Evaluating the Accuracy of Self-Ratings of Language in Adults with Aphasia and Non-Brain Injured Adults: A Pilot Study

Sarah Arnett, B.A.[1] **and Jennifer Mozeiko, Ph.D., CCC-SLP**[1]

ABSTRACT

Rating scales are frequently used in research and clinical practice with people with aphasia (PWA) to characterize communication in the home environment. However, it remains unclear whether responses provided on rating scales accurately reflect the communication that occurs. We aim to evaluate the accuracy of PWA's self-perceptions of verbal language use as measured by a rating scale and determine whether this accuracy is different from that of non–brain-injured (NBI) participants. Four PWA and four NBI participants completed a rating scale estimating their amount of verbal language production as compared with their communication partner. Audio recordings from participants' home environments were analyzed for proportion of words and conversational turns contributed by the participant, which were compared with rating scale estimates. Perceptions of verbal language output among both PWA and NBI participants showed variable accuracy, with discrepancies between estimates and objective data across both groups. The reliability of rating scales in quantifying language output appears questionable, suggesting they may not accurately represent naturalistic language environments of PWA. Additional research with larger sample sizes is warranted to investigate whether this trend is consistent across a larger population of individuals with aphasia.

KEYWORDS: aphasia, rating scales, home environment, self-perception

[1]Department of Speech, Language, and Hearing Sciences, University of Connecticut, North Windham, Connecticut.

Address for correspondence: Sarah Arnett, B.A., Department of Speech, Language, and Hearing Sciences, University of Connecticut, 61 Hilltop Dr. Apt. 2, North Windham, CT 06256 (e-mail: sarah.arnett@uconn.edu).

Data-Based Research Articles.

Learning Outcomes: As a result of this activity, the reader will be able to (1) explain the limitations of subjective rating scales in characterizing the verbal language use of people with aphasia; (2) compare the reliability of self-perception of language use between adults with aphasia and non–brain-injured adults; (3) expand foundational knowledge base in a way that assists in making clinical judgments as to when the use of subjective rating scales may or may not be appropriate in clinical practice.

Though much of aphasia rehabilitation occurs during intervention in clinical settings, a considerable portion of the process is also influenced by language use outside of clinical settings and the communicative environment of the homes of people with aphasia (PWA). It is likely that understanding the way that PWA use language in nonclinical environments, particularly at home, would provide valuable information for speech-language pathologists looking to maximize language gains and optimize the rehabilitation process. This might take the form of assigning more meaningful home practice programs, counseling PWA and their families on specific ways to create a language-rich environment, or even as an additional means of tracking progress and treatment efficacy. Given that most of our time is spent in the home environment, it makes intuitive sense that what happens there would contribute to, or potentially impede, rehabilitative progress. Secondary outcomes in a variety of studies support this notion. As an example, Meinzer et al[1] found that participants who received intensive constraint-induced language therapy showed greater maintenance of functional gains post-treatment if their families had been involved in the treatment process.

At this time, however, self-reports from PWA and their caregivers are the only accessible means of obtaining this information, and it is not clear whether these self-reports are reliable. The use of rating scales to quantify functional communication, quality of life, and verbal output has a significant historical role in aphasia rehabilitation and research.[2–4] However, these rating scales are often completed by clinicians or caregivers, and both the reliability of direct patient reports and the applicability of other respondents' answers to a PWA's perspective on their language use remain in question. If clinicians are able to get a reliable representation of the richness of the communication environments at home, it might provide

great clinical utility in designing treatment plans, providing communication partner training, and optimizing elements of the home environment for greater language facilitation.

Rating Scales

Rating scales do give clinicians a subjective look into the communication dynamics of the home environment, but there are many limitations to the validity and reliability of the information provided through such measures. The American Speech, Language, and Hearing Association's Functional Assessment of Communication Skills for Adults (ASHA FACS),[5] for example, reports high levels of interrater reliability and moderate external validity when measured against the standardized Western Aphasia Battery-Revised Aphasia Quotient (WAB-R AQ),[6] but there is far less agreement between its scores and the functional communication ratings reported by clinicians and family members.[7,8] Other scales, such as the ASHA Quality of Communication Life Scale (ASHA QCL)[9] and Communication Confidence Rating Scale for Aphasia (CCRSA)[9] have also shown relatively strong psychometric properties in initial evaluations, but have variable levels of accuracy in the assessment of gains pre- and posttreatment.[10]

Other rating scales that assess the linguistic or communicative content of spontaneous speech have been shown to be too insensitive to reflect changes in functional communication.[11] This lack of sensitivity may stem from a variety of causes, including subjective variability, rater bias, or a lack of adequate parameters for ranking elements of spontaneous speech. Despite these observations, rating scales are regularly used as a comparison measure between pretreatment and posttreatment speech.[12–14]

