

## Research Article

# Discourse Characteristics in Aphasia Beyond the Western Aphasia Battery Cutoff

Davida Fromm,<sup>a</sup> Margaret Forbes,<sup>a</sup> Audrey Holland,<sup>b</sup> Sarah Grace Dalton,<sup>c</sup> Jessica Richardson,<sup>c</sup> and Brian MacWhinney<sup>a</sup>

**Purpose:** This study examined discourse characteristics of individuals with aphasia who scored at or above the 93.8 cutoff on the Aphasia Quotient subtests of the Western Aphasia Battery–Revised (WAB-R; Kertesz, 2007). They were compared with participants without aphasia and those with anomia.

**Method:** Participants were from the AphasiaBank database and included 28 participants who were not aphasic by WAB-R score (NABW), 92 participants with anomia, and 177 controls. Cinderella narratives were analyzed using the Computerized Language Analysis programs (MacWhinney, 2000). Outcome measures were words per minute, percent word errors, lexical

diversity using the moving average type–token ratio (Covington, 2007b), main concept production, number of utterances, mean length of utterance, and proposition density.

**Results:** Results showed that the NABW group was significantly different from the controls on all measures except MLU and proposition density. These individuals were compared to participants without aphasia and those with anomia.

**Conclusion:** Individuals with aphasia who score above the WAB-R Aphasia Quotient cutoff demonstrate discourse impairments that warrant both treatment and special attention in the research literature.

People with aphasia (PWA) who score at or above the cutoff of 93.8 on the Aphasia Quotient (AQ) subtests of the Western Aphasia Battery (WAB; Kertesz, 1982) or the Western Aphasia Battery–Revised (WAB-R; Kertesz, 2007) are regularly included in research and seen in clinics because they consider themselves to be impaired. The description of these individuals in the WAB-R manual as “normal or nonaphasic” (Kertesz, 2007, p. 91) fails to capture the mild deficits that they experience. However, many questions remain to be answered about this population. If they are still aphasic, do they manifest a characteristic set of deficits? Are they an even milder version of mild anomia? Should they be included in our aphasia research as PWA? Should insurance companies cover their treatment? These are just some of the many questions that need to be examined.

It is important to be able to generalize the results of small group studies to some larger target population. With

aphasia, the target population is a collection of individuals with highly variable types and severities of impairments. In order to make the generalization from study groups to populations more accurate, researchers attempt to analyze aphasia groups by type. This group of PWA with WAB-R AQ scores at or above 93.8 does not yet have a widely accepted “type” or a name, although some authors have used “not aphasic by WAB” (NABW; Dalton & Richardson, 2015; Fromm, Forbes, Holland, & MacWhinney, 2013; Richardson et al., 2015). With the exception of these few studies, these individuals have not been treated as a separate group in most aphasia research studies, many of which use the WAB to describe their samples. Some studies report the number of PWA who scored above the WAB’s cutoff and still include them in the aphasia group (Cruice, Pritchard, & Dipper, 2014; Papanicolaou, Moore, Deutsch, Levin, & Eisenberg, 1988; Ross & Wertz, 2003; Sekine & Rose, 2013; Ulatowska et al., 2001; Ulatowska, Reyes, Santos, & Worle, 2011; Wilson et al., 2012). Other studies use the 93.8 cutoff to exclude those who score at or above it (Fergadiotis & Wright, 2011; Pedersen, Vinter, & Olsen, 2004; Sekine, Rose, Foster, Attard, & Lanyon, 2013). Some articles that report WAB results for their participants are not explicit about whether the cutoff was used or not (Bryant et al., 2013; Fergadiotis, Wright, & West, 2013). Even the WAB manual presents conflicting references to these individuals.

<sup>a</sup>Carnegie Mellon University, Pittsburgh, PA

<sup>b</sup>University of Arizona, Tucson

<sup>c</sup>University of New Mexico, Albuquerque

Correspondence to Davida Fromm: fromm@andrew.cmu.edu

Editor: Krista Wilkinson

Associate Editor: Daniel Kempler

Received May 3, 2016

Revision received September 6, 2016

Accepted November 29, 2016

[https://doi.org/10.1044/2016\\_AJSLP-16-0071](https://doi.org/10.1044/2016_AJSLP-16-0071)

**Disclosure:** The authors have declared that no competing interests existed at the time of publication.

Kertesz (2007) explains that these patients with mild aphasia who tested above the cutoff were sometimes considered another control group, sometimes included with the aphasia group, and sometimes excluded from both groups in their studies.

