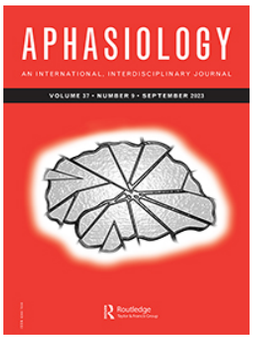

Aphasiology

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

Differences in Connected Speech Outcomes Across Elicitation Methods

Tatiana T. Schnur & Sharon Wang

To cite this article: Tatiana T. Schnur & Sharon Wang (2023): Differences in Connected Speech Outcomes Across Elicitation Methods, Aphasiology, DOI: 10.1080/02687038.2023.2239509

To link to this article: https://doi.org/10.1080/02687038.2023.2239509

Published online: 12 Aug 2023.

APHASIOLOGY
https://doi.org/10.1080/02687038.2023.2239509

Routledge
Taylor & Francis Group

Differences in Connected Speech Outcomes Across Elicitation Methods

Tatiana T. Schnur[a,b] and Sharon Wang[a]

[a]Department of Neurosurgery, Baylor College of Medicine, Houston, Texas; [b]Department of Neuroscience, Baylor College of Medicine, Houston, Texas

ABSTRACT

Background: Connected speech is often used to assess many aspects of an individual's language abilities after stroke. However, it is unknown the degree to which elicitation methods differ in generating structural and syntactic aspects of connected speech, two critical components of successful communication. Quantifying the degree to which elicitation methods differ in eliciting structurally, syntactically, and lexically complex connected speech at the earliest stage of stroke before reorganization and rehabilitation of function independent of clinical diagnosis of aphasia has not been examined to date. Addressing this gap has implications for early clinical intervention as well as empirical studies of connected speech production.

Aims: We compared two common elicitation methods, picture description and storytelling on lexical, structural, and syntactic measures of connected speech in speakers during the acute stage of left hemisphere stroke.

Methods & Procedures: We measured connected speech using an automated quantitative production analysis approach (Fromm et al., 2021) in 71 native-English speaking participants (27 female; 59 ± 13 years) within an average 3.9 days from left hemisphere stroke onset. We tested the degree of agreement and consistency between elicitation methods for lexical, structural, and syntactic measures of connected speech, as well as the degree of concordance in classifying deficits across individuals.

Outcomes & Results: Storytelling elicited significantly more words and more structurally complex, lexically diverse, and syntactically accurate speech in comparison to picture description. Elicitation methods differed in measuring outcomes across participants for the lexical and syntactic, but not structural complexity aspects of connected speech where storytelling classified more participants with impairments in comparison to picture description.

Conclusions: These differences suggest storytelling provides assessment of connected speech abilities more reflective of real-world abilities where its use is particularly critical for examining individual differences and providing diagnoses of acute stroke language deficits. As a result, using storytelling as a connected

ARTICLE HISTORY
Received 26 May 2022
Revised 11 Jul 2023
Accepted 17 Jul 2023

KEYWORDS
acute stroke; connected speech; language; syntax; methods

CONTACT Tatiana T. Schnur ttschnur@gmail.com Department of Neurosurgery, Baylor College of Medicine, 1 Baylor Plaza, Houston, TX, 77030

speech elicitation method more effectively captures a patient's language capabilities after stroke, consequently informing clinical diagnosis and treatment.

Introduction

Approximately 30% of adults who survive a left hemisphere stroke experience language impairments that can affect retrieval of words, their meanings, their associated grammatical information, and their sounds (Flowers et al., 2013; Grönberg et al., 2022). In comparison to assessing speech with picture naming, connected speech assessments like picture description and storytelling are useful for capturing many aspects of an individual's language abilities including accessing and organizing words into more structurally and syntactically complex combinations. These language abilities are critical for communication during connected speech. Quantifying the degree to which language abilities during connected speech are impaired independent of clinical diagnosis of aphasia at the earliest stage of stroke before reorganization and rehabilitation of function has implications for early clinical intervention and empirical studies of connected speech production. This is because connected speech impairments may be differentially observed depending on the elicitation task. To our knowledge, differences in the complexity of language elicited during descriptive and narrative discourse tasks after stroke have not been examined. Although insight comes from investigations in other clinical populations including mild cognitive impairment and neurodegenerative disease (cf. Clark et al., 2021; Lavoie et al., 2021), it is unclear which elicitation method is a more sensitive measurement of connected speech in speakers after stroke. Addressing this gap is the primary focus of this paper.

Assessments of language deficits after left hemisphere stroke as well as other clinical populations including neurodegenerative disease generally involve eliciting spontaneous speech from speakers viewing scenes which are continuously available during description or to a lesser extent producing spontaneous speech via storytelling or in response to autobiographical questions (Bryant et al., 2016; Bryant et al., 2017; Filiou et al., 2020). The most commonly presented scenes include a picture of the picnic scene used in the Western Aphasia Battery (WAB; Kertesz, 2007) and a picture of the cookie theft scene used in the Boston Diagnostic Aphasia Examination (BDAE; Goodglass et al., 1983) and the National Institutes of Health Stroke Scale (Brott et al. 1989; Meyer & Lyden, 2009). In the storytelling task, speakers elaborate on a well-known story, such as the Cinderella fairy tale, after viewing a wordless picture book to remind them of the story (Rochon et al., 2000; Saffran et al., 1989). In these methods, speakers are encouraged to produce full sentences if possible. Because picture description requires fewer materials and less elicited content in comparison to storytelling (Alyahya et al., 2020; Ash et al., 2013; Bose et al., 2022; cf. Stark, 2019), it is easier to administer and relatively easier to score (Alyahya et al., 2020; Ash et al., 2013; Bose et al., 2022; Clarke et al., 2021; Lavoie et al., 2021). Critically, structural deficits like producing less elaborated sentences with fewer embeddings and syntactic deficits like producing fewer syntactically well-formed sentences and grammatical words are hallmarks of clinical characterizations of impaired speech in

speakers with stroke (Berndt et al., 1997; Gordon, 2006; Matchin et al., 2020). Thus, it is important to assess connected speech using an elicitation method which allows for rich structural and syntactic output to better diagnose language impairments.

In clinical populations with language deficits as a result of neurodegenerative disease and mild cognitive impairment, picture description has been compared to storytelling yielding mixed results concerning which method elicits richer structural and syntactic output. When comparing picture description with sequential picture storytelling in speakers with either frontal-temporal dementia or different types of primary progressive aphasia (n=32), structural and syntactic complexity did not differ as measured by the mean length of utterance, the number of dependent clauses per utterance and grammatically well-formed sentences (Ash et al. 2013). Other studies provide mixed results. To assess connected speech in Bengali speakers with probable Alzheimer's Disease (AD; n=6) in comparison to controls (n=8), Bose and colleagues (2022) used Quantitative Production Analysis (QPA; Saffran et al., 1989; Rochon et al., 2000; Gordon, 2006; Wilson et al., 2010). The QPA procedure provides a detailed approach to analyze connected speech in an objective way focusing on the frequency of a number of features of connected speech such as the degree to which noun and verb phrases are elaborated, how often sentences are well-formed or how often determiners are used when required. Here, sequential picture storytelling vs. picture description elicited greater group differences across multiple syntactic and structural measures. Speakers with AD produced fewer closed-class words, decreased quality of noun inflections, and shorter sentences with fewer embeddings in comparison to controls. In another study using QPA but comparing storytelling without visual stimuli to picture description in a group of 13 speakers with logopenic variant of primary progressive aphasia (lvPPA) and 13 unimpaired speakers, storytelling elicited greater differences between groups in the proportion of well-formed sentences produced (Lavoie et al., 2021). Using multiple automated approaches to quantify connected speech and a machine learning analytic approach, storytelling without visual stimuli outperformed picture description in predicting whether a speaker was diagnosed with mild cognitive impairment (n=12) and/or mild Alzheimer's disease (n=12) in comparison to controls (n=25). However, the differences were mostly in lexical aspects including phonological complexity of words (the number of syllables per word), macrolinguistic discourse aspects like the density of ideas in the narrative (the proportion of content information units to overall words produced) and semantic coherence of the narrative (Clarke et al., 2021). Although the results are suggestive that storytelling allows better opportunity to produce structurally and syntactically complex connected speech, the mixed neurodegenerative etiology within and across participant samples, modest sample sizes, and lack of evidence from stroke populations leave it unclear whether these results extend to those with language deficits after stroke.