In addition to inherent issues of reliability and validity regarding communicative function, some commonly used rating scales, such as the

Communicative Effectiveness Index (CETI), are designed for caregivers to complete, as opposed to the PWA.[4] This introduces additional questions about the applicability of information gleaned from these rating scales, as it reflects only the caregiver's perception of the PWA's language, and may miss important elements present from the PWA's perspective on their own language use. It has been well-documented that PWA and their caregivers often have very different perceptions of many social and behavioral elements of the PWA's communication impairment, including their communicative needs and the extent of their disability.[15–17] It logically follows that these differences in perspective may be present in each person's perception of the PWA's communication abilities as well. Indeed, a comparison of the results of the CETI when given to PWA and to their communication partners directly showed some disparity between the two perspectives. Though the differences found did not quite meet the level of significance within the scope of the study, the authors noted that the disparity was close enough to the threshold to indicate that a larger sample might, in fact, reveal a more significant difference.[18] This pattern also occurs in rating scales reflecting quality of life, as it has been demonstrated that the communication partners tend to rate PWA as more significantly impacted by their aphasia than PWA rate themselves.[19] This is not to say that either PWA or their caregivers are universally more accurate in reporting language use at home; rather, the fact that individual rating scales may result in the reporting of different information based on the target audience further highlights the difficulty clinicians may face in obtaining reliable and valid information about language productivity in the home environment.

Furthermore, because many surveys require a level of linguistic comprehension, it can be difficult to ascertain whether some questions are phrased with adequate clarity for a rater, particularly one with aphasia, to understand what is being asked and to respond appropriately. Indeed, in multiple research studies, some questions from rating scales have had to be excluded due to high levels of ambiguity or rater confusion.[10,20] Other considerations may also limit the reliability and validity of subjective rating scales, such as the day-to-day variability in responses or the possibility of a participant exaggerating their responses in an attempt to please the clinician. Given these considerations in the use of rating scales, it would be beneficial if clinicians had an objective means of assessing the communicative environment of the homes of PWA.

Aims

The present study aims to add to the literature surrounding the use of rating scales as outcome measures through the collection of naturalistic conversation samples obtained in the home environments of PWA and their communication partners. Through comparison of participant responses on a scale rating their perceived verbal language use to the actual amount of verbal language produced in their home environments, we contribute data relevant to the ecological validity of rating scales in clinical use. The specific experimental questions were as follows:

1. In terms of total words produced and number of conversational turns, how accurate are PWA's perceptions of their own language use?
2. Is this accuracy notably different from an equivalent comparison with non–brain-injured (NBI) individuals?

METHODS

Note: Due to the emergence of the COVID-19 virus and associated precautionary measures taken beginning in March 2020, some logistical elements of the study procedure were adapted for the final two participants, NBI03 and NBI04. While every attempt was made to keep the procedure as consistent as possible, some changes were necessary to ensure the health and safety of the researchers and participants. These adaptations will be described in the following sections alongside the original procedure.

Participants

We recruited a total of four PWA from the University of Connecticut Aphasia Groups

Table 1 Participants Characteristics

Participant	Gender	Age	Time post onset (Y;M)	Education (years)
PWA01	Male	73	6;1	16
PWA02	Male	57	3;8	16
PWA03	Male	64	6;3	12
PWA04	Male	78	2;4	22
NBI01	Female	50	N/A	14
NBI02	Female	54	N/A	16
NBI03	Male	59	N/A	16
NBI04	Male	83	N/A	19

Abbreviations: NBI, non–brain-injured; PWA, people with aphasia.

and four additional NBI participants. To ensure the communication environments were as comparable as possible, all participants were required to live with exactly one other adult for inclusion in the study. All participants had no self-reported history of any premorbid language or learning disability. Finally, to meet inclusion criteria, participants had to have hearing thresholds of no more than 40 dB at 500, 1,000, 2,000, and 3,000 Hz frequencies. For pre-COVID participants, this was confirmed by the administration of a hearing screening; for post-COVID participants, since all interactions occurred over video conference, this was confirmed by self-report. See Table 1 for participant characteristics.

Procedure

STANDARDIZED ASSESSMENT
We obtained informed consent from each participant and their primary communication partner. A demographic survey and a confirmation of typical hearing were administered to each participant to ensure their eligibility for the study. To characterize each participant's language abilities and perceived quality of life, PWA completed the Quick Aphasia Battery (QAB),[21] the Discourse Comprehension Test (DCT),[22] and the Stroke and Aphasia Quality of Life Scale (SAQOL).[23] NBI participants completed the QAB, the DCT, and the Quality of Life Scale (QOLS).[24] The QAB was selected

for this study due to its high rate of correlation with results from the WAB and its clinical feasibility based on the shorter length of administration.[21] The DCT asks participants to recall stated and implied details and main ideas from a verbal narrative, which allows for a better depiction of the participant's comprehension and retention skills. Finally, the quality of life scales were chosen to give us insight into any external factors that might affect results, such as general emotional state and social investment. All participants also completed an aphasia-friendly communication rating scale designed specifically for this study to closely align with the measures that could be obtained from the home recordings, an excerpt of which can be found in Fig. 1.

