International Journal of Language & Communication Disorders

INT J LANG COMMUN DISORD, JANUARY/FEBRUARY 2020,
VOL. 55, NO. 1, 110–120

Research Report

Evaluating informative content and global coherence in fluent and non-fluent aphasia

Audrey A. Hazamy† and Jessica Obermeyer‡

†Department of Communication Arts, Sciences, and Disorders, Brooklyn College, Brooklyn, NY, USA
‡Department of Communication Sciences and Disorders, Temple University, Philadelphia, PA, USA

(*Received May 2019; accepted September 2019*)

Abstract

Background: Discourse analysis is an important component of aphasia assessment because it can provide an insight into functional communication abilities. However, there are many unknowns regarding the levels of discourse breakdowns that occur across aphasia types. The purpose of the current study is to explore the possible differences in discourse-level communication in persons with fluent and non-fluent aphasia during picture description.
Aims: To examine if persons with fluent and non-fluent aphasia differ on utterance-level discourse measures when evaluating informative content and global coherence. Additionally, to evaluate and compare the types of global coherence violations made by each group.
Methods & Procedures: Data from 31 people with aphasia was collected from AphasiaBank, which included 13 people with fluent aphasia and 18 people with non-fluent aphasia. Discourse samples from three picture descriptions were analysed. Discourse outcomes included utterances with new information (UNIs—relevant utterances containing new information) and global coherence (the extent to which each utterance maintained the overall discourse theme). Additionally, seven types of errors were identified to explore the nature of breakdowns in global coherence.
Outcomes & Results: People with fluent aphasia produced significantly higher proportions of UNIs and had significantly higher average global coherence ratings than those with non-fluent aphasia. Differences in global coherence violations were identified with people with fluent aphasia producing more non-specific, incorrect and off-topic utterances and people with non-fluent aphasia producing more incomplete utterances. One of the most common global coherence error types in both groups was commentary.
Conclusions & Implications: Although people with fluent aphasia produced more types of global coherence errors, including incorrect, non-specific and off-topic utterances, the group was still rated significantly higher on utterance relevance and topic maintenance, indicating that the ability to produce a complete utterance plays an important role in some aspects of discourse production. Additionally, these findings provide an insight into potential targets for intervention.

Keywords: aphasia, discourse production, information content, coherence.

What this paper adds

What is already known on the subject

The majority of studies on discourse in stroke aphasia have emphasized examining impairments in higher level language processing by comparing client performance with neurologically healthy controls. These findings have suggested that people with aphasia demonstrate breakdowns in conveying informative content and maintaining unified semantic themes in their discourse production.

Address correspondence to: Audrey A. Hazamy, Brooklyn College, Department of Communication Arts, Sciences, and Disorders, 2900 Bedford Avenue, Brooklyn, NY 11210, USA; e-mail: audrey.hazamy68@brooklyn.cuny.edu

International Journal of Language & Communication Disorders
ISSN 1368-2822 print/ISSN 1460-6984 online
DOI: 10.1111/1460-6984.12507

What this paper adds to existing knowledge
The paper evaluated the differences in conveying utterance-level informative content and maintaining global coherence in the discourse of people with fluent and non-fluent aphasia. Additionally, the types of global coherence errors were identified and compared to provide an insight into the nature of global coherence breakdowns between the two groups.

What are the potential or actual clinical implications of this work?
The findings suggest that individuals with fluent aphasia were rated higher in their ability to produce informative and coherent discourse in response to picture-description tasks than individuals with non-fluent aphasia. Furthermore, it demonstrates group commonalities as well as differences in their patterns of errors which contribute to low discourse coherence. We suggest that fluency itself is an important facet of discourse ability and provide possible unique targets for assessment and intervention to improve communicative competency in these groups.

Introduction

Historically, the majority of research and clinical focus in persons with aphasia concerned assessment and intervention at the word and sentence levels, yet the ultimate goal of rehabilitation in this population is discourse functionality. Discourse, often broadly defined as language beyond the single clause or individual sentence level (e.g., Armstrong 2000), is central to everyday communication and establishing and maintaining relationships (Pritchard *et al.* 2017). Although many treatment approaches have successfully intervened at the word and/or sentence levels of language, improvements in these linguistic abilities do not necessarily translate to improvements in discourse ability and modest treatment-related improvements in discourse may be missed on traditional standardized tests (Armstrong 2000). Furthermore, persons with aphasia may demonstrate paradoxical discourse functionality despite their language impairment (e.g., Olness and Ulatowska 2011). As such, discourse analysis has emerged as a means of evaluating communicative effectiveness and functionality and there is increasing recognition of the importance of generalization effects to connected speech in response to intervention (Dietz and Boyle 2018). However, despite ongoing research into the discourse of individuals with aphasia, there is limited information regarding how persons with different types of aphasia differ on discourse-level communication breakdowns. As aphasia types can vary widely in their discourse abilities, identifying potential differences in linguistic breakdowns is imperative to establishing comparative norms and subsequently identifying targets for intervention to maximize communicative competence.