In speakers after stroke, comparisons between picture description and storytelling elicitation have exclusively assessed lexical as opposed to structural-syntactic characteristics. In speakers with chronic aphasia after stroke (n=46), sequential picture storytelling elicited a larger number and diversity of words in comparison to single picture description perhaps unsurprisingly given the increased visual depictions of events (Alyahya et al. 2020; Alyahya et al., 2021; cf. Wright & Capilouto, 2009 for similar increases in number and diversity of words in sequential picture storytelling vs. picture description in unimpaired speakers). In a direct comparison of storytelling without visual stimuli and picture

description, both speakers with chronic aphasia (n=90) and unimpaired speakers (n=84) produced more verbs, adjectives and adverbs out of the total number of words during storytelling but only unimpaired speakers produced more words during storytelling (Stark, 2019). In support of using either elicitation method to measure connected speech in stroke, speakers with chronic aphasia did not differ in their elicited total number of words, open vs. closed class words, or nouns vs. verbs across methods. Although these studies did not analyze the structural and syntactic aspects of connected speech, in comparison to answers from quality of life questions, picture description was better at eliciting increased numbers of predicates and arguments (26 speakers with chronic aphasia, Dipper et al., 2018). However, to our knowledge, the degree to which spontaneous storytelling and picture description differ in elicitation of structural and syntactic elements has not been explored in speakers with stroke.

Current Study

Our goal was to measure how connected speech elicitation methods compared in eliciting structurally and syntactically, as well as lexically rich speech while quantifying degree of impairment in speakers after acute stroke. This study offers advantages over previous work in several ways. First, we compared connected speech elicited during picture description and spontaneous storytelling without visual stimuli from a large cohort of speakers during the acute phase of a left hemisphere stroke with varying degrees of language impairment (n = 71). By recruiting participants independent of a clinical diagnosis of aphasia, we included a wide variety of unimpaired to impaired speech samples, providing a more expansive speech sample and one more representative of those clinically assessed acutely for deficits. Second, we used an automated version of the QPA approach (Fromm et al., 2021) to quantify microlinguistic connected speech deficits (Rochon et al., 2000; Saffran et al., 1989). The QPA provides an objective and rigorous analysis of different measures of the content, structure, and syntax of connected speech in comparison to often used categorical, subjective measures (e.g., assessments of fluency on a 0-10 scale in the BDAE) and has been used to measure connected speech across different clinical populations (e.g., Gordon et al., 2020; Mirman et al., 2019; Lavoie et al., 2021; Wilson et al., 2010). Using the validated automated QPA (C-QPA) approach affords the advantages of less sensitivity to bias and less inconsistent evaluation (Fromm et al., 2021). Lastly, we provide the first assessment in participants with stroke of whether elicitation tasks differ in the number of participants categorized with lexical-syntactic deficits relative to controls which clarifies the degree to which assessment tasks are biased. Results will inform researchers and clinicians to facilitate better assessment of lexical, structural, and syntactic deficits in connected speech after stroke.

Methods

Participants

In this retrospective study, 71 participants with stroke recruited from three comprehensive stroke centers in the Texas Medical Center, Houston, Texas participated in this study (44 males; 60 right-handed; 66 ischemic stroke; 3 hemorrhagic stroke; 2 undetermined;

age: M= 59.4 years; S.D.= 13.5; range= 20-83; education: M= 14.0 years; S.D.= 3.0; range= 6-23). All were native monolingual English speakers who experienced a left hemisphere acute stroke within a mean of 3.9 days of behavioral testing (SD= 2.5 days, range=1-13 days) without other health conditions which could impact cognition (i.e., tumor, dementia, alcohol and/or drug dependency). For this study, further inclusion criteria included if participants completed behavioral testing for both picture description and storytelling tasks and their speech was intelligible enough to be transcribed. For the control group, we enrolled 12 non-brain damaged participants (two male; 10 right-handed) matched in mean age (M = 56.8, SD = 13.9, range = 37-78) and education (M = 15.4, SD = 1.9, range = 12-18; *t*'s < 1.58, *p*-values > 0.11) to the 71 participants with stroke. Informed consent was approved by the Baylor College of Medicine Institutional Review Board.

We assessed participants for the degree of apraxia of speech to limit its contribution to the assessment of connected speech measures. Sixty-one participants and all controls completed subtest 5 of the Second Edition of the Apraxia Battery for Adults (Dabul, 2000). We were unable to administer the Apraxia Battery for ten participants because of in-patient clinical care scheduling conflicts. For these participants, we analyzed speech errors to assess for the presence of apraxia of speech using speech samples from picture description and storytelling (see below for task details). A speech language pathologist used a four-point scale to score apraxia of speech: 1 no apraxia (no speech sound errors), 2 mild apraxia (speech errors on < 25% of words), 3 moderate apraxia (speech errors on 25-50% of words), and 4 severe apraxia (speech errors on > 50% of words). Sixty-four participants with stroke and all control participants were confirmed without apraxia of speech. The remaining seven had the following scores: 2 (n=4) and 3 (n = 3).

We did not clinically assess for the presence of aphasia as this study was part of a larger study which included a longer battery of carefully designed tests of specific cognitive abilities where correlational analyses required a range of abilities (Ding et al., 2020; Ding & Schnur, 2022; Fromm et al., 2020; Martin et al., 2021; Martin & Schnur, 2019; Schnur & Lei, 2022; Zahn et al., 2023). We did not incidentally administer an aphasia testing battery because testing time was restricted given the acute care setting. Aphasia diagnoses as part of clinical care were not available at the time of testing as testing occurred close to hospital admission.

Procedure

Participants were tested in the hospital or in part at the participant's home if the participant was discharged before testing was completed. We administered and audio recorded picture description followed by storytelling.

Tasks

Storytelling. Participants viewed a picture book of the Cinderella story (Jeffers et al., 2004) for as long as they liked with the words concealed. Participants were then instructed to close the book and retell the story of Cinderella. Experimenters encouraged participants to produce full sentences and continue speaking if they produced long hesitations mid-narrative.

Picture Description. Participants viewed the picnic scene from the WAB (Kertesz, 2007) and were instructed to describe what was happening in the picture. They were permitted to look at the picture while they were describing the picture. As was done for the storytelling task, participants were encouraged to produce full sentences and continue speaking if they produced long hesitations mid-description.

Transcription & Scoring

For picture description, S.W. followed transcription procedures from Martin and Schnur (2019, p. 63) based on the rules from the QPA (Gordon, 2006, p.189; Berndt et al., 2000, pp. 8-12; cf. Rochon et al., 2000, p. 198; Saffran et al., 1989, pp. 450-451) to identify words uttered, separate narrative words from other uttered words, and segment narrative words into utterances. For storytelling, we used transcriptions by two raters from Martin and Schnur which were transcribed using the same approach (also reported in Ding & Schnur, 2022).

Following procedures described by Fromm et al. (2021), S.W. morphologically parsed the picture description transcription to indicate utterances that were imperatives, missing required determiners, missing subjects, and not syntactically well-formed. We then extracted 12 different measures of lexical-syntactic accuracy and structural complexity using a semi-automated procedure of the QPA (C-QPA) via CLAN (Fromm et al., 2021; MacWhinney, 2000). For storytelling, we used the automated C-QPA results conducted by Fromm et al. (2021) on the transcriptions from Martin and Schnur (2019) and Ding and Schnur (2022). We refer to these variables grouped into three major aspects of connected speech based on principal component loadings reported by Ding, Martin, Hamilton, and Schnur (2020; cf. Gordon, 2006, 2020): lexical selection, structural complexity, and syntactic accuracy. See Table 1 for C-QPA variable definitions. For an example transcription, segmentation, and morphological parsing, see Appendix.

Inter-rater Reliability

For storytelling by participants and controls, inter-rater reliability for manual transcriptions and utterance segmentation was previously conducted by Martin and Schnur (2019) and further described by Ding and Schnur (2022).

S.W. completed all picture description transcriptions and coding. To establish inter-rater reliability between S.W. and the raters for storytelling on transcription accuracy and utterance segmentation. S.W. re-transcribed the storytelling tasks for a randomly selected sub-sample of five participants who produced more than 10 utterances on the storytelling task and who also completed the picture description task. S.W. reached 97% agreement with the raters from Martin and Schnur (2019) on the number of narrative words (range 93-100%) where an average 91% of narrative words were identical (range 79-98%). When segmenting narrative words into utterances, S.W. showed an average agreement of 93% (range 75%-100%) with Martin and Schnur[1] To establish intra-rater reliability for the picture description transcription and utterance segmentation, we randomly selected

[1]Fromm et al. (2021) demonstrated good agreement between manual scoring and automated C-QPA in all QPA measures except for auxiliary complexity..