The few studies that have treated these individuals as a separate group have found that they do not always perform similarly to the relevant comparison group: nonaphasic controls, a larger aphasia sample, or subsets with specific subtypes of aphasia. In a recent example, Dalton and Richardson (2015) measured core lexicon (using a measure they devised called CoreLex) and story gist production (using main concept [MC] analysis) for 235 PWA and 166 controls on a picture description task (Broken Window; Menn et al., 1998). The CoreLex measure consisted of unique lemmas (e.g., *run* for *run*, *runs*, *running*, and *ran*) used by 50% or more of the normative sample for the discourse task. MC scoring was done using a multidimensional point scoring system: 0 for MCs that were missing; 1 for an inaccurate and incomplete concept; 2 for an inaccurate and complete concept or an accurate and incomplete concept; and 3 for an accurate and complete concept. The overall MC score was a total of the assigned points. Overall, speakers with aphasia (including those who were NABW) had significantly lower CoreLex scores than did controls. The CoreLex scores of the speakers who were NABW were significantly different from those of controls as well as those with anomic, Broca's, conduction, and Wernicke's aphasia. MC scores of PWA (including those who were NABW) were significantly different from scores of controls. To be specific, MC scores for the NABW group were significantly lower (i.e., fewer attempted MCs) than the MC scores of controls and significantly higher (i.e., more attempted MCs) than the MC scores of individuals with Broca's and Wernicke's aphasia. These results suggest that the NABW group and those with anomic and conduction aphasia did not differ in their ability to communicate the gist of the story, but they differed significantly in the typicality of words used. It is important to note that the NABW group differed from controls on both measures, producing fewer MCs and fewer words of the type produced by nonaphasic speakers describing the same picture (e.g., more general words such as *hit* instead of *kick* a soccer ball, or leaving out a noun such as *lamp* in a picture where a lamp is a prominent feature of the nonaphasic response) in their narratives. More studies examining language and discourse characteristics of large samples of PWA are needed to determine the characteristics of this overlooked group to better ascertain their assessment and intervention needs, accurately interpret existing research findings, and properly design future studies.

The NIH-funded AphasiaBank Project makes it possible to access a large shared database of discourse samples from PWA, as well as persons NABW and controls (MacWhinney, Fromm, Forbes, & Holland, 2011). The standard discourse protocol includes free speech samples (stroke story and coping, important event), picture descriptions (broken window, refused umbrella, cat rescue), a

story narrative (Cinderella), and a procedural discourse task (peanut butter and jelly sandwich). AphasiaBank also collects demographic data and test results using a variety of formal and informal measures. Detailed information about the protocol and associated testing is available at the AphasiaBank website (<http://aphasia.talkbank.org/protocol/>).

The purpose of this study was to use the AphasiaBank database to examine the discourse characteristics of the NABW group and compare them with those of the groups that would be most relevant for purposes of differentiation: controls and individuals with anomic aphasia. Controls score in the same range on the WAB-R AQ (93.8–100.0) as do those in the NABW group; individuals with anomic aphasia are the next highest scoring group on the WAB-R AQ and the ones with the most similar profile of WAB-R AQ subtest scores. This research will indicate whether PWA who score above the WAB-R AQ cutoff differ from these two comparison groups on a selection of discourse measures that are likely to be sensitive to discourse differences in higher functioning participants: words per minute, percent word errors, lexical diversity (according to the moving average type–token ratio [MATTR]; Covington, 2007b), MC production, number of utterances, mean length of utterance (MLU) in words, and proposition density in the Cinderella narrative. These measures represent a range of functional discourse skills (e.g., productivity, lexical processing, and informativeness of discourse) that have been used in many studies of discourse in aphasia (specifically in fluent, anomic, and mild or recovering PWA) and in younger and older non-brain-damaged controls (Andretta, Cantagallo, & Marini, 2012; Capilouto, Wright, & Maddy, 2016; Fergadiotis & Wright, 2011; Marini, Andretta, del Tin, & Carlomagno, 2011; Nicholas & Brookshire, 1993; Richardson & Dalton, 2016; Wright, Capilouto, Wagovich, Cranfill, & Davis, 2005). Most of these measures can be calculated using automated analyses that rapidly produce highly accurate and replicable results, and that do not require human coders with high levels of training or linguistic knowledge. The Cinderella narrative was selected because it typically yields reasonably long and varied speech samples and can be administered in a standardized, reliable fashion for comparison across participants.