For pre-COVID participants, these assessments were administered in person as outlined in the testing manuals. For post-COVID participants, slight variations were made to the assessments to make administration possible over video conference. These variations included providing the DCT audio files for the participant to play on their own device to avoid audio feedback and numbering the picture array for the naming portion of the QAB so that participants could indicate their selection with a number instead of physically pointing. See Table 2 for assessment data of participants.

HOME RECORDING
The Language Environment Analysis (LENA) system was used to collect naturalistic audio recordings from the participants' home environment. Though this system was originally designed for language acquisition studies in children, it has also proved useful in data collection with adult populations.[25–27] This tool is composed of a small audio recorder (digital language processor [DLP]) designed to be slipped into a clothing pocket and accompanying software that analyzes the recording for metrics such as total adult word count, number of conversational turns, and a representation of the auditory environment in terms of the proportions of meaningful speech, distant speech, noise, and TV/electronic noise. In the current study, this software was used to isolate small sections of naturalistic conversational discourse

In one-on-one conversations, at home, how much do you speak in comparison to the other person?
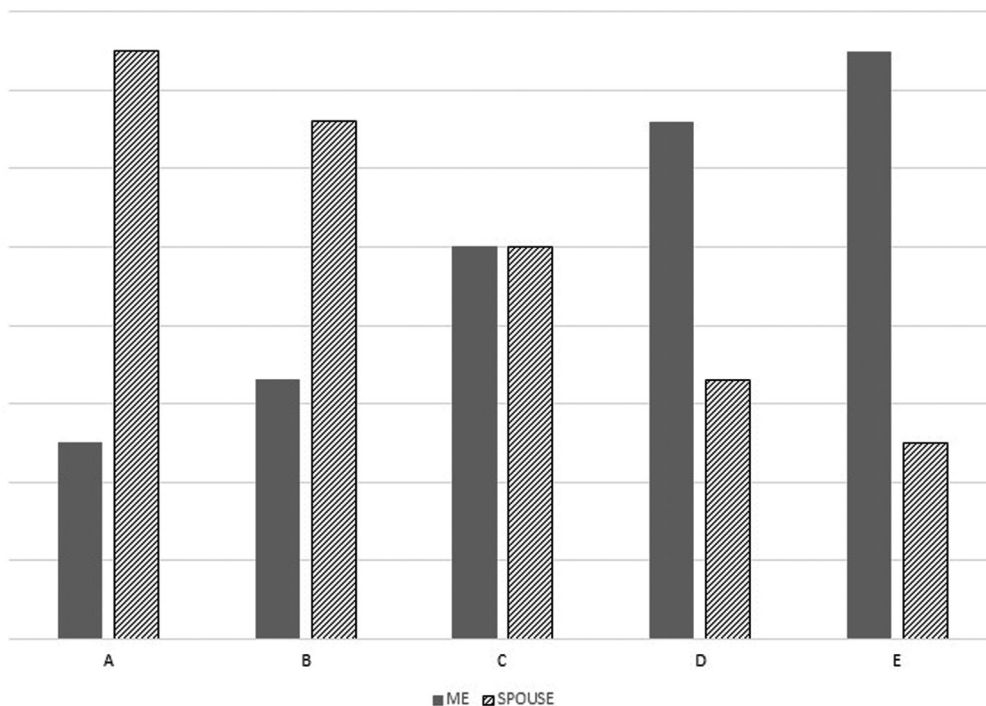
- A. I talk significantly less – (1/4)
- B. I talk moderately less – (1/2)
- C. We talk about the same amount – (50-50)
- D. I talk moderately more – (x2)
- E. I talk significantly more – (x4)

Figure 1 Excerpt of the communication rating scale.

from multi-hour recordings in order for us to analyze it manually.

After completing the preliminary testing, participants and communication partners were instructed in the use of the LENA DLP and were asked to demonstrate the use of the DLP to ensure proficiency. Participants were asked to take it home and set it to record for the length of their time awake over the course of 2 days. For post-COVID participants, instruction was

conducted via video conference and DLPs were mailed to their home address.