Table 1. Connected speech C-QPA variable definitions and results comparing storytelling and picture description elicitation methods. An '*' indicates significant disagreement at the group level (t's > 3.86; p's < .0002), inconsistency (lack of significant correlation) across participants (r's < .12; p's > .18; note significant correlations indicated by '~', p's < .002), and significant differences classifying participant deficits (χ^2's > 8.60, p's < .0009) after controlling for multiple comparisons (alpha = 0.05/12 = 0.004), where more participants classified with deficits by assessment task are indicated in parentheses. '†' Indicates significant at the uncorrected for multiple comparisons p < .05 level. Imp. = Impaired.

Connected Speech C-QPA Variables	Definition	Participants	Disagreement (*p's < .0002)	Inconsistency (* p's > .18) Consistency (~ p's < .002)	Inconsistent Deficit Classification (* p's < .0009)
Lexical Selection					
Proportion pronouns	# pronouns/ (# nouns + pronouns)	70	*	*	* (# imp. story > pic)
Proportion verbs	# verbs/(# nouns + verbs)	70	*	*	* (# imp. story > pic)
Proportion closed-class words	# closed-class words/# narrative words	70	†	*	
Structural complexity					
Mean utterance length	# words in utterances/# utterances	70	*		* (# imp. pic > story)
Mean sentence length	# words in sentences/# sentences	68	*		
Sentence elaboration	Subject noun phrase + verb phrase elaboration where phrase elaboration = (number of open class words and pronouns in a phrase/number of phrases) − 1	69	*		
Embedding index	# embeddings/# sentences	70	*		* (# imp. story > pic)
Narrative words	# words directly contributing to narrative	70	*		
Syntactic accuracy					
Proportion well-formed sentences	# syntactically well-formed sentences/# sentences	71	*	*	
Proportion words in sentences	# words in sentences/# narrative words	70	*	*	* (# imp. pic > story)
Determiner index	# nouns requiring determiners, with determiners/# nouns requiring determiners	68	*	*	
Auxiliary complexity	(Auxiliary score/# matrix verbs) − 1	70	*	*	

five patients for S.W. to score twice, with the second timepoint ranging between one to six months after the first. S.W. reached 99% agreement on the number of narrative words (range 96-100%) where an average 96% of narrative words were identical (range 93-99%). When segmenting narrative words into utterances, S.W. showed an average agreement of 97% (range 92-100%).

Statistical Analysis

We tested whether 12 QPA measures of lexical-syntactic aspects of connected speech elicited using storytelling vs. picture description differed along three dimensions, correcting for multiple comparisons (alpha = 0.05/12 = 0.004). First, to assess the degree to which the two elicitation methods significantly differed from one another in measurements of connected speech, we compared the 12 C-QPA measures via paired *t*-test. Second, we calculated Pearson correlation coefficients to determine the degree to which the 12 C-QPA variable scores generated from the two elicitation methods were consistent across individuals. Finally, we conducted a chi-square test of independence for each C-QPA measure to investigate if the number of people classified as impaired via storytelling was significantly different than the number of people classified as impaired via picture description. We set the level of impairment as < -2 SDs from the control performance mean for all variables except for the proportion pronouns, proportion verbs, and proportion of closed-class words produced as these scores are bivalent. For these variables, we also classified scores as impaired at > 2 SDs from the control performance mean as they indicate problems producing primarily nouns in comparisons to closed-class words, pronouns and verbs. Negative values indicate the reverse (i.e. problems producing words with grammatical function; cf. Ding et al., 2020). We tested for outliers for each of the 12 C-QPA variables using the criteria of studentized residuals greater than 3 (extreme values) or a Cook's D greater than 1 (influential).

Results

Table 1 summarizes the results comparing C-QPA outcomes for picture description and storytelling elicitation methods averaged across tasks (via paired *t*-test), across participants (via Pearson *r* coefficients), and when identifying participants as impaired or spared in comparison to controls (via chi-square). Outlier analyses identified between 0 and 3 participants as outliers across C-QPA variables.

Agreement and Consistency Between Tasks

Task Agreement

At the group level, picture description and storytelling measurements were significantly different for 11 of 12 C-QPA lexical-syntactic connected speech variables. As seen in Figure 1, A, regarding lexical selection, picture description elicited smaller proportions of pronouns and verbs produced; pronouns: picture description, *M*= 0.2, *SD*= 0.1, range= 0-0.49; storytelling, *M*= 0.5, *SD*= 0.1, range= 0.13-0.75; *t*(69)= 18.27, *p*= <.0001; verbs: picture description, *M*= 0.4, *SD*= 0.1, range= 0.18-0.59; storytelling,

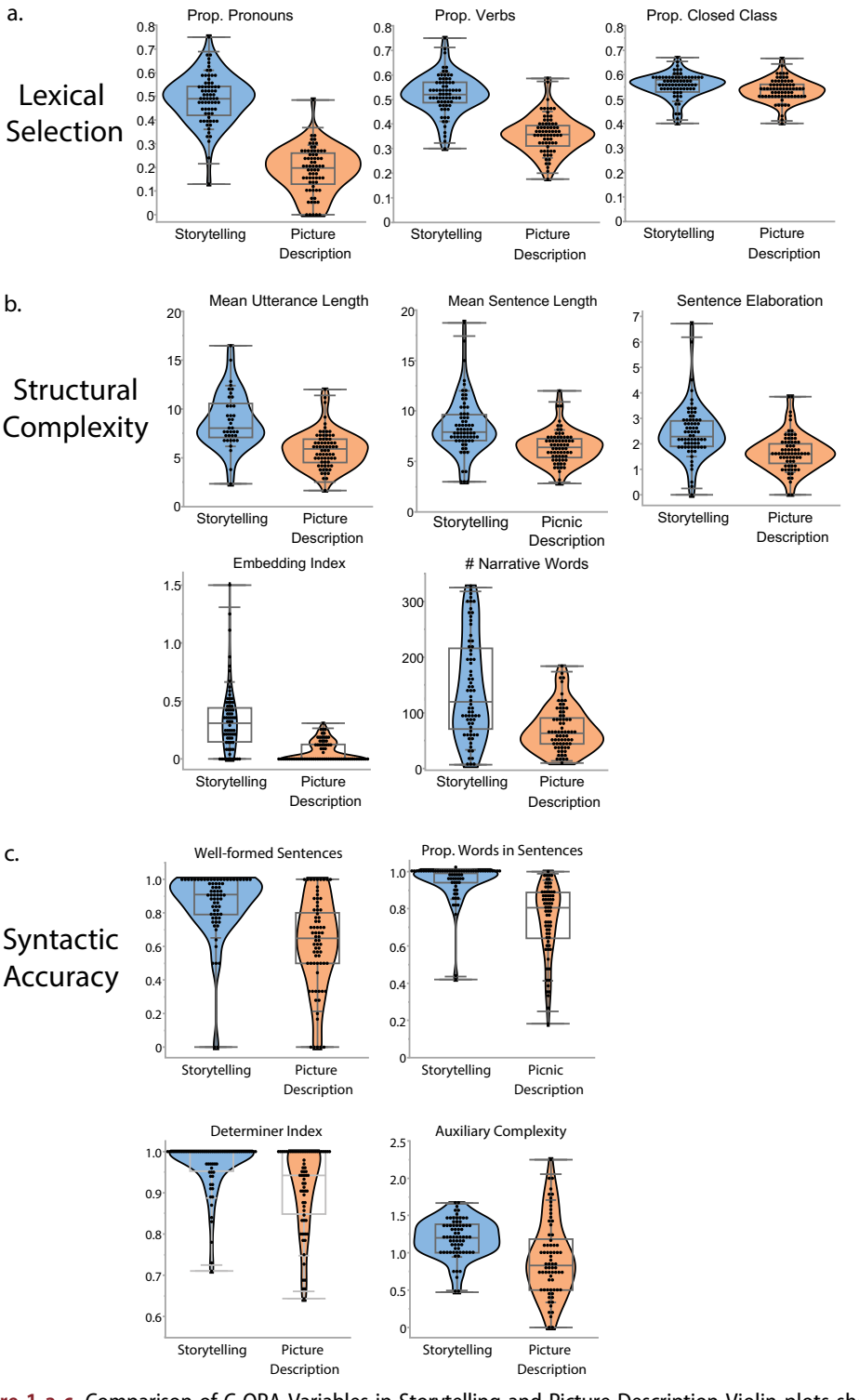

Figure 1 a-c. Comparison of C-QPA Variables in Storytelling and Picture Description Violin plots show the density distribution of performance across participants (indicted by '•') where thicker vs. thinner sections indicate higher vs. lower frequency of performance. Grey lines summarize performance within