## Method

### Participants

Potential participants were PWA (including those who scored above the WAB-R AQ cutoff) and nonaphasic speakers from the AphasiaBank database who met the inclusion criteria and produced a minimum of 20 words on the Cinderella task. Inclusion criteria for both groups were no history of cognitively deteriorating conditions (e.g., Alzheimer's or Parkinson's disease), hearing and vision (aided or unaided) adequate for testing, fluent speakers of English, and no depression at time of testing. The control group had the additional inclusion criterion of no neurological condition (e.g., stroke, head injury). All PWA had

suffered a stroke and were native English speakers. They considered themselves to be people with aphasia and chose to participate in aphasia groups and/or aphasia research.

Some participants within each group spoke English plus a second language learned by 6 years of age: four in the anomic group, two in the NABW group, and one in the control group. The second language was not consistent across participants (e.g., Hungarian, Italian, Ukrainian, Spanish, Hebrew, French). In addition, some participants were late bilinguals, speaking English plus another language learned after 6 years of age: four in the anomic group, two in the NABW group, and seven in the control group. Again, the languages varied (e.g., Spanish, Portuguese, Indonesian, Norwegian, Italian, Japanese, French). Seven participants in the control group and two participants in the anomic group reported that they were multilingual, speaking three or more languages fluently.

Data from any individual with aphasia who was assessed on more than one occasion were included only from the visit in which that individual was first classified as NABW or anomic based on WAB-R scores. With only one exception, all of the individuals in the NABW group were classified as NABW the first time they were assessed. The exception was a participant who was initially (5 years earlier) classified as anomic with a WAB-R AQ of 89.6. For the anomic group, all were classified as anomic the first time they were assessed except for four participants who were initially classified as conduction, one initially classified as transcortical motor and another as transcortical sensory. The number of participants who met these criteria for each group was as follows: 27 NABW, 87 anomic, 177 controls.

## Procedure

All participants completed the standard AphasiaBank discourse protocol. The participants in the anomic and NABW groups also completed the associated testing, which included the WAB-R and other tests of naming, repetition, and auditory comprehension (MacWhinney et al., 2011). With few exceptions (six control sessions recorded with audiotape only), all sessions were recorded on videotape. All participants signed consent forms for the testing and data collection and gave approval for the data to be available for research and teaching.

Transcriptions were completed by trained and experienced transcribers using the Codes for the Human Analysis of Transcripts (CHAT) format, which allows for analyses using the Computerized Language Analysis (CLAN) programs (MacWhinney, 2000). Following the guidelines developed and explained by Saffran, Berndt, and Schwartz (1989) for use in the quantitative production analysis, utterances were segmented on the basis of the following hierarchy of indices: syntax, intonation, pause, semantics. All of these factors may be considered in the decision, but more weight is given to syntax (well-formed sentences) and prosody (falling intonation) than to pauses and semantic criteria. Furthermore, transcribers are urged to be

conservative in defining utterance boundaries and, if unsure, to make shorter rather than longer utterances. Two transcribers reviewed each transcription, and the two reached forced-choice agreement on any discrepancies. Complete transcripts and videos are available at the AphasiaBank website (<http://aphasia.talkbank.org/>).<sup>1</sup> Word errors were coded using the error coding system described at the AphasiaBank website. Examples of types of common errors include:

- Phonemic paraphasias, nonwords; for example, /pɪnts/ for *prince*, /slɪfə/ for *slipper*
- Phonemic paraphasias, words; for example, *coast* for *coach*, *sipper* for *slipper*
- Semantic paraphasias, related; for example, *grandmother* for *godmother*, *shoe* for *foot*
- Semantic paraphasias, unrelated; for example, *pitcher* for *slipper*, *trucks* for *dress*
- Nonword, known target; for example, /tsenz/ for *dance*, /kʌmpəl/ for *pumpkin*.