Data Analysis

DATA PROCESSING

When the DLPs were returned, the recordings were transferred to the LENA software and LENA's Advanced Data Extractor (ADEX)

Table 2 Participant Assessment Data

Participant	QAB score	DCT score	SAQOL mean score	QOLS mean score
PWA01	8.21	35	4.33	N/A
PWA02	8.77	31	3.82	N/A
PWA03	5.97	29	3.56	N/A
PWA04	4.99	19	4.19	N/A
NBI01	9.55	33	N/A	4.56
NBI02	10.00	38	N/A	5.81
NBI03	9.92	39	N/A	6.50
NBI04	9.86	38	N/A	6.88

Abbreviations: DCT, Discourse Comprehension Test; NBI, non–brain-injured; PWA, people with aphasia; QAB, Quick Aphasia Battery; QOLS, Quality of Life Scale, reported as the mean score out of 7; SAQOL, Stroke and Aphasia Quality of Life Scale, reported as the mean score out of 5.

was used to isolate 30-second intervals of speech in descending order of maximum verbal output. ADEX can be used to isolate variables, types of interactions, and particular intervals of time on a more specific level than can be performed in the LENA software itself, and thus was highly useful in separating these short sections of maximum speech. This approach was based on the procedure used by Ramírez-Esparza et al, which was demonstrated to be effective in processing data from children and adults recorded using the LENA system.[28] Sections of speech that included conversation between at least two speakers were selected in descending order until a total time of 60 minutes of conversation was identified.

TRANSCRIPTION
Identified sections were manually transcribed verbatim by three trained research assistants. If conversation extended beyond the 30 seconds of an identified section, transcription was extended until the conversation's natural end.

QUANTITATIVE ANALYSIS
Though the LENA system produces automated counts of adult words and conversational turns from recordings, it has not yet been demonstrated that this automatic analysis accurately reflects the language of PWA. To maximize data validity, LENA was used only to identify sections of naturalistic speech, and each transcribed section was manually analyzed by the first author for the total number of produced

words and total number of conversational turns for each speaker. These multiple measures were used to account for PWA's possible different presentations of communication quantity based on aphasia subtype. Words were identified using the rules laid out by Nicholas and Brookshire.[29] We defined conversational turns as occasions where one conversational partner spoke after another in a way that contributed to the content of the conversation; short interjections indicating affirmation or understanding (e.g., the insertion of "yeah" while the other partner is telling a story) were not counted as conversational turns. Introduction of a new idea by either speaker, regardless of who had spoken last, was also counted as a conversational turn. Conversational turns were counted only in sections in which both the participant and their primary communication partner produced at least one turn. To increase consistency across sections, conversational turn data were taken only for the participant and their primary communication partner; any conversational turns attributed to other speakers were not included in the data.

The obtained quantitative data were compared with participants' rating scale responses. Each measure was converted to a proportion of the participant's contribution in relation to the contribution of their communication partner. These objective proportions were then compared with the corresponding estimates provided in each participant's rating scale responses on the level of percent difference. This allowed us to investigate the accuracy of participants'

Table 3 Average Percentage of Words Produced by Participants with Aphasia Compared with Primary Communication Partner

Participant	Estimated proportion of conversational contribution[a]	Actual proportion of words produced[b]	%Difference in words produced[c]
PWA01	66%	68.58%	−02.58%
PWA02	33%	38.07%	−05.07%
PWA03	33%	47.86%	−14.86%
PWA04	50%	17.76%	+32.23%

[a]Estimated proportion of conversational contribution refers to the estimate made by the participant with aphasia wearing the recording device.
[b]Actual proportion of words produced refers to manual calculation of words as captured by the recording device.
[c]%Difference refers to the percent difference between the two aforementioned metrics.

perceptions of their language use and measures of conversational contribution. To understand the magnitude of difference between participants' rating scale responses and their production in actual conversation, the proportion of the participant's production as compared with their conversation partner was subtracted from their original estimate. As such, a negative percent difference corresponds with an individual underestimating their conversational contribution, while a positive percent difference comes from an overestimation.

The proportion of the participant's production of words in conversation was reported in two ways. The first was the participant's average percentage of words, which refers to the result of taking the average of the participant's proportional contribution in each individual section. The second was percent of total words, which refers to the proportion of the total words spoken by the participant across all sections when compared with the total words spoken by all speakers across all sections. Data for proportion of conversational turns were reported in similar ways, with the participant's average percentage of conversational turns and their percentage of total conversational turns.

RESULTS

Comparisons of the relative language produced by each participant and their conversation partner are reported in Tables 345 to 6. The quantitative analysis protocol outlined earlier was used to obtain the data both in terms of words (see Tables 3 and 4) and in conversational turns (see Tables 5 and 6,), though these results are reported separately for clarity.