M= 0.5, *SD*= 0.1, range= 0.3-0.75; *t*(69)= 12.55, *p*= <.0001. However, the difference between elicitation methods for proportion of closed-class words produced did not reach the alpha corrected threshold for multiple comparisons; picture description, *M*= 0.5, *SD*= 0.1, range= 0.4-0.67; storytelling, *M*= 0.6, *SD*= 0.1, range= 0.4-0.67; *t*(69)= 2.73, *p*= 0.008. Regarding structural complexity, as seen in Figure 1, B picture description elicited utterances with fewer words; mean utterance length: picture description, *M*= 5.8, *SD*= 2.3, range= 1.7-12; storytelling, *M*= 8.3, *SD*= 2.9, range= 2.3-18.8; *t*(68)= 7.83, *p*= <.0001, sentences with fewer words (mean sentence length: picture description, *M*= 6.4, *SD*= 2.4, range= 2.8-12; storytelling, *M*= 8.5, *SD*= 2.8, range= 3-18.8; *t*(67)= 6.9, *p*= <.0001, less elaborated sentences (sentence elaboration: picture description, *M*=1.7, *SD*=0.9, range= 0-3.9; storytelling, *M*= 2.5, *SD*= 1.1, range= 0-6.7; *t*(68)= 6.75, *p*= <.0001, less sentential embedding, embedding index values: picture description, *M*= 0.1, *SD*= 0.1, range= 0-0.3; storytelling, *M*= 0.3, *SD*= 0.3, range= 0-1.5; *t*(69)= 8.40, *p*= <.0001), and significantly fewer narrative words (number of narrative words: picture description, *M*= 70.8, *SD*= 44.4, range= 10-184; storytelling, *M*= 144.1, *SD*= 91.7, range= 7-325; *t*(69)= 7.33, *p*= <.0001. Regarding syntactic accuracy, as seen in Figure 1, C picture description elicited a significantly smaller proportion of well-formed sentences and proportion words in sentences in comparison to storytelling (proportion well-formed sentences: picnic description, *M*= 0.6, *SD*= 0.3, range= 0-1; storytelling, *M*= 0.9, *SD*= 0.2, range= 0-1; *t*(70)= 5.79, *p*< .0001; proportion words in sentences: picture description, *M*= 0.7, *SD*= 0.2, range= 0.18-1; storytelling, *M*= 1.0, *SD*= 0.1, range= 0.42-1.00; *t*(69)= 3.97, *p* < .0001. Participants were less likely to produce determiners during picture description as indicated by a significantly smaller determiner index for picture description (*M*= 0.9, *SD*= 0.2, range= 0.64-1) compared with storytelling (*M*= 1.0, *SD*= 0.1, range= 0.71-1), *t*(67)= 3.97, *p*= .0002. Picture description also elicited less auxiliary complexity in comparison to storytelling (auxiliary complexity: picture description, *M*= 0.9, *SD*= 0.6, range= 0-2.3; storytelling, *M*= 1.2, *SD*= 0.3, range= 0.47-1.7; *t*(69)= 3.86, *p*= .0003. In summary, storytelling elicited greater values for all C-QPA variables compared to picture description at statistically significant levels except for proportion of closed-class words.

Consistency across Participants

We calculated Pearson correlation coefficients to assess whether lexical, structural, and syntactic C-QPA measures were consistently measured across individuals for storytelling and picture description elicitations. See Figure 2 for scatter plots and Table 1 for summary of results. Of the 12 variables, five (42%; mean sentence length, mean utterance length, sentence elaboration, embedding index, and number of narrative words) were significantly correlated across methods, using the same Bonferroni-corrected threshold as for

quantile box plots where median is the center of the box, 25th percentile the bottom, 75th percentile the top, lower bars 10th and 2.5th percentile, and higher bars 90th and 97.5th percentile. All comparisons were statistically significant after multiple-comparison correction (Storytelling (in blue) > Picture Description (in orange); *p*'s < .004) save for proportion closed-class words (*p*= 0.008). **I attached the three panels (a.b.c.) in .eps format so that you can assure that they are the same size in the document. Currently panel C is not the same size as the others.

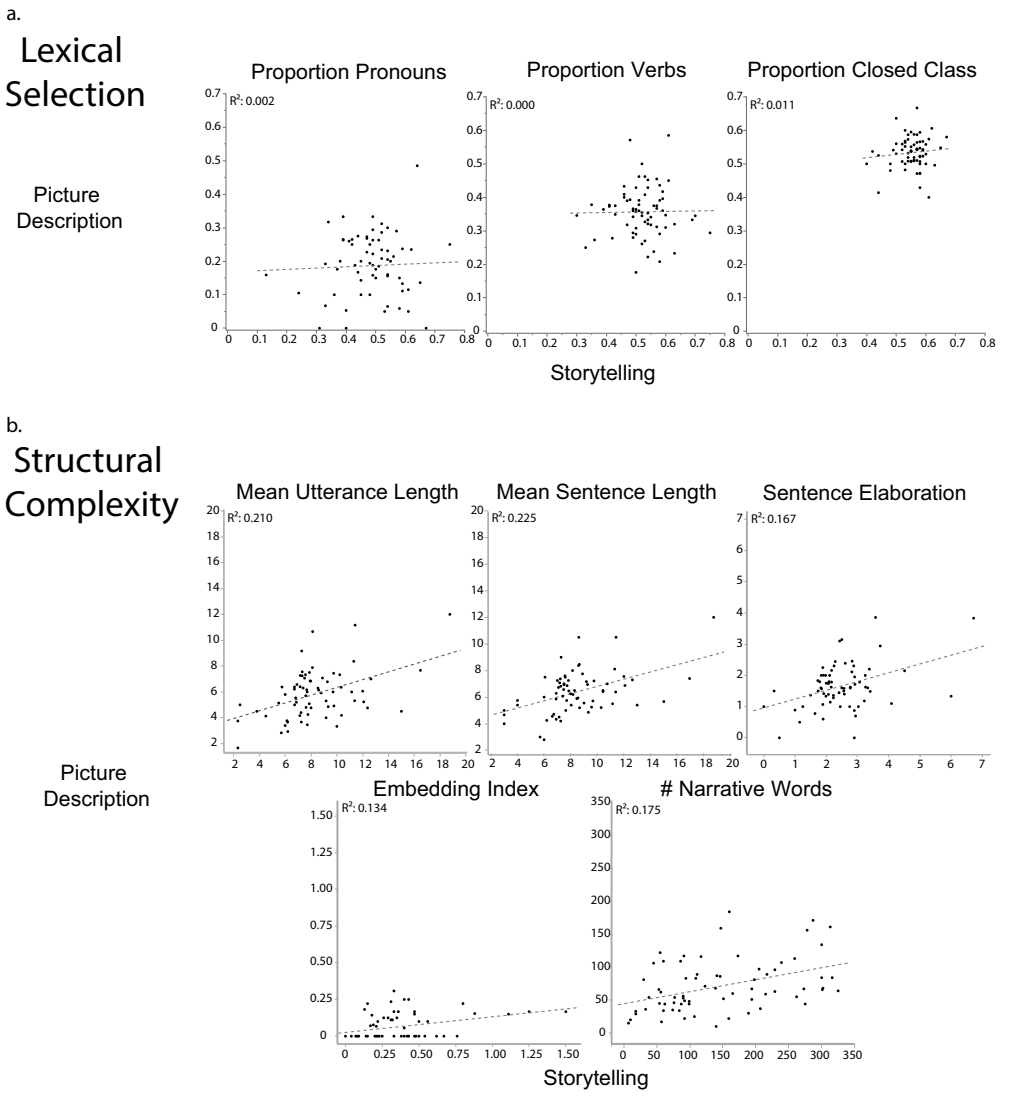

Figure 2. C-QPA Variable Performance across Participants on Storytelling vs. Picture Description Individual performance (indicted by '•') for C-QPA outcomes related to lexical selection (A), structural complexity (B), and syntactic accuracy (C) during storytelling (x-axis) vs. picture description (y-axis). Proportion variance explained (R2) in upper left corners. Structural complexity variables were significantly correlated across elicitation methods after correcting for multiple comparisons (r's > 0.37, p's < .002). Other comparisons were not-significant (r's < 0.12; p's > .18). **I attached the three panels (a. b.c.) in .eps format so that you can assure that they are the same size in the document.**I attached the three panels (a.b.c.) in .eps format so that you can assure that they are the same size in the document. Currently panel C is not the same size as the others.

c.

