clinical Archives of Communication Disorders*, *2*(1), 23–29.

Miller, J. F., & Iglesias, A. (2008). *Systematic analysis of language transcripts (SALT), English & Spanish (Version 9)[Computer software]*. University of Wisconsin—Madison, Waisman Center. Language Analysis Laboratory.

Myrberg, K., Hydén, L. C., & Samuelsson, C. (2018). Different approaches in aphasia assessments: A comparison between test and everyday conversations. *Aphasiology*, *32*(4), 417–435.

Newhoff, M., Tonkovich, J. D., Schwartz, S. L., & Burgess, E. K. (1982). Revision strategies in aphasia. In *Clinical aphasiology: Proceedings of the Conference 1982* (pp. 83–84). BRK Publishers.

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.10.1044/jshr.3602.338

Nicholas, L. E., & Brookshire, R. H. (1995). Presence, completeness, and accuracy of main concepts in the connected speech of non-brain-damaged adults and adults with aphasia. *Journal of Speech, Language, and Hearing Research*, *38*(1), 145–156. https://doi.10.1044/jshr.3801.145

Oelschlaeger, M. L., & Thorne, J. C. (1999). Application of the correct information unit analysis to the naturally occurring conversation of a person with aphasia. *Journal of Speech, Language, and Hearing Research*, *42*(3), 636–648. https://doi.10.1044/jslhr.4203.636

Peach, R. K., & Coehlo, C. A. (2016). Linking inter- and intra-sentential processes for narrative production following traumatic brain injury: Implications for a model of discourse processing. *Neuropsychologia*, *80*, 157–164. https://doi.10.1016/j.neuropsychologia.2015.11.015

Prinz, P. M. (1980). A note on requesting strategies in adult aphasics. *Journal of Communication Disorders*, *13*(1), 65–73. https://doi.10.1016/0021-9924(80)90023-4

Pritchard, M., Hilari, K., Cocks, N., & Dipper, L. (2017). Reviewing the quality of discourse information measures in aphasia. *International Journal of Language and Communication Disorders*, *52*(6), 689–732. https://doi.10.1111/1460-6984.12318

Ramsberger, G., & Rende, B. (2002). Measuring transactional success in the conversation of people with aphasia. *Aphasiology*, *16*(3), 337–353. https://doi.org/10.1080/02687040143000636

Richardson, J. D., & Dalton, S. G. (2016). Main concepts for three different discourse tasks in a large non-clinical sample. *Aphasiology*, *30*(1), 45–73. https://doi.org/10.1080/02687038.2015.1057891

Rochon, E., Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (2000). Quantitative analysis of aphasic sentence production: Further development and new data. *Brain and Language*, *72*(3), 193–218. https://doi.10.1006/brln.1999.2285

Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (1989). The quantitative analysis of agrammatic production: Procedure and data. *Brain and Language*, *37*(3), 440–479. https://doi.10.1016/0093-934x(89)90030-8

Schienberg, S., & Holland, A. (1980). Conversational turn-taking in Wernicke aphasia. In *Clinical aphasiology: Proceedings of the Conference 1980* (pp. 106–110). BRK Publishers.

Shadden, B. B., & Agan, J. P. (2004). Renegotiation of identity: The social context of aphasia support groups. *Topics in Language Disorders*, *24*(3), 174–186.

Shadden, B. B., & Hagstrom, F. (2007). The role of narrative in the life participation approach to aphasia. *Topics in Language Disorders*, *27*(4), 324–338.

Simmons-Mackie, N., Elman, R., Holland, A., & Damico, J. (2007). Management of discourse in group therapy for aphasia. *Topics in Language Disorders*, *27*, 5–23. https://doi.10.1097/00011363-200701000-00003

Swinburn, K., Porter, G., & Howard, D. (2005). *The Comprehensive Aphasia Test*. Psychology Press.

Tanaka, T., Branch, W., Dalton, S. G., & Richardson, J. D. (2016, November). *It's raining cats: Further development of discourse checklists for picture description tasks* [Poster presentation]. American-Speech-Language-Hearing Association Convention, Philadelphia, PA.

Tetnowski, J. T., Tetnowski, J. A., & Damico, J. S. (2020). Patterns of conversation trouble source and repair as indices of improved conversation in aphasia: A multiple-case study using conversation analysis. *American Journal of Speech-Language Pathology*, 1–18. Advance online publication. https://doi.org/10.1044/2020_AJSLP-19-00100

Ulatowska, H. K., Allard, L., Reyes, B. A., Ford, J., & Chapman, B. (1992). Conversational discourse in aphasia. *Aphasiology*, *6*(3), 325–330.