Trends Across Groups

Two PWA (PWA01 and PWA02) estimated their production in conversations with their spouse quite reliably in proportion of words produced, but these participants' estimates were less accurate when compared in terms of conversational turns. The third PWA (PWA03)

Table 4 Average Percentage of Words Produced by Non–Brain-Injured (NBI) Participants Compared with Primary Communication Partner

Participant	Estimated proportion of conversational contribution[a]	Actual proportion of words produced[b]	%Difference in words produced[c]
NBI01	50%	32.43%	+17.57%
NBI02	50%	67.96%	−17.96%
NBI03	50%	37.16%	+12.84%
NBI04	25%	65.80%	−40.80%

[a]Estimated proportion of conversational contribution refers to the estimate made by the non–brain-injured participant wearing the recording device.
[b]Actual proportion of words produced refers to manual calculation of words as captured by the recording device.
[c]%Difference refers to the percent difference between the two aforementioned metrics.

Table 5   Average Percentage of Conversational Turns Produced by Participants with Aphasia Compared with Primary Communication Partner

Participant	Estimated proportion of conversational contribution[a]	Actual proportion of conversational turns[b]	%Difference in conversational turns[c]
PWA01	66%	54.88%	+11.11%
PWA02	33%	48.17%	−15.17%
PWA03	33%	51.94%	−18.94%
PWA04	50%	37.02%	+12.98%

[a]Estimated proportion of conversational contribution refers to the estimate made by the participant with aphasia wearing the recording device.
[b]Actual proportion of conversational turns refers to manual calculation of conversational turns as captured by the recording device.
[c]%Difference refers to the percent difference between the two aforementioned metrics.

Table 6   Average Percentage of Conversational Turns Produced by Non–Brain-Injured Participants Compared with Primary Communication Partner

Participant	Estimated proportion of conversational contribution[a]	Actual proportion of conversational turns[b]	%Difference in conversational turns[c]
NBI01	50%	46.10%	+03.90%
NBI02	50%	51.75%	−01.75%
NBI03	50%	48.21%	+01.79%
NBI04	25%	57.81%	−32.81%

[a]Estimated proportion of conversational contribution refers to the estimate made by the non–brain-injured participant wearing the recording device.
[b]Actual proportion of conversational turns refers to manual calculation of conversational turns as captured by the recording device.
[c]%Difference refers to the percent difference between the two aforementioned metrics.

showed a consistently greater underestimation of his conversational contribution both when measured in terms of words and conversational turns. On the other hand, the fourth PWA (PWA04) had rather large overestimations of his conversational contribution in both metrics.

The data collected on the words produced by NBI participants were inconsistent with their estimations in all cases; the percent difference in results showed both underestimation and overestimation by NBI when compared with recorded data. In contrast, when the estimates of NBI participants were compared with their proportion of conversational turns, higher rates of accuracy were observed, with the data of three NBI participants lining up very closely with their rating scale responses. However, the fourth NBI participant (NBI04) showed a particularly inaccurate estimate even in conversational turns, with a considerable underestimation of his contribution across both metrics. Of note, though the accuracy of their

estimates on conversational contribution was higher than that of the participants with aphasia when measured by conversational turns, the overall accuracy of NBI participants in estimating their verbal output in terms of proportion of words was in approximately the same range as participants with moderate to severe aphasia and was in fact notably lower than that of participants with mild aphasia.

Trends Across Metrics

A greater variance was seen in the proportion of words produced across participants than in the proportion of conversational turns. While the proportion of words in conversation spoken by participants ranged dramatically, the proportion of conversational turns produced by all participants were clustered more closely around 50%. Overall, NBI participants showed greater accuracy in their estimated proportion of conversational contribution; three NBI participants

demonstrated consistency between their estimations and the calculated results within a few percentage points. However, one NBI participant (NBI04) underestimated his contribution by a considerable amount, which marks a notable difference from the high rates of agreement seen in the other NBI participants. For PWA, the accuracy of estimates of their contribution as compared with their recorded conversational turns included both underestimations and overestimations, with no estimates aligning particularly closely with recorded data. However, two PWA (PWA01 and PWA02) had estimates that very closely aligned with the recorded data in terms of the proportion of words they produced.

DISCUSSION

In this study, we explore a novel perspective on the reliability of language production rating scales, the range of accuracy in self-perception of language use among both PWA and NBI participants, and the potential utility of examining natural communication trends in the everyday lives of PWA.

Research Question 1: In terms of total words produced and number of conversational turns, how accurate are PWA's perceptions of their own language use?

There was a significant range in the accuracy of PWA's self-perceptions of their language output and conversational contribution. In some cases, these estimates were highly accurate, but other participants significantly underestimated or overestimated their production to such a degree that their rating scale responses revealed little about their actual recorded communication. This was the case across both metrics of language use. The participants who most significantly underestimated or overestimated their production tended to have these inaccurate estimates across the measures of both words and conversational turns. However, even among those whose estimates of their language proved quite accurate when measured by one metric, proportion of words produced, the same estimate was shown in many cases to be inaccurate when language output was measured by a different metric, proportion of conversational turns. This high-

ly varied range of accuracy across PWA suggests that rating scales based on self-perception, despite their clinical pervasiveness, have no guarantee of reflecting verbal language use or communicative environments accurately.