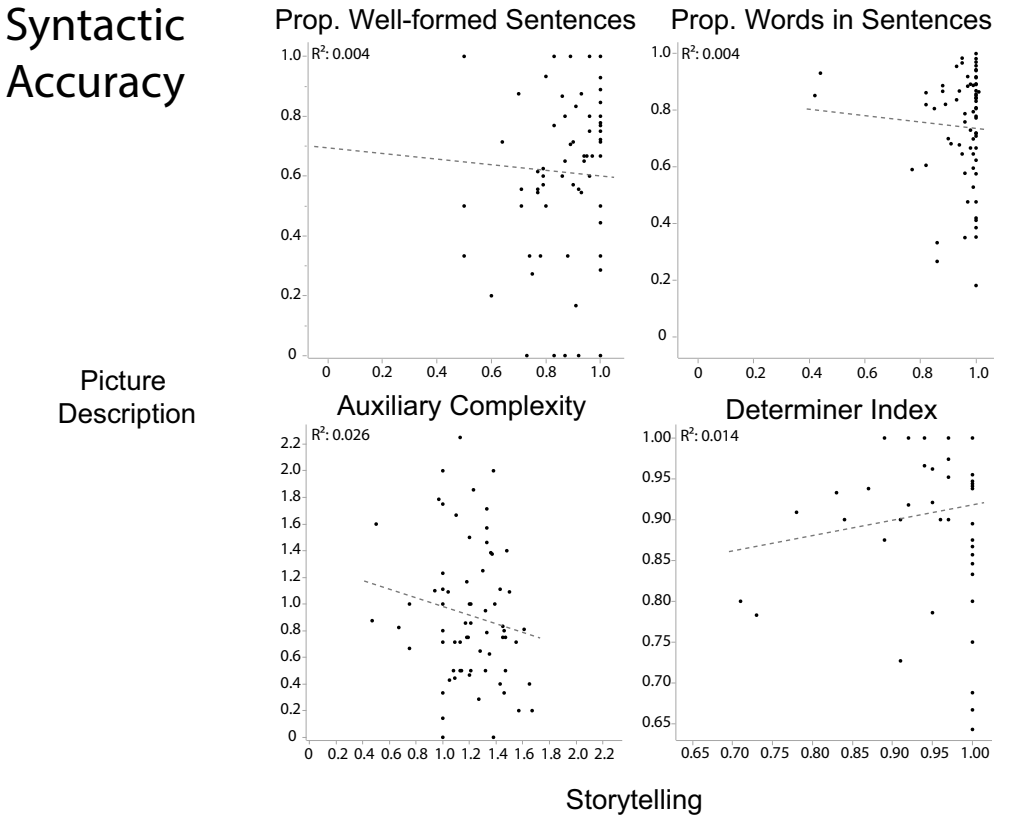

Figure 2. (Continued).

the *t*-tests, *r*'s > 0.37, *p*'s < .002. These variables, all proxies for structural complexity, showed moderate, positive correlations. Remaining C-QPA variables were not significantly correlated across methods, *r*'s < 0.12, *p*'s > 0.18. In summary, although elicitation methods were consistent in reflecting the degree of structural complexity produced across individuals, methods were not consistent in measuring lexical selection and syntactic accuracy abilities.

Consistency in Deficit Identification

We conducted chi-square tests of independence to determine the consistency of picture description and storytelling when identifying impaired and spared speakers across lexical, structural, and syntactic C-QPA measures of connected speech. See Table 1 and Figure 3 for summary of results. Elicitation methods significantly differed in the number of individuals classified as impaired/spared across all three types of connected speech measures. For two of the three lexical selection measures, storytelling classified more participants as impaired than picture description, proportion of verbs, χ^2 (1, N = 70) = 9.57, p = .002; proportion of pronouns, χ^2 (1, N = 70) = 14.96, p < .0001. Picture description classified more participants as impaired for the structural complexity measure related to the

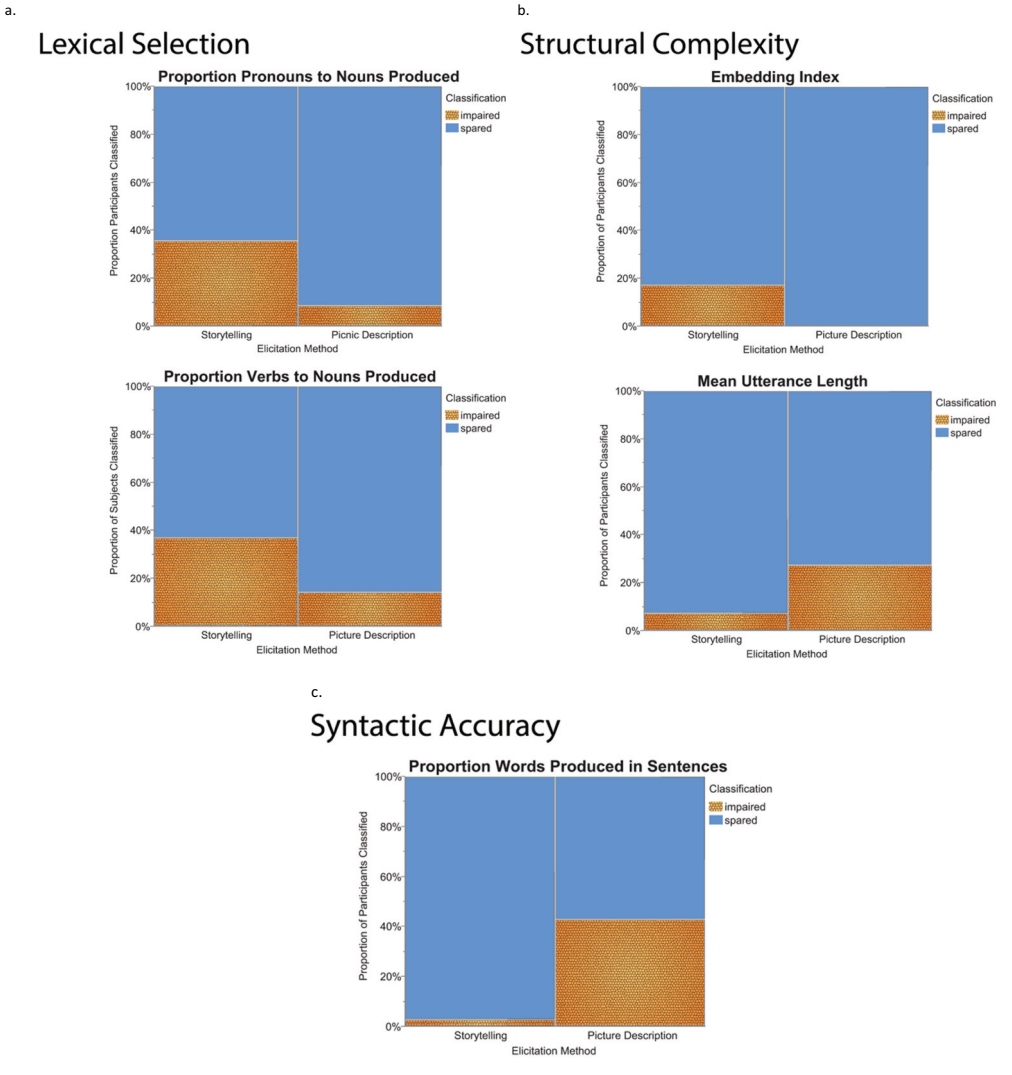

Figure 3. Impairment Classification Differences across Elicitation Methods Significant differences between storytelling and picture description (x-axis) in the proportion of participants classified as impaired (pattern orange) and spared (solid blue; y-axis) for lexical selection (a), structural complexity (b), and syntactic accuracy (c) C-QPA variables (χ^2 's > 9.6, p's < .002).

number of words produced, mean utterance length, χ^2 (1, N = 69) = 10.99, p = .0009. However, storytelling classified more individuals as impaired in the degree of sentence embedding, χ^2 (1, N = 70) = 13.13, p = .0003. Regarding syntactic accuracy, picture description classified more individuals as impaired for the proportion of words produced in sentences χ^2 (1, N = 70) = 31.76, p < .0001. No other impairment classification differences were statistically significant (proportion of closed-class words, χ^2 (1, N = 70) < 1; mean sentence length, χ^2 (1, N = 68) < 1; sentence elaboration, χ^2 (1, N = 69) < 1; number of narrative words produced, χ^2 (1, N = 70) = 4.12, p = .04; proportion of well-formed sentences, χ^2 (1, N = 71) < 1; determiner index, χ^2 (1, N = 68) = 1.11, p= .29; and auxiliary complexity, χ^2 (1, N = 70) < 1. In summary, storytelling classified individuals as

impaired on more connected speech measures (n=3) in comparison to picture description (n=2).