The Cinderella narrative was extracted from the complete transcripts for CLAN analyses. CLAN analyses excluded any non-task-related comments made by the participants during the task (e.g., comments about the task). Lexical diversity was measured using MATTR (Covington, 2007b), which has recently been demonstrated to be a reliable, valid, and unbiased measure of lexical diversity in aphasia (Fergadiotis et al., 2013). In their confirmatory factor analysis of four lexical diversity measures, Fergadiotis et al. found several advantages in using MATTR. It is a strong indicator of lexical diversity, with no evidence of systematic sample length effects, and small residual variances. MATTR calculates type-token ratios across consecutive nonoverlapping segments (called window lengths) of a sample and then averages them. A window length of 20 words was used here because 20 words was the shortest sample in the study, similar to the technique employed by Fergadiotis et al., who chose a window length of 17 words corresponding to the shortest sample in their study. The *proposition density score*, or the number of ideas per total number of words, was calculated in CLAN according to the rules developed by Covington and colleagues (Brown, Snodgrass, & Covington, 2007; Covington, 2007a) for Computerized Propositional Idea Density Rater. MC scores were measured using the techniques described previously, assigning points for Inaccurate/Incomplete, Inaccurate/Complete, Accurate/Incomplete, and Accurate/Complete concepts (Dalton & Richardson, 2015; Richardson & Dalton, 2016). Concept points were summed to generate an overall MC score for each participant. Inter- and intrarater reliabilities for MC coding in transcripts from

---

<sup>1</sup>AphasiaBank membership allows for access to password-protected participant data. To become a member, send an e-mail request to Brian MacWhinney ([macw@cmu.edu](mailto:macw@cmu.edu)).

control participants were reported to range from 91% to 96.3% (Richardson & Dalton, 2016).

### Statistical Analysis

Exploratory data analysis revealed that some of the distributions had outlier variables, with the data being skewed right (e.g., exceptionally long stories with more total words, more utterances). To address this, log transformations were applied to all of the data (with the exception of percent word errors, which will be explained below) to make the distributions more symmetrical and normal. Statistical tests were run on both the log-transformed data and the nontransformed data. Results (in terms of significant group differences on each variable) did not differ when calculated on the log-transformed and nontransformed data, but statistical test results will be reported on the basis of the log-transformed data. Tables, however, will show the nontransformed data, for ease of interpretation.

Group differences were analyzed using analysis of variance; post hoc testing was done using Tukey's honest significant difference (HSD) test. The Tukey HSD test is a conservative test, specifically designed for pairwise comparisons. For percent word errors, the controls mostly had scores of 0. Therefore, a chi-square test was used to determine whether the groups differed based on the simple dichotomy of having 0% word errors or greater than 0% word errors.

Because the range of possible WAB-R AQ scores for the anomic aphasia group can range by over 30 points in comparison with the limited 6.2-point range of possible AQ scores for the NABW group (93.8–100.0), a secondary analysis was conducted between the NABW group and the top-scoring group of participants from the anomic aphasia group (anomic-top subgroup,  $n = 36$ ) using two-tailed  $t$  tests. Therefore, anyone from the anomic group whose WAB-R AQ was between 87.5 and 93.7 (same 6.2-point range) was included in the anomic-top subgroup for this separate analysis. This comparison could help determine whether the NABW group was similar to those with the mildest anomic aphasia. The significance levels for all statistical testing will be reported, but all  $p$  values that are less than .001 will be given as  $p < .001$ .

### Results

Demographic characteristics of the participants appear in Table 1. The three primary groups (excluding the anomic-top subgroup) did not differ significantly in age ( $p = .62$ ) or years of education ( $p = .59$ ). Mean ages ranged from 62.3 years (anomic) to 64.0 years (controls), and mean years of education ranged from 15.4 years (controls) to 15.9 years (NABW). Time post-onset for those in the NABW (mean = 5.3 years) and anomic (mean = 4.5 years) groups did not differ significantly ( $p = .30$ ). The distribution of men and women across groups was not statistically significant ( $p = .14$ ), with men accounting for

56% of the anomic group, 46% of the control group, and 37% of the NABW group.

The absolute sample sizes were significantly different across groups,  $F(2, 288) = 46.73, p < .0001$ . Mean number of words produced for the Cinderella story telling for the anomic group was 224 ( $SD = 161$ ), for the NABW group was 241 ( $SD = 104$ ), and for the control group was 476 ( $SD = 284$ ). Tukey HSD post hoc testing revealed significant differences between the control group and both the anomic ( $Q = 13.25, p = .001$ ) and NABW groups ( $Q = 6.05, p = .001$ ), but not between the anomic and NABW groups ( $Q = 2.2, p = .27$ ).