Walker, G. M., & Schwartz, M. F. (2012). Short-form Philadelphia Naming Test: Rationale and empirical evaluation. *American Journal of Speech-Language Pathology*, *21*(2), S140–S153. https://doi.10.1044/1058-0360(2012/11-0089)

Worrall, L., Sherratt, S., Rogers, P., Howe, T., Hersh, D., Ferguson, A., & Davidson, B. (2011). What people with aphasia want: Their goals according to the ICF. *Aphasiology*, *25*(3), 309–322.

**APPENDIX A**

**Language Samples**

A.B.: Pretreatment

The cat got in the tree was chased by the dog. She's crying for the cat so the man go up. And he gonna get the cat but the ladder fall down and somebody called the firemen and the uh uh they coming to put the ladder so the man could get down. I don't know how they gonna get the cat out. And the bird is watching. He over there tweeting.

A.B.: Posttreatment

The girl call her dad for help to get the cat out of the tree. The dad c-climbed the ladder but the ladder. He got scared of the dog chasing the tree and ladders fell. And somebody called the cops and the cops bring another ladder. And they run they gonna run to get the man out of the tree and get the cat out of the tree. What else? The man was getting the cat And the dog was chasing the man Somehow or another he he he knocked the ladder down And then the people called the fire engine and the firemen bringing the ladder to put back up to take the man out of the tree Which will take the cat out of the tree There's a bird over here I don't know what he's got to do with it. X Nobody with it for the bird.

C.D.: Pretreatment

Woman, man tar ... tar ... bird, cat ... baby, and uh woman. The uh bird, no the one two one man, man, walkin and go with to [unintelligible].

*(prompt: anything else you can tell me about what's happening?)*

Well uh I forget. Yeah.

C.D.: Posttreatment

Dog goes here and ... woman no man goes goo and uh goes beautiful come on one two pull goes here and uh bird cat baby ... forget ... pull the and go this this let's see and that's it ... well no.

C.D.: Pretreatment Conversation

| | Words | Conversation Turns | Maintenance | Substantive | Uncodeable |
|---|---|---|---|---|---|
| C: So can you tell me a little bit about what you did this weekend? | | | | | |
| P: Uh yeah | 1 | 1 | | 1 | |
| P: Uh uh there was uh (wife) and (son) and and uh (another relative) and uh (another) and uh I forget the other | 3 | | | 1 | |
| P: I forget the other | 4 | | | 1 | |
| C: mhm | | | | | |
| P: But were I was making the cut | 7 | 1 | | | 1 |
| P: Sh it was cold uh everybody hot | 2 | | | 1 | |

*(continues)*

C.D.: Pretreatment Conversation *(Continued)*

| | Words | Conversation Turns | Maintenance | Substantive | Uncodeable |
|---|---|---|---|---|---|
| P: (XYZ names) had cold and (name) | 3 | | | 1 | |
| P: Uh it went like that | 4 | | | | 1 |
| P: Wow it go boom boom and uh | 5 | | | | 1 |
| P: And uh uh so so fun uh and uh get uh be home | 5 | | | 1 | |
| P: Uh wait a minute | 3 | | 1 | | |
| P: Yeah they go and uh uh hours | 5 | | | 1 | |
| P: And uh (wife) and me was waiting and | 4 | | | 1 | |
| P: (name) Go boom boom boom boom boom and I have | 3 | | | | 1 |
| P: And then I said holy shit | 6 | | | 1 | |
| P: Boom boom this this this this | 4 | | | | 1 |
| P: Coming car | 2 | | | 1 | |
| C: Yeah? | | | | | |
| P: An and then yay! | 3 | 1 | | 1 | |
| P: And then uh | 0 | | | | 1 |
| C: Yeah nice | | | | | |
| (345 seconds) — Totals: | 64 | 3 | 1 | 11 | 6 |

C.D.: Posttreatment Conversation Transcript

| | Words | Conversation Turns | Maintenance | Substantive | Uncodable |
|---|---|---|---|---|---|
| (I conversational partner) So how was your weekend? | | | | | |
| Good, boat | 2 | 1 | | 1 | |
| No, capin | 2 | | | 1 | |
| And theres uh | 2 | | 1 | | |
| 1 2 3 4 5 6 7 8 9 | 9 | | | 1 | |
| The uh | 0 | | 1 | | |
| And uh | 0 | | 1 | | |
| This this ***(points to map of home town on iPad)*** | 2 | | | 1 | |
| That's right right right | 2 | | 1 | | |
| Yeah yeah yeah | 2 | | 1 | | |
| Island pond road | 3 | | | 1 | |