However, higher rates of agreement between estimations and retrieved objective data were found in the participants with more mild aphasia. This may be associated with generally higher levels of error awareness, as greater insight into their communicative behaviors might enable them to more accurately gauge their proportional contributions to conversations. Previous research has shown that people with mild aphasia tend to perform better in a range of executive functioning tasks[30] and have greater insight into deficits[31]; we theorize that this higher level of insight and self-monitoring may be related to the more accurate self-perceptions seen in this study by people with mild aphasia. Further research should be conducted to determine whether the reliability of self-perception in people with mild aphasia is high enough to justify the use of clinical rating scales in this population.

Research Question 2: Is this accuracy notably different from an equivalent comparison with NBI individuals?

Both PWA and NBI participants showed a great deal of variability in the accuracy of their estimates. Some PWA demonstrated a good deal of accuracy depending on the metric used to measure their conversational contribution; however, each one also had some results that were drastically different from their estimates. The inaccuracies of these perceptions involved both underestimation and overestimation across the range of participants. The lack of a consistent trend suggests that the inaccuracy of self-perception of language is not predictable. In addition, none of the NBI participants accurately estimated their verbal output when this output was measured by proportion of words. Both considerable underestimations and overestimations were noted across the range of participants, supporting the previous conclusion regarding the lack of predictability in errors of self-perception.

NBI participants did show a higher rate of accuracy overall in estimating their

conversational contribution when this was measured by proportion of conversational turns. Under these circumstances, three NBI participants had estimates that aligned very closely with their recorded data (within <4 percentage points), while there was a much larger range of disparity among the PWA in their estimates and their proportion of conversational turns. Furthermore, the high rate of agreement between the estimates and conversational turn data of NBI participants was not seen when their data were measured in proportion of words.

Bearing these mixed results in mind, it is perhaps worthwhile to evaluate the ways in which the two metrics used in this study might quantify different aspects of communicative contribution. By our definitions, any response that contributed even a small degree of meaningful content to the conversation was counted as a conversational turn. As such, a response of three words and a response of three sentences could each be counted as one conversational turn, even though the amount of meaningful content they contributed to the conversation might differ dramatically. This indicates that it might reasonably be said that proportion of words could better quantify the amount of content contributed to the conversation, while proportion of conversational turns could reflect the overall continued engagement of the participant in conversation.

We believe this may be part of the reason that the proportion of conversational turns across participants had a smaller degree of disparity and were clustered more closely around 50%, while a considerably wider range was seen for the proportion of words produced by each participant. If this is the case, it might suggest that the tendency for NBI participants to estimate their conversational contribution around equal with that of their partner is a possible reason for the marked increase in accuracy when using this metric. To this effect, it is also worth noting that the only NBI participant whose estimate of his conversational contribution was notably inaccurate when measured in terms of conversational turns (NBI04) was also the only NBI participant to estimate his contribution as something other than 50%.

Furthermore, this leads to the question of whether PWA and people without brain injury actually think about conversational contribution in different ways. If we assume that proportion of words spoken is more closely tied to contribution of content and that proportion of conversational turns is more related to engagement and participation in conversation, these results might indicate that people without aphasia are more likely to think about the dynamic of their conversations with a partner by comparing their respective engagement, while PWA are more likely to consider the amount of content they are contributing in comparison to the communication partner. To our knowledge, little to no prior research has investigated whether perceptions of conversational contribution differ between people with and without aphasia, but the differences across groups and metrics observed in this study have implications that may justify further research into this topic.

While the varying accuracy among all participants calls into question the applicability of rating scales in any population, it is noteworthy that the estimates made by NBI participants were in approximately the same range of accuracy as those made by participants with moderate to severe aphasia when measured by proportion of words, and were in fact less accurate than those made by participants with mild aphasia using this metric. These results may reflect the possibility that PWA, especially those with mild aphasia, are more attentive to how their language deficits affect their contribution of content to their everyday conversations. Because communication requires a greater amount of concentration and cognitive resources for PWA,[32,33] we theorize that it is possible that they have a higher level of insight into their verbal contributions than do people without brain injury, who require fewer cognitive resources for effective communication and thus may not attend to conversational interactions to the same degree.

Clinical Implications

While information about the way PWA use language in their home environments can have great clinical utility, the preliminary results of

this study lead to questions as to whether the self-perceptions of PWA as reflected by rating scales accurately represent this information. However, the possibility that increased levels of insight into deficits and overall language use might be tied to increased reliability of self-perception of language use in the home environment holds clinical implications of its own, particularly regarding the inclusion of insight-related goals in aphasia treatment as appropriate. The results of this study support the use of insight-related goals in treatment as a potential for increased validity in clinical information gathering as well as direct language gains.