Table 2 shows means and standard deviations on discourse outcome measures for all groups. Analysis of variance testing for the NABW, anomic, and control groups revealed significant differences ( $p < .001$ ) across groups on all discourse measures. Post hoc pairwise comparison results appear in Table 3. All groups differed significantly from each other on words per minute, main concepts, and lexical diversity (MATTR). For total utterances, controls differed significantly from both the NABW and anomic groups, but the NABW and anomic groups were statistically similar. For MLU and proposition density, controls and NABW groups differed significantly from the anomic group, but the controls and NABW groups were statistically similar. Chi-square results revealed that the groups differed significantly ( $\chi^2 = 143.13, p < .001$ ) on percent word errors.

Comparing the NABW group with the anomic-top subgroup revealed significant differences for words per minute ( $t = -5.15, p < .001$ ), MLU ( $t = 3.43, p = .001$ ), percent word errors ( $t = -3.92, p < .001$ ), proposition density ( $t = -2.46, p = .02$ ), and MATTR ( $t = -2.47, p = .02$ ). The only outcome measures on which the NABW and anomic-top subgroup did not differ were total utterances and main concepts.

### Discussion

Individuals who consider themselves to be aphasic but who test above the WAB-R AQ cutoff for aphasia demonstrated discourse characteristics that significantly differed from nonaphasic controls. In their Cinderella stories, they had fewer total utterances, fewer words per minute, less lexical diversity, more word errors, and fewer main concepts. They did not differ on MLU or proposition density. On the other hand, the individuals in the NABW group also differed significantly from the full anomic group on all discourse measures except total utterances. They had more words per minute, more lexical diversity, fewer word errors, more main concepts, longer MLU, and more proposition density. When compared with those scoring in the top of the anomic group, they still differed significantly on all measures except total utterances and main concepts. The full anomic group differed from the controls on all discourse measures. To be clear, this is not intended to suggest that these individuals do not demonstrate word-finding problems, nor does it mean they do not demonstrate

**Table 1.** Demographic characteristics: Means and standard deviations.

Parameter	Anomic (n = 87)	Anomic-Top (n = 36)	NABW (n = 27)	Controls (n = 177)
Age (yrs.)	62.2 (11.9)	61.8 (11.5)	62.4 (13.8)	64.0 (17.0)
Education (yrs.)	15.6 (2.6)	15.4 (2.4)	15.9 (2.9)	15.4 (2.4)
Sex (% male)	56	53	37	46
Time post-onset	4.5 (3.7)	5.3 (4.5)	5.3 (3.8)	—
WAB-R AQ	85.0 (6.5)	90.8 (1.7)	96.2 (1.7)	—

Note. Em dash indicates data not applicable. NABW = not aphasic by WAB; WAB-R = Western Aphasia Battery-Revised; AQ = Aphasia Quotient.

a mild form of what many would consider to be a fluent, anomic aphasia. The results of this study suggest that the discourse of these participants who score above the WAB-R AQ cutoff score cannot be considered unimpaired or nonaphasic, yet most of their language impairments are significantly milder than those who score in the anomic aphasia category on the WAB.

The NABW group in this sample produced discourse that was slower, with restricted vocabulary, more word errors (e.g., *dollars* for *daughters*, *sepmother* for *stepmother*, *stepsitters* for *stepsisters*, *feets* for *feet*, *mouses* for *mice*, *blippers* for *slippers*), fewer utterances, and less information about the gist of the topic than nonaphasic control speakers. Most of the word errors could be characterized as simple phonemic paraphasias (e.g., *blippers* for *slippers*, *sepmother* for *stepmother*, *stepsitters* for *stepsisters*, *weave* for *leave*). The reduced words per minute in these speakers is affected by the presence of pauses, fillers, revisions, and repetitions that are not counted as actual words. Although little research has singled out this group for analysis, Dalton and Richardson (2015) also reported significant differences between the NABW and control groups on main concepts, and Fromm et al. (2013) reported significantly reduced speech rate (as measured by words per minute) for the NABW group in comparison with controls on a procedural discourse task (how to make a peanut butter and jelly sandwich). Such limitations, as evidenced by the discourse data for NABW participants, may create difficulties navigating daily communication encounters and engaging in conversation with others.