*(continues)*

C.D.: Posttreatment Conversation Transcript *(Continued)*

| | Words | Conversation Turns | Maintenance | Substantive | Uncodable |
|---|---|---|---|---|---|
| No no no | 2 | | | 1 | |
| Within 18 (city), (state) | 4 | | | 1 | |
| Well yeah | 2 | | 1 | | |
| Good no no no no | 2 | | 1 | | |
| Yeah | 1 | | 1 | | |
| Oh boy | 2 | | 1 | | |
| No no no | 1 | | | 1 | |
| Warm o yeah | 1 | | 1 | | |
| Uh uh let's see | 2 | | 1 | | |
| It was huh | 0 | | 1 | | |
| 6 | 0 | | | 1 | |
| Yup | 0 | | 1 | | |
| Yeah | 0 | | | | |
| No no this | 2 | | 1 | | |
| (city) | 1 | | | 1 | |
| Well uh ok | | | 1 | | |
| Goes uh shit | | | 1 | | |
| See get uh | | | 1 | | |
| What wait a minute uh | 4 | | 1 | | |
| ***Draws on paper to illustrate how many tens of people had been to his house the weekend before for a party, with a circle indicating each group of 10*** | | | | | |
| One, two, three, four, five, six, seven, eight, | 8 | | | 1 | |
| This uh | | | 1 | | |
| C: mmhmm | | | | | |
| This is work work work work | 4 | 1 | | 1 | |
| That's right yeah yeah yeah yeah | 3 | | 1 | | |
| Many | 1 | | 1 | | |
| Yeah yeah yeah | 1 | | 1 | | |
| Beautiful yeah yeah yeah | 1 | | 1 | | |
| Uhh yeah maybe | 1 | | 1 | | |
| Yeah no | 1 | | 1 | | |
| Well no no because uh we were shuh | | | 1 | | |
| Was it | | | 1 | | |
| Kids great great but I wouldn't go to goes goes to the jers | 6 | | | 1 | |

*(continues)*

C.D.: Posttreatment Conversation Transcript *(Continued)*

| | Words | Conversation Turns | Maintenance | Substantive | Uncodable |
|---|---|---|---|---|---|
| No non no no . . . no no no no no | | | 1 | | |
| It was pretty good | 4 | | | 1 | |
| Oh yeha yeah yeah yeah | | | 1 | | |
| But it was pretty good we went in boat boat captain, | 4 | | | 1 | |
| And uh it goes this . . . this here this | 4 | | 1 | | |
| This ***(drew a big island pond)*** | | | | 1 | |
| That's the boat and then we did this | 7 | | | 1 | |
| Right right right right, yeah yeah uh | 1 | | 1 | | |
| Yeah well you see hahaha uh work work work (wife) and me work work work work because because you don't you don't oh yeah yeah yeah | 10 | | | 1 | |
| Because why because boat boat | 2 | | | 2 | |
| No no no no | 1 | | 1 | | |
| That's correct | 2 | | 1 | | |
| No no but tracker | 1 | | 1 | | |
| Well come over! | 3 | | | 1 | |
| Yeah that's right hahaa | 3 | | 1 | | |
| Oh boy nice . . . uh working because uh yeah | 3 | | 1 | | |
| No no no no no | 1 | | 1 | | |
| Wes' gon his this yea this is work work and then uh | 3 | | | 1 | |
| This ***(pointed to 31)*** and yes | 2 | | | 1 | |
| No gone gone gone ok now 6 | 5 | | | 1 | |
| There's 70 | | | | | 1 |
| And her gonna gonna we're gonna (drew 11:00–6:00) | 2 | | | | 1 |
| Oh shit no jeez boat boat | 3 | | | 1 | |
| Well yeah and uh (wife) | 1 | | | 1 | |
| partner | 1 | | | 1 | |

*(continues)*

C.D.: Posttreatment Conversation Transcript *(Continued)*

| Conversation Turns | Words | Maintenance | Substantive | Uncodable |
|---|---|---|---|---|
| But uh working and uh goes to | 2 | | 1 | |
| We get a rib rib steak and lopster | 3 | | 1 | |
| Yeah yeah yeah yeah | | 1 | | |
| And uh and I mean big big big | 2 | | 1 | |
| So that's alright and then boom when it's when it goes sl sleep (wife) me kkk yeah | 10 | | 1 | |
| Ooh yeah yeah yeah but uh after that gone | 5 | 1 | | |
| Yeah yeah | 1 | 1 | | |
| (675 seconds) Totals: | 160 | 40 | 31 | 2 |