Discussion

Comparing methods used to assess the degree to which speakers produce impaired connected speech after acute stroke before reorganization and rehabilitation of function is important to inform empirical studies of connected speech production and guide early therapeutic strategies. However, the degree to which connected speech elicitation methods differ in eliciting structural and syntactic content is not well-understood. Here we investigated how picture description and storytelling elicitation approaches compared on lexical, structural, and syntactic microlinguistic aspects of connected speech in a large group of speakers during the acute phase of left hemisphere stroke independent of clinical aphasia diagnosis. We compared the degree to which the methods differed at the group level, the degree of consistency at the individual level, and the degree of consistency when categorizing speakers as impaired or spared. In comparison to picture description, storytelling produced more lexically and structurally complex content and more syntactically accurate content, which resulted in different connected speech outcomes for participants across methods and increased the number of participants classified with impairments. These results inform methodological approaches for the study of connected speech deficits in this rare and clinically important patient population of acute stroke.

Participants with acute stroke produced more structurally and syntactically complex and lexically diverse content during storytelling compared to picture description. The pattern of connected speech differences between elicitation methods likely results from the inherent nature of the tasks. During picture description, participants viewed many visually depicted objects and several actors performing actions in a scene, in contrast to storytelling where the narrative was related without visual assistance. This prompted participants during picture description to describe objects and the few depicted actions more often with shorter, ungrammatical utterances (e.g., "couple having a picnic, house with a garage and a car, driveway, sailboat, someone sailing out in the water, a dog by the kid with the kite, a tree"). Consequently, participants produced fewer verbs and pronouns in comparison to nouns, less structural complexity (shorter sentences, shorter utterances, less sentential elaboration and embeddings), and less syntactically accurate speech (fewer syntactically well-formed sentences, words in sentences, and required determiners). Our results are consistent with studies in speakers with neurodegenerative disease that found that picture description in comparison to storytelling with and without pictures or semi-structured interviews produced less syntactically rich and complex speech but was more able to expose semantic deficits (Bose et al., 2022; Clarke et al, 2021; Lavoie et al., 2021; Sajjadi et al., 2012; cf. Ash et al., 2013). Others have also noted that picture description elicits list-like labeling (Bose et al., 2022; Wright & Capilouto, 2009). Wright and Capilouto demonstrated in unimpaired adult speakers that changing the instructions for picture description to emphasize temporal-causal information increased the number of main events described, although the syntactic and structural complexity of speech (i.e., the number of words and clauses within a main clause) did not change. This leaves open for

future research whether picture description with tailored instructions might elicit speech as structurally and syntactically complex as that elicited by spontaneous storytelling.

The differences we observed in content elicited between methods had consequences for measurement consistency and impairment classification. First, regarding lexical selection, on average participants produced fewer nouns in comparison to pronouns and verbs during storytelling vs. picture description. As a result, not only were elicitation methods inconsistent measuring individual capacity to produce nouns in comparison to pronouns and verbs, but further, more participants were classified as outside the range of controls in producing nouns during storytelling (31% of participants during storytelling vs. 3% during picture description). We hypothesize that storytelling captured more noun production deficits because without the depicted objects to facilitate lexical selection, participants were less successful in lexical retrieval. With regards to deficits producing pronouns and verbs in comparison to nouns, far fewer people showed deficits in either storytelling (1% and 6% respectively) or picture description (7% and 13% respectively), in line with the less agrammatic presentation of connected speech in these speakers with acute stroke. It is seemingly paradoxical that a method that elicits list-like labeling (picture description) should less frequently categorize lexical selection impairment in comparison to storytelling. However, this result highlights that storytelling, by virtue of its lack of visual cues, is a more taxing language production task which more closely approaches real-world communication. Thus, we suggest that storytelling is a more ecologically valid measure than picture description when assessing lexical selection ability during spontaneous connected speech in speakers with more mild speech deficits.

Second, although storytelling elicited overall more structurally complex speech, elicitation methods were consistent measuring individual capacities. That is, if one person produced more words per utterance and longer sentences compared to another person, they were also likely to do so during storytelling. However, whether participants were assessed as impaired on structural complexity varied depending on the elicitation method. Storytelling classified more individuals as impaired in the number of embeddings per sentence produced in comparison to controls. This is likely because picture description elicited a more compressed range in performance and fewer overall sentential embeddings. During picture description 60% of participants produced no embeddings at all, but performance was within the control range (and thus no one was classified as impaired). However, 7% of all participants produced no embeddings during storytelling and 17% overall were impaired relative to controls. Thus, because picture description elicited shorter utterances with few to no embeddings for both patients and control participants in comparison to storytelling, storytelling appears the better tool to measure the ability to produce embeddings.

In contrast, participants were more likely to be classified as impaired during picture description for variables that were more dependent on the overall number of narrative words produced (mean number of words per utterance and the proportion of narrative words produced within vs. outside of sentences, a syntactic accuracy variable). Again, the inherent nature of the task is likely responsible for these impairment classifications. As described above, we conjecture that participants approached picture description listing objects in short ungrammatical sequences, an approach controls did not consistently take, which resulted in more participants classified as impaired for the proportion of words produced in sentences.

Lastly, participants performed near ceiling during storytelling for most syntactic accuracy measures, but performance was more variable during picture description. As a result, methods were inconsistent when measuring syntactic accuracy across individuals. That said, across most syntactic measures, the number of people classified as impaired did not differ between elicitation methods (cf. Clarke et al., 2021 for contrasting results in a small group of speakers with mild cognitive impairment and Alzheimer's Disease). Although null results occur for many reasons, elicitation measures may not have differed classifying syntactic impairments because our participant cohort as a whole generally had more mild speech deficits.

In sum, differences between elicitation method language outcomes arose at least in part due to the inherent nature of the tasks. Participants produced more diverse speech and performed substantially better on structural complexity and syntactic accuracy aspects of connected speech during storytelling. It remains an open question if differences in language measures were due to the difference in output, as during storytelling participants produced twice as many narrative words as picture description. However, elsewhere participants with chronic stroke produced more verbs, adjectives and adverbs during storytelling compared to picture description when there was no difference in overall output between elicitation methods (Stark, 2019). While speech samples still require manual transcription, because of recent advancements in automated analysis of the lexical, structural, and syntactic elements of connected speech via the C-QPA and CLAN (Fromm et al., 2021; cf. other approaches used by Clarke et al., 2021), evaluating connected speech for deficits in a clinical or research context is a more reasonable possibility. Given the yield and the possibilities of using automated analysis via the C-QPA, it may be beneficial to administer a lengthier task like storytelling.

Other studies have explored differences in connected speech outcomes between elicitation methods in speakers with chronic stroke, but the methodological designs and variables critically differ between studies, which makes direct comparisons difficult to interpret (e.g., single vs. sequential picture descriptions or answers to questions; Alyahya et al., 2020, 2021; Ash et al., 2013; Dipper et al., 2018). We are aware of one study which directly compared connected speech outcomes between storytelling and picture description in speakers with stroke. When describing a single picture of firefighters rescuing a cat in comparison to narration of the Cinderella tale (Stark, 2019; part of the AphasiaBank protocol, cf. MacWhinney et al., 2011), speakers with chronic stroke did not significantly differ in the total number of words, mean length of utterance and proportion nouns vs. verbs produced across methods in contrast to our results (although both studies found no differences in the proportion of closed vs. open class words produced across elicitation methods). We conjecture this may be due to differences between participant groups in terms of the degree of language impairment (speakers recruited because of clinically diagnosed aphasia after stroke vs. speakers recruited after acute stroke independent of a clinical diagnosis of aphasia). Thus, differences in stimuli used to elicit connected speech, derived outcome measures (notably a previous singular focus on lexical variables) and participant populations make comparisons between studies difficult.

Future Directions

Our study examined connected speech in large group of participants during the acute stage of left hemisphere stroke independent of clinical diagnosis of aphasia.

Although this participant sample provides important information regarding the nature of language deficits before reorganization and rehabilitation of function, by testing performance before recovery during the acute stage of stroke, we *de facto* excluded participants with more severe speech deficits, for example, those who could not produce enough connected speech for analysis or those whose speech was difficult to transcribe due to motor or articulatory deficits. These participants with more severe language deficits are likely included in chronic stroke studies as reorganization and rehabilitation afford improved speech output during recovery. Additionally, because testing was conducted in neuro-intensive care units as part of a larger battery of neuropsychological assessments during the same testing session, we were unable to counterbalance test administration. Thus, the results we report here should be replicated in participants with clinical diagnoses of aphasia with a counterbalanced order of task presentation to avoid potential impact of task order and importantly, extend conclusions to those with increased language deficits.