A slightly different set of similarities and differences was observed for the NABW group and those with anomic aphasia. The discourse analyses demonstrated similarities

between these two groups in the quantity of output, as measured by total words and total utterances. However, the NABW group was less impaired on all other measures, including rate, utterance length, word errors, lexical diversity, proposition density, and communicating the gist of the story. With the exception of this final variable, the NABW group even differed significantly from the mildest subgroup of those with anomic aphasia. Examples of word errors from the transcripts of speakers with anomic aphasia include various mispronunciations of *Cinderella* and *stepsisters*, pronoun errors (e.g., *he* for *she*, *she* for *he*, *she* for *they*), and semantic errors (e.g., *king* for *prince*, *princess* for *godmother*, *foot* for *shoe*).

These results provide confirmation and additional information about the discourse differences that exist for this group of PWA who test above the WAB-R AQ cutoff. They also help to define the needs of this group and reveal the deficits that likely warrant continued treatment and coverage by third-party payers in spite of their standardized test scores. For example, spoken language marked by word errors, slower rate of speech, and decreased essential content will affect conversation, communication, and/or participation. In addition, as other researchers have urged, the results reinforce the need to carefully select multiple types and levels of discourse measures (micro- and macro-linguistic) because “normal” performance on one does not necessarily mean that no discourse deficits are present (Coelho, Grela, Corso, Gamble, & Feinn, 2005; Marini et al., 2011; Sherratt, 2007). Although the WAB-R, like other formal aphasia tests, is designed to assess many skills, the assessment of fluency in connected speech is impressionistic and not highly sensitive. Future investigations should

**Table 2.** Discourse data: Means and standard deviations.

Parameter	Anomic (n = 87)	Anomic-Top (n = 36)	NABW (n = 27)	Controls (n = 177)
Number of utterances	31.6 (20.5)	32.8 (21.8)	27.3 (11.7)	49.4 (34.7)
MLU (words)	7.0 (2.0)	7.4 (1.7)	8.8 (1.7)	9.9 (2.3)
MATTR	0.80 (0.05)	0.81 (0.04)	0.84 (0.03)	0.87 (0.02)
Proposition density	0.46 (0.06)	0.46 (0.05)	0.49 (0.03)	0.50 (0.03)
% Word errors	4.9 (4.5)	4.9 (4.4)	1.5 (1.3)	0.06 (0.13)
Words per minute	56.0 (27.5)	54.5 (26.6)	86.9 (24.6)	134.3 (30.9)
Main concepts	27.1 (15.6)	33.4 (15.4)	42.3 (18.4)	55.9 (18.9)

Note. NABW = not aphasic by WAB; MLU = mean length of utterance; MATTR = moving average type-token ratio.

**Table 3.** Discourse data: ANOVA and Tukey's HSD post hoc test results.

Parameter	Anomic vs. NABW	Anomic vs. Controls	NABW vs. Controls
Number of utterances			
<i>Q</i>	.30	7.58	5.13
<i>p</i>	<i>ns</i>	.001	.001
MLU (words)			
<i>Q</i>	6.81	16.13	2.96
<i>p</i>	.001	.001	<i>ns</i>
MATTR			
<i>Q</i>	5.93	20.54	6.70
<i>p</i>	.001	.001	.001
Proposition density			
<i>Q</i>	4.04	9.06	1.43
<i>p</i>	.013	.001	<i>ns</i>
Words per minute			
<i>Q</i>	9.33	28.88	8.36
<i>p</i>	.001	.001	.001
Main concepts			
<i>Q</i>	6.15	16.31	3.78
<i>p</i>	.001	.001	.022

Note. ANOVA = analysis of variance; HSD = honest significant difference; NABW = not aphasic by WAB; MLU = mean length of utterance; MATTR = moving average type-token ratio; *ns* = nonsignificant.

consider probing further within each discourse level, for example, focusing on other indices of fluency besides words per minute (e.g., repetitions, revisions, false starts, pauses) or other indices of narrative analysis (e.g., cohesion and local and global coherence).