These results also suggest that it may be more productive for SLPs to avoid the use of subjective measures and data subject to interpretation where possible, and perhaps instead to ask patients with aphasia more specific questions and introduce more specific functional goals and suggestions for home programming. For example, suggesting that a patient "talk more at home" or "practice talking with anyone you can" leaves room for a great deal of varied interpretation, and the way that a PWA might interpret and execute these tasks could be very different from the way the clinician intended depending on their self-perception of language use in other settings. More specific functional targets, such as "give your spouse a compliment today" or "practice ordering coffee three times with your daughter before ordering for yourself at a coffee shop," limit the variability of interpretation that may be influenced by variations in self-perception of language use.

Limitations

This preliminary data provide insight but remain imperfect. The logistics of using data from participants' natural environments presented some atypical challenges and limitations, which are detailed in the following paragraphs.

From an applicability perspective, the rating scale developed for this study was specifically designed to target measures of conversation that could be easily identified by the LENA technology used. In this sense, though it is based on many rating scales in clinical use, the results cannot be compared with those that might be retrieved using the CETI or other

common rating scales. Though this study provides valuable foundational investigation, were a replication to be performed using validated clinical rating scales, the results would be more applicable to direct clinical use.

Furthermore, in an attempt to make the questions on the rating scale as aphasia-friendly and intuitive as possible, participants were asked to give ranges or approximate which of several benchmark proportions was most reflective of their communication. These estimates were appropriate for comparison within the scope of this study, but it should be noted that the participants' estimates were not as exact as they might be if the questions had been asked in a more open-ended fashion.

It is also worth noting that the sample size of this study was relatively small, consisting of four PWA dyads and four NBI dyads. Given the breadth of different communication styles and dynamics that may be present between any given person and their primary communication partner, there is no guarantee that this sample is representative of a larger population, which limits the ability to generalize findings of this study to overall clinical practice. Additionally, since all PWA were recruited from a university aphasia group, potential sampling bias may further complicate the generalizability of findings due to the potential presence of confounding variables (e.g., PWA who regularly attend an outpatient aphasia group could be more motivated to improve their language than those who do not, which might also be reflected in language productivity at home).

Finally, the need to modify the study procedure for the last two NBI participants due to the emergence of the COVID-19 virus placed some limitations on the consistency of our methods of conducting the study across all participants. While all attempts were made to keep the modified procedure as consistent with the original as possible, it is important to recognize that the results associated with these two participants might still have been affected by these adaptations.

CONCLUSION

Though rating scales are frequently used in clinical assessment of PWA, there are

significant concerns about the reliability and validity of results obtained from these subjective measures. In the present pilot study, it was found, upon comparing rating scale responses to objective data retrieved from the everyday conversations of PWA, the accuracy of their perceptions of their verbal language output was highly variable. Notably, the rating scale responses of the two adults with mild aphasia tended to be more accurate with regard to the proportion of words spoken than those with moderate to severe aphasia or those without brain injury; further research should be conducted to determine if rating scales may have more clinical utility when administered to people with mild aphasia. Furthermore, an equivalent comparison with NBI participants revealed that this lack of consistent accuracy was also the case for people without communication disorders. These preliminary results call into question the reliability of subjective rating scales in quantifying verbal communicative output, and further research is needed to evaluate whether their use in clinical settings may or may not produce an accurate representation of PWA's everyday language use. In the meantime, speech-language pathologists may be better off not relying on subjective measures and instead asking their clients with aphasia more specific questions based on functional and achievable goals.

REFERENCES

1. Meinzer M, Djundja D, Barthel G, Elbert T, Rockstroh B. Long-term stability of improved language functions in chronic aphasia after constraint-induced aphasia therapy. Stroke 2005; 36(07):1462–1466

2. Crockett DJ. A comparison of empirically derived groups of aphasic patients on the Neurosensory Center Comprehensive Examination for Aphasia. J Clin Psychol 1977;33(1, S1)194–198

3. Hilari K, Wiggins R, Roy P, Byng S, Smith S. Predictors of health-related quality of life (HRQL) in people with chronic aphasia. Aphasiology 2003; 17(04):365–381

4. Lomas J, Pickard L, Bester S, Elbard H, Finlayson A, Zoghaib C. The communicative effectiveness index: development and psychometric evaluation of a functional communication measure for adult aphasia. J Speech Hear Disord 1989;54(01): 113–124

5. Frattali C, Thompson CK, Holland AL, Wohl CB, Ferketic M. Functional Assessment of Communication Skills for Adult. Rockville, MD: American Speech-Language-Hearing Association; 1995