We adopted a common picture description stimulus to compare with storytelling, the black-and-white picnic picture description task from the Western Aphasia Battery (Kertesz, 2007). However, the recently updated, color version of the "cookie theft" scene from the Boston Diagnostic Aphasia Examination (Berube et al., 2019) that also includes more depicted actions, may be more stimulating and consequently elicit different responses from those elicited from the picnic picture description. In future work we endeavor to establish whether other elicitation methods are more effective at capturing language impairments compared to storytelling.

Lastly, we focused on the microlinguistic features of connected speech, as deficits in these lexical, structural, and syntactic aspects often occur after stroke (Berndt et al., 1997; Gordon, 2006; Matchin et al., 2020). However, deficits to convey a consistent theme across the discourse, i.e. global coherence, and relate successive utterances to each other, i.e., local coherence (Kintsch & Van dijk, 1978; Marini et al., 2011) among deficits to other macrolinguistic features of discourse also occur after stroke and impact communication (e.g., Alyahya et al., 2022; Barker et al., 2017; Glosser & Deser, 1991; Schneider et al., 2021; Sherratt & Bryan, 2012). A fruitful direction for future research will be to extend current investigations of how connected speech tasks differ in eliciting deficits in macrolinguistic features of discourse to larger samples of people after acute and chronic stroke (Fergadiotis & Wright, 2012; Leaman & Edmonds, 2023; Olness, 2007; Wright et al., 2014).

Conclusions

Successful connected speech is critical for communication but methods of assessment after stroke vary with unknown consequences for quantification of deficits. Here, we report for the first time a detailed comparison of the structural and syntactic outcomes for two common approaches, spontaneous storytelling and picture description in a large group of speakers during the acute phase of left hemisphere stroke. In comparison to picture description, storytelling produced more lexically diverse, structurally complex and syntactically accurate content which increased the number of participants classified with lexical and structural impairments in the acute stage after stroke. This suggests that storytelling may be a task better suited to uncover linguistic deficits early post stroke

and may provide a better reflection of the abilities needed to communicate effectively in a real-world setting.

Acknowledgements

This work was presented at the Academy of Aphasia (2021). We thank Jolie Anderson, Miranda Brenneman, Cris Hamilton, Danielle Rossi, and Chia-Ming Lei for data collection. We thank Saketh Katta for his help implementing CLAN and Chia-Ming Lei for his guidance on following the QPA manual. We thank Lynn Maher for her help in diagnosing apraxia of speech in our participant population. Lastly, we thank our participants and their caregivers for their time and effort which made this work possible.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This work was supported by the National Institute on Deafness and Other Communication Disorders [R01DC014976].

ORCID

Tatiana T. Schnur http://orcid.org/0000-0002-4651-7721

Data Availability Statement

Anonymized data that support the findings of this study will be publicly available on Open Science Framework.

CRediT authorship contribution statement

Tatiana T. Schnur: Conceptualization, Visualization, Methodology, Writing – original draft, Supervision, Funding acquisition.

Sharon Wang: Formal analysis, Visualization, Writing – original draft.

References

Alyahya, R. S. W., Halai, A. D., Conroy, P., & Lambon, M. A. (2020). A unified model of post-stroke language deficits including discourse production and their neural correlates. *Brain*, *143*(5), 1541–1554. https://doi.org/10.1093/brain/awaa074.

Alyahya, R. S. W., Conroy, P., Halai, A. D., & Ralph, M. A. L. (2022). An efficient, accurate and clinically-applicable index of content word fluency in Aphasia. Aphasiology, 36(8), 921–939. https://doi.org/10.1080/02687038.2021.1923946

Alyahya, R. S. W., Halai, A. D., Conroy, P., & Ralph, M. A. L. (2021). Content word production during discourse in aphasia: Deficits in word quantity, not lexical–semantic complexity. *Journal of Cognitive Neuroscience*, *33*(12), 2494–2511. https://doi.org/10.1162/jocn_a_01772.

Ash, S., Evans, E., O'Shea, J., Powers, J., Boller, A., Weinberg, D., … Grossman, M. (2013). Differentiating primary progressive aphasias in a brief sample of connected speech. *Neurology*, *81*(4), 329–336. https://doi.org/10.1212/WNL.0b013e31829c5d0e.

Barker, M. S., Young, B., & Robinson, G. A. (2017). Cohesive and coherent connected speech deficits in mild stroke. *Brain and Language*, *168*, 23–36. https://doi.org/10.1016/j.bandl.2017.01.004

Berndt, R. S., Mitchum, C. C., Haendiges, A. N., & Sandson, J. (1997). Verb retrieval in aphasia: 1. Characterizing single word impairments. *Brain and Language*, *56*(1), 68–106. https://doi.org/10.1006/brln.1997.1727 .

Berndt, R. S., Wayland, S., Rochon, E., Saffran, E. M., & Schwartz, M. F. (2000). *Quantitative production analysis: A training manual for the analysis of aphasic sentence production*. Hove, East Sussex UK: Psychology Press.

Berube, S., Nonnemacher, J., Demsky, C., Glenn, S., Saxena, S., Wright, A., Tippett, D. C., & Hillis, A. E. (2019). Stealing Cookies in the Twenty-First Century: Measures of Spoken Narrative in Healthy Versus Speakers With Aphasia. *American journal of speech-language pathology*, *28*(1S), 321–329. https://doi.org/10.1044/2018_AJSLP-17-0131.

Boschi, V., Catricalà, E., Consonni, M., Chesi, C., Moro, A., & Cappa, S. F. (2017). Connected speech in neurodegenerative language disorders: A review. *Frontiers in Psychology*, *8*(MAR). https://doi.org/10.3389/fpsyg.2017.00269.

Bose, A., Dutta, M., Dash, N. S., Nandi, R., Dutt, A., & Ahmed, S. (2022). Importance of Task Selection for Connected Speech Analysis in Patients with Alzheimer's Disease from an Ethnically Diverse Sample. *Journal of Alzheimer's disease: JAD*, *87*(4), 1475–1481. https://doi.org/10.3233/JAD-220166.

Brisebois, A., Brambati, S. M., Boucher, J., Rochon, E., Leonard, C., Désilets-Barnabé, M., Desautels, A., & Marcotte, K. (2022). A longitudinal study of narrative discourse in post-stroke aphasia. *Aphasiology*, *36*(7), 805–830. https://doi.org/10.1080/02687038.2021.1907295

Brott, T., Adams, H. P., Jr, Olinger, C. P., Marler, J. R., Barsan, W. G., Biller, J., Spilker, J., Holleran, R., Eberle, R., & Hertzberg, V. (1989). Measurements of acute cerebral infarction: a clinical examination scale. *Stroke*, *20*(7), 864–870. https://doi.org/10.1161/01.str.20.7.864.

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

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

Clarke, N., Barrick, T. R., & Garrard, P. (2021). A Comparison of Connected Speech Tasks for Detecting Early Alzheimer's Disease and Mild Cognitive Impairment Using Natural Language Processing and Machine Learning. *Frontiers in Computer Science*, 3. https://doi.org/10.3389/fcomp.2021.634360.

Dabul, B. (2000). *Apraxia battery for adults (2nd ed.)*. Austin, TX: Pro-Ed.

Ding, J., Martin, R. C., Cris Hamilton, A., & Schnur, T. T. (2020). Dissociation between frontal and temporal-parietal contributions to connected speech in acute stroke. *Brain*, *143*(3), 862–876. https://doi.org/10.1093/brain/awaa027 .

Ding, J., & Schnur, T. T. (2022). Anterior connectivity critical for recovery of connected speech after stroke. *Brain Communications*, *4*(6), fcac266. https://doi.org/10.1093/braincomms/fcac266.

Dipper, L., Pritchard, M., Walkden, E., & Cruice, M. (2018). How do speakers with and without aphasia use syntax and semantics across two discourse genres? *Aphasiology*, *32*(6), 720–738. https://doi.org/10.1080/02687038.2018.1447642 .

Fergadiotis, G., & Wright, H. H. (2011). Lexical diversity for adults with and without aphasia across discourse elicitation tasks. *Aphasiology*, *25*(11), 1414–1430. https://doi.org/10.1080/02687038.2011.603898

Filiou, R.P, Bier, N., Slegers A., Houzé, B., Belchior, P. & Brambati, S.M. (2020) Connected speech assessment in the early detection of Alzheimer's disease and mild cognitive impairment: a scoping review, *Aphasiology*, *34*(6), 723–755, DOI: https://doi.org/10.1080/02687038.2019.1608502 .

Flowers, H. L., Silver, F. L., Fang, J., Rochon, E., & Martino, R. (2013). The incidence, co-occurrence, and predictors of dysphagia, dysarthria, and aphasia after first-ever acute ischemic stroke. *Journal of communication disorders*, *46*(3), 238–248. https://doi.org/10.1016/j.jcomdis.2013.04.001.