This study has a number of potential implications concerning research and treatment. Given that those who score above the WAB-R AQ cutoff still demonstrate impairments in discourse performance, although not to the same degree as the majority of individuals with anomic aphasia who score below the cutoff, researchers are advised to specify whether they are including or excluding these individuals, especially for studies of speech production. Even when the anomic group was limited to the most mildly impaired, they were still significantly different from the NABW group on everything but total number of utterances and main concepts. Further research comparing various types of discourse tasks performed by larger numbers of individuals who are NABW is encouraged to help clarify their speech and language characteristics and to systematically explore functional differences in this group. Although this study focused on the WAB-R, other commonly used batteries that determine presence or absence of aphasia or make a determination of “normal” functioning may also not identify these individuals as having aphasia. Furthermore, these tests are often used for determining treatment needs. Although these batteries assess many essential language skills impaired in aphasia, they often use subjective rating scales to assess the varied and complex characteristics of connected speech. Attention to higher-level language processes and validation of their impact is necessary for making appropriate decisions about good treatment and fair payment.

In addition to the data-based differences between individuals who are NABW and their control counterparts,

in informal exchanges they frequently comment to the effect that although they may “appear to be back to normal, they aren’t there yet.” They often express frustration that they do not qualify for research studies, disability benefits, or insurance coverage for treatment but are often unable to return to work because of language difficulties. Clinicians and families should take those comments quite seriously and validate such concerns. Furthermore, even if some of these individuals manage to receive therapy, few clinical tasks designed for aphasia treatment seem to be appropriate for them, and approaches to their treatment should be carefully reconsidered. These might include reading parts in plays to normalize speech rate and reduce disfluencies (Whitney & Goldstein, 1989), recording and critiquing conversations to produce fuller and more accurate discourse, and learning personalized scripts to use for communication in daily life.

To conclude, measurement of the mild but persistent deficits observed in this population may help to advance our understanding of some of the difficulties such individuals face if and when they return to employment. This will also require continued development and refinement of informative and efficient tools that are clinically efficient and practical to use for measuring these mild deficits.

## Acknowledgment

This work was funded by National Institute on Deafness and Other Communication Disorders Grant R01-DC008524 (2012–2017), awarded to Brian MacWhinney.

## References

- Andreetta, S., Cantagallo, A., & Marini, A. (2012). Narrative discourse in anomic aphasia. *Neuropsychologia*, *50*, 1787–1793.