6. Kertesz A. The Western Aphasia Battery. New York: Grune & Stratton; 1982

7. Spreen O, Risser AH. Assessment of Aphasia. New York, NY: Oxford University Press; 2003

8. Prins R, Bastiaanse R. Analyzing the spontaneous speech of aphasic speakers. Aphasiology 2004;18 (12):1075–1091

9. Paul DR, Frattali CM, Holland AL, Thompson CK, Caperton CJ, Slater SC. ASHA Quality of Communication Life Scale (QCL). Rockville, MD: American Speech-Language-Hearing Association; 2003

10. Cherney LR, Babbitt EM, Semik P, Heinemann AW. Psychometric properties of the communication Confidence Rating Scale for Aphasia (CCRSA): phase 1. Top Stroke Rehabil 2011;18 (04):352–360

11. Grande M, Hussmann K, Bay Eet al.. Basic parameters of spontaneous speech as a sensitive method for measuring change during the course of aphasia. Int J Lang Commun Disord 2008;43(04): 408–426

12. Fox S, Armstrong E, Boles L. Conversational treatment in mild aphasia: a case study. Aphasiology 2009;23(7–8):951–964

13. Jacobs BJ. Social validity of changes in informativeness and efficiency of aphasic discourse

following linguistic specific treatment (LST). Brain Lang 2001;78(01):115–127

14. Raymer AM, McHose B, Smith KG, Iman L, Ambrose A, Casselton C. Contrasting effects of errorless naming treatment and gestural facilitation for word retrieval in aphasia. Neuropsychol Rehabil 2012;22(02):235–266

15. Gillespie A, Murphy J, Place M. Divergences of perspective between people with aphasia and their family caregivers. Aphasiology 2010;24(12):1559–1575

16. Horowitz A, Goodman CR, Reinhardt JP. Congruence between disabled elders and their primary caregivers. Gerontologist 2004;44(04):532–542

17. Walters K, Iliffe S, Tai SS, Orrell M. Assessing needs from patient, carer and professional perspectives: the Camberwell Assessment of need for Elderly people in primary care. Age Ageing 2000;29(06):505–510

18. Williamson DS, Richman M, Redmond SC. Applying the correlation between aphasia severity and quality of life measures to a life participation approach to aphasia. Top Stroke Rehabil 2011;18(02):101–105

19. Hilari K, Owen S, Farrelly SJ. Proxy and self-report agreement on the stroke and aphasia quality of life scale-39. J Neurol Neurosurg Psychiatry 2007;78(10):1072–1075

20. Crockford C, Lesser R. Assessing functional communication in aphasia: clinical utility and time demands of three methods. Eur J Disord Commun 1994;29(02):165–182

21. Wilson SM, Eriksson DK, Schneck SM, Lucanie JM. A quick aphasia battery for efficient, reliable, and multidimensional assessment of language function. PLoS One 2018;13(02):e0192773

22. Brookshire RH, Nicholas LE. Discourse Comprehension Test. Communication Skill Builders. 1993

23. Hilari K, Byng S, Lamping DL, Smith SC. Stroke and Aphasia Quality of Life Scale-39 (SAQOL-39): evaluation of acceptability, reliability, and validity. Stroke 2003;34(08):1944–1950

24. Burckhardt CS, Anderson KL. The Quality of Life Scale (QOLS): reliability, validity, and utilization. Health Qual Life Outcomes 2003;1(01):60

25. Li L, Vikani AR, Harris GC, Lin FR. Feasibility study to quantify the auditory and social environment of older adults using a digital language processor. Otol Neurotol 2014;35(08):1301–1305

26. Wang Z, Pan X, Miller KF, Cortina KS. Automatic classification of activities in classroom discourse. Comput Educ 2014;78:115–123

27. Wang Z, Miller K, Cortina K. Using the LENA in teacher training: promoting student involvement through automated feedback. Unterrichtswissenschaft. 2013;4:290–305

28. Ramírez-Esparza N, García-Sierra A, Kuhl PK. Look who's talking: speech style and social context in language input to infants are linked to concurrent and future speech development. Dev Sci 2014;17(06):880–891

29. Nicholas LE, Brookshire RH. A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. J Speech Hear Res 1993;36(02):338–350

30. Fridriksson J, Nettles C, Davis M, Morrow L, Montgomery A. Functional communication and executive function in aphasia. Clin Linguist Phon 2006;20(06):401–410

31. van der Stelt CM, Fama ME, Mccall JD, Snider SF, Turkeltaub PE. Intellectual awareness of naming abilities in people with chronic post-stroke aphasia. Neuropsychologia 2021;160:107961

32. Fischler I. Attention, resource allocation, and language. . In: Aphasia and Language: Theory to Practice. The Science and Practice of Neuropsychology. New York, NY: Guilford Press; 2000:348–371

33. Murray LL. Review attention and aphasia: theory, research and clinical implications. Aphasiology 1999;13(02):91–111