Fromm, D., Forbes, M., Holland, A., & MacWhinney, B. (2020). Using AphasiaBank for Discourse Assessment. *Seminars in speech and language*, *41*(1), 10–19. https://doi.org/10.1055/s-0039-3399499 .

Fromm, D., Katta, S., Paccione, M., Hecht, S., Greenhouse, J., MacWhinney, B., & Schnur, T. T. (2021). A Comparison of Manual Versus Automated Quantitative Production Analysis of Connected Speech. *Journal of Speech, Language, and Hearing Research*, *64*(4), 1271–1282. https://doi.org/10.1044/2020_JSLHR-20-00561.

Glosser, G., & Deser, T. (1991). Patterns of discourse production among neurological patients with fluent language disorders. *Brain and Language*, *40*(1), 67–88. https://doi.org/10.1016/0093-934x(91)90117-j

Goodglass, H., Barresi, B., & Kaplan, E. (1983). *The Boston Diagnostic Aphasia Examination*. Philadelphia: Lippincott Williams and Wilkins.

Gordon, J. K. (2006). A quantitative production analysis of picture description. *Aphasiology*, *20*(2–4), 188–204. https://doi.org/10.1080/02687030500472777.

Gordon, J. K. (2020). Factor Analysis of Spontaneous Speech in Aphasia. *Journal of Speech, Language, and Hearing Research*, *63*(12), 4127–4147. https://doi.org/10.1044/2020_JSLHR-20-00340.

Grönberg, A., Henriksson, I., Stenman, M., & Lindgren, A. G. (2022). Incidence of Aphasia in Ischemic Stroke. *Neuroepidemiology*, *56*(3), 174–182. https://doi.org/10.1159/000524206 .

Jeffers, S., Perrault, C., & Ehrlich, A. (2004). *Cinderella*. New York: Dutton Children's Books.

Kintsch, W., & van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological Review*, *85*(5), 363–394. https://doi.org/10.1037/0033-295X.85.5.363

Kertesz, A. (2007). *Western Aphasia Battery - Revised*.

Lavoie, M., Black, S. E., Tang-Wai, D. F., Graham, N. L., Stewart, S., Leonard, C., & Rochon, E. (2021). Description of connected speech across different elicitation tasks in the logopenic variant of primary progressive aphasia. *International journal of language & communication disorders*, *56*(5), 1074–1085. https://doi.org/10.1111/1460-6984.12660.

Leaman, M. C., & Edmonds, L. A. (2023). Analyzing language in the picnic scene picture and in conversation: The type of discourse sample we choose influences findings in people with aphasia. *American Journal of Speech-Language Pathology*, 1–18. https://doi.org/10.1044/2023_AJSLP-22-00279

MacWhinney, B. (2000). *The CHILDES Project: Tools for analyzing talk* (Third Edit). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

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

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. https://doi.org/10.1080/02687038.2011.584690

Martin, R. C., Ding, J., Hamilton, A. C., & Schnur, T. T. (2021). Working Memory Capacities Neurally Dissociate: Evidence from Acute Stroke. *Cerebral Cortex Communications*, *2*(2), tgab005. https://doi.org/10.1093/texcom/tgab005 .

Martin, R. C., & Schnur, T. T. (2019). Independent contributions of semantic and phonological working memory to spontaneous speech in acute stroke. *Cortex*, *112*, 58–68. https://doi.org/10.1016/j.cortex.2018.11.017.

Matchin, W., Basilakos, A., Stark, B. C., den Ouden, D. B., Fridriksson, J., & Hickok, G. (2020). Agrammatism and Paragrammatism: A Cortical Double Dissociation Revealed by Lesion-Symptom Mapping. *Neurobiology of language (Cambridge, Mass.)*, *1*(2), 208–225. https://doi.org/10.1162/nol_a_00010 .

Meyer, B. C., & Lyden, P. D. (2009). The modified National Institutes of Health Stroke Scale: its time has come. *International journal of stroke: official journal of the International Stroke Society*, *4*(4), 267–273. https://doi.org/10.1111/j.1747-4949.2009.00294.x.

Olness, G. S. (2006). Genre, verb, and coherence in picture-elicited discourse of adults with aphasia. *Aphasiology*, *20*(2–4), 175–187. https://doi.org/10.1080/02687030500472710

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.org/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.org/10.1016/0093-934X(89)90030-8

Schneider, F., Marcotte, K., Brisebois, A., Townsend, S. A. M., Smidarle, A. D., Loureiro, F., da Rosa Franco, A., Soder, R. B., Nikolaev, A., Marrone, L. C. P., & Hübner, L. C. (2021). Neuroanatomical correlates of macrolinguistic aspects in narrative discourse in unilateral left and right hemisphere stroke: A voxel-based morphometry study. *Journal of Speech, Language, and Hearing Research*, *64*(5), 1650–1665. https://doi.org/10.1044/2020_JSLHR-20-00500

Schnur, T. T., & Lei, C.-M. (2022). Assessing naming errors using an automated machine learning approach. *Neuropsychology*, *36*(8), 709–718. https://doi.org/10.1037/neu0000860 .

Sherratt, S., & Bryan, K. (2012). Discourse production after right brain damage: Gaining a comprehensive picture using a multi-level processing model. *Journal of Neurolinguistics*, *25*(4), 213–239. https://doi.org/10.1016/j.jneuroling.2012.01.001

Stark, B. C. (2019). A comparison of three discourse elicitation methods in aphasia and age-matched adults: Implications for language assessment and outcome. *American Journal of Speech-Language Pathology*, *28*(3), 1067–1083. https://doi.org/10.1044/2019_AJSLP-18-0265

Wilson, S. M., Henry, M. L., Besbris, M., Ogar, J. M., Dronkers, N. F., Jarrold, W., … Gorno-Tempini, M. L. (2010). Connected speech production in three variants of primary progressive aphasia. *Brain*, *133*(7), 2069–2088. https://doi.org/10.1093/brain/awq129

Wright, H. H., & Capilouto, G. J. (2009). Manipulating task instructions to change narrative discourse performance. *Aphasiology*, *23*(10), 1295–1308. https://doi.org/10.1080/02687030902826844.

Wright, H. H., Koutsoftas, A. D., Capilouto, G. J., & Fergadiotis, G. (2014). Global coherence in younger and older adults: Influence of cognitive processes and discourse type. Neuropsychology, development, and cognition. *Aging, Neuropsychology and Cognition*, *21*(2), 174–196. https://doi.org/10.1080/13825585.2013.794894

Zahn, R., Schnur, T.T., & Martin, R.C. (2023). Contributions of semantic and phonological working memory to narrative language independent of single word production: Evidence from acute stroke. *Cognitive Neuropsychology*, *39*(5–8), 296–324. https://doi.org/10.1080/02643294.2023.2186782

Appendix

An example of picture description transcription, utterance segmentation, and morphological parsing.

Original Transcription

Participant: Uh man and woman are having a picnic and a man is flying a kite with his dog running beside him and uh someone is sailing on a sailboat and someone's fishing on a pier and this person is playing in the sand this child (3 sec) at a house and a tree and a flag and trees and the water
Experimenter: Anything else
Participant: The man is reading a book and the woman is pouring a drink and they're listening to the radio
Experimenter: Is that it
Participant: yes

Words Uttered

man and woman are having a picnic and a man is flying a kite with his dog running beside him and someone is sailing on a sailboat and someone is fishing on a pier and this person is playing in the sand this child at a house and a tree and a flag and trees and the water
The man is reading a book and the woman is pouring a drink and they are listening to the radio
Yes

Narrative Words

man and woman are having a picnic and a man is flying a kite with his dog running beside him and someone is sailing on a sailboat and someone is fishing on a pier and this person is playing in the sand this child at a house a tree a flag trees the water
The man is reading a book and the woman is pouring a drink and they are listening to the radio

Utterances

[man and woman are having a picnic] [a man is flying a kite with his dog running beside him] [someone is sailing on a sailboat] [someone is fishing on a pier] [this person is playing in the sand] [this child at a house] [a tree] [a flag] [trees] [the water]
[The man is reading a book] [the woman is pouring a drink] [they are listening to the radio]

Morphological Parsing following Fromm et al. (2021)

0det man and 0det woman are having a picnic
a man is flying a kite with his dog running beside him
someone is sailing on a sailboat
someone is fishing on a pier
this person is playin in the sand
this child at a house [+ gram]
a tree [+ gram]
a flag [+ gram]
trees [+ gram]
the water [+ gram]
the man is reading a book
the woman is pouring a drink
they are listening to the radio