- Brown, C., Snodgrass, T., & Covington, M.** (2007). *Computerized Propositional Idea Density Rater 3 (CPIDR 3)* [Computer software]. Athens, GA: University of Georgia, CASPR Project, Artificial Intelligence Center.
- Bryant, L., Spencer, E., Ferguson, A., Craig, H., Colyvas, K., & Worrall, L.** (2013). Propositional Idea Density in aphasic discourse. *Aphasiology, 27*, 992–1009.
- Capilouto, G. J., Wright, H. H., & Maddy, K. M.** (2016). Micro-linguistic processes that contribute to the ability to relay main events: Influence of age. *Aging, Neuropsychology, and Cognition, 23*, 445–463.
- Coelho, C. A., Grela, B., Corso, M., Gamble, A., & Feinn, R.** (2005). Microlinguistic deficits in the narrative discourse of adults with traumatic brain injury. *Brain Injury, 19*, 1139–1145.
- Covington, M. A.** (2007a). *CPIDR 3 user manual*. Athens, GA: University of Georgia Artificial Intelligence Center.
- Covington, M. A.** (2007b). *MATTR user manual* (CASPR Research Report 2007–05). Athens, GA: University of Georgia Institute for Artificial Intelligence.
- Cruice, M., Pritchard, M., & Dipper, L.** (2014). Verb use in aphasic and non-aphasic personal discourse: What is normal? *Journal of Neurolinguistics, 28*, 31–47.
- Dalton, S. G., & 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*(Suppl.), S923–S938. [https://doi.org/10.1044/2015\\_AJSLP-14-0161](https://doi.org/10.1044/2015_AJSLP-14-0161)
- Fergadiotis, G., & Wright, H. H.** (2011). Lexical diversity for adults with and without aphasia across discourse elicitation tasks. *Aphasiology, 25*, 1414–1430.
- Fergadiotis, G., Wright, H. H., & West, T. M.** (2013). Measuring lexical diversity in narrative discourse of people with aphasia. *American Journal of Speech-Language Pathology, 22*, 397–408.
- Fromm, D., Forbes, M., Holland, A., & MacWhinney, B.** (2013, May). *PWAs and PBJs: Language for describing a simple procedure*. Poster presented at the Clinical Aphasiology Conference, Tucson, AZ.
- Kertesz, A.** (1982). *Western Aphasia Battery test manual*. New York, NY: Grune & Stratton.
- Kertesz, A.** (2007). *Western Aphasia Battery–R*. New York, NY: Grune & Stratton.
- MacWhinney, B.** (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Erlbaum.
- MacWhinney, B., Fromm, D., Forbes, M., & Holland, A.** (2011). AphasiaBank: Methods for studying discourse. *Aphasiology, 25*, 1286–1307.
- Marini, A., Andreetta, S., del Tin, S., & Carlomagno, S.** (2011). A multi-level approach to the analysis of narrative language in aphasia. *Aphasiology, 25*, 1372–1392.
- Menn, L., Reilly, K. F., Hayashi, M., Kamio, A., Fujita, I., & Sasanuma, S.** (1998). The interaction of preserved pragmatics and impaired syntax in Japanese and English aphasic speech. *Brain and Language, 61*, 183–225.
- Nicholas, L. E., & Brookshire, R. H.** (1993). A system for scoring main concepts in the discourse of non-brain-damaged and aphasic speakers. *Clinical Aphasiology, 21*, 87–99.
- Papanicolaou, A. C., Moore, B. D., Deutsch, G., Levin, H. S., & Eisenberg, H. M.** (1988). Evidence for right-hemisphere involvement in recovery from aphasia. *Archives of Neurology, 45*, 1025–1029.
- Pedersen, P. M., Vinter, K., & Olsen, T. S.** (2004). Aphasia after stroke: Type, severity and prognosis. The Copenhagen Aphasia Study. *Cerebrovascular Diseases, 17*, 35–43.
- Richardson, J. D., & Dalton, S. G.** (2016). Main concepts for three different discourse tasks in a large non-clinical sample. *Aphasiology, 30*, 45–73.
- Richardson, J. D., Dalton, S. G. H., Schultz, C., Harlow, V., Haley, K., Jacks, A., & Silverman, M.** (2015, November). *Story grammar analysis in persons with aphasia*. Poster presented at the American-Speech-Language-Hearing Association convention, Denver, CO.
- Ross, K. B., & Wertz, R. T.** (2003). Discriminative validity of selected measures for differentiating normal from aphasic performance. *American Journal of Speech-Language Pathology, 12*, 312–319.
- Saffran, E. M., Berndt, R. S., & Schwartz, M. F.** (1989). The quantitative analysis of agrammatic production: Procedure and data. *Brain and Language, 37*, 440–479.
- Sekine, K., & Rose, M.** (2013). The relationship of aphasia type and severity to gesture productions in people with aphasia. *American Journal of Speech-Language Pathology, 22*, 662–672.
- Sekine, K., Rose, M., Foster, A., Attard, M., & Lanyon, L.** (2013). Gesture production patterns in aphasic discourse: In-depth description and preliminary predictions. *Aphasiology, 27*, 1031–1049.
- Sherratt, S.** (2007). Multi-level discourse analysis: A feasible approach. *Aphasiology, 21*, 375–393.
- Ulatowska, H. K., Olness, G. S., Wertz, R. T., Thompson, J. L., Keebler, M. W., Hill, C. L., & Auther, L. L.** (2001). Comparison of language impairment, functional communication, and discourse measures in African-American aphasic and normal adults. *Aphasiology, 15*, 1007–1016.
- Ulatowska, H. K., Reyes, B. A., Santos, T. O., & Worle, C.** (2011). Stroke narratives in aphasia: The role of reported speech. *Aphasiology, 25*, 93–105.
- Whitney, J., & Goldstein, H.** (1989). Using self-monitoring to reduce dysfluencies in speakers with mild aphasia. *Journal of Speech and Hearing Disorders, 54*, 576–586.
- Wilson, K. R., O'Rourke, H., Wozniak, L. A., Kostopoulos, E., Marchand, Y., & Newman, A. J.** (2012). Changes in N400 topography following intensive speech language therapy for individuals with aphasia. *Brain and Language, 123*, 94–103.
- Wright, H. H., Capilouto, G., Wagovich, S., Cranfill, T., & Davis, J.** (2005). Development and reliability of a quantitative measure of adults' narratives. *Aphasiology, 19*, 263–273.