Discourse production is a complex process with multiple levels of representation involving semantic and linguistic information (Sherratt 2007). Sherratt (2007) describes a hypothetical multilevel discourse processing model, adapted from Frederiksen *et al.* (1990), which involves both top-down and bottom-up processing. According to this model, discourse production begins with the selection/retrieval of a discourse frame from memory (*frame/schema generation*), which represents the type of discourse (e.g., narrative, procedural) to be produced. Following, details associated with events, such as the setting and actors, are inserted into the framework (*insertion of semantic information*). Semantic knowledge and information stored in long-term memory are then integrated (*integration of semantic information*) to promote the connectedness of discourse. Processes then *select and prioritize information*, such that determinations are made regarding what needs to be explicitly stated and the order in which information is to be shared. Propositions are then specified through the encoding of selected information (*generation, selection and chunking of propositions*). Finally, during *linguistic formulation,* syntactic analysis and lexical/morphological processing takes place (e.g., linguistic encoding of propositions, application of cohesive mechanisms) before the discourse is produced (*articulation*). The complexity of this process illustrates important considerations in terms of clinical care and research in aphasia. Given the multilevel nature and interconnectivity of processes, production at the discourse level presents unique challenges for various aphasia types not present at the word or isolated sentence level. Moreover, it highlights the importance of identifying the types of breakdowns that occur across different aphasia types in order to tailor intervention strategies to best support improvements in functional communication. For instance, Sherratt (2007) documents a relationship between the use of non-specific elements and the appropriateness of discourse grammar. She goes on to note that difficulty during schema and frame generation may influence a speaker's ability to chunk propositions for syntactic encoding resulting in an increased usage of non-specific elements, thus identifying a possible higher level conceptual target for intervention in patients who demonstrate a high incidence of empty phrase and indefinite word use during discourse production.

Discourse in aphasia

Discourse research has identified several impairments in persons with aphasia compared with non-neurologically impaired controls. For instance, individuals with aphasia convey less information across several informativeness measures and various discourse tasks (for reviews, see Armstrong 2000 and Linnik *et al.* 2016). For example, Nicholas and Brookshire (1993) introduce the now often used correct information unit (CIU) as a word-level content quantification measure and report that persons with aphasia produce fewer CIUs and per cent CIUs compared with non-brain damaged controls. Discourse deficits are also reported in cohesion, or the semantic connectedness of discourse units. Andreetta *et al.* (2012) report more errors of cohesion during a narrative discourse task in a group with anomic aphasia compared with controls. Studies have also documented deficits across various other microlinguistic (i.e., lexical and syntactic) skills in persons with aphasia during connected speech, including measures of productivity (e.g., number of words, mean length of utterance) and grammatical structuring (e.g., omissions, per cent complete sentences) (e.g., Andreetta and Marini 2015). Adults with aphasia may also demonstrate breakdowns in macrolinguistic (i.e., pragmatic) measures of discourse such as the topic maintenance measure of global coherence, which reflects the extent to which a discourse unit relates to an overarching semantic topic (Glosser and Deser 1991) (for a review, see Ellis *et al.* 2016). For instance, in a study using story retell, Wright and Capilouto (2012) report that, compared with controls, speakers with aphasia exhibit significantly lower overall global coherence scores characterized by utterances which lack substantive information and thus require more inferencing on the part of the listener. Furthermore, given the complex and dynamic nature of discourse processing, it is perhaps not surprising that researchers also note the interaction between various discourse skills. For example, macrolinguistic skills have been linked to breakdowns in microlinguistic measures in this population (e.g., Wright and Capilouto 2012, Sherratt 2007, Andreetta and Marini 2015).

Discourse across aphasia types

Despite the growing collection of studies providing performance comparisons with controls, few have directly attempted to differentiate discourse-level characteristics in persons with different types of aphasia. In an early study, Christiansen (1995) explores coherence violations amongst persons with mild anomic, conduction and Wernicke's aphasia. Analysing discourse elicited using cartoon stories, Christiansen notes that the three subgroups evidenced different patterns of violations. While individuals with anomic aphasia presented with violations predominantly related to information completeness, persons with conduction aphasia primarily produced errors associated with repetition of propositional content. Moreover, individuals with Wernicke's aphasia produced more errors related to utterance relevance. Christiansen suggests that while the various fluent aphasia types all demonstrated deficits in coherence, the differential qualitative error patterns may be associated with different underlying impairments and/or compensatory strategies. She goes on to propose that the errors made by those with anomic and conduction aphasia may reflect different adaptive strategies to compensate for surface-level (i.e., lexical and syntactic) deficits; however, the relevance errors noted in persons with Wernicke's aphasia may be directly attributed to underlying impairments in coherence.

While Christiansen (1995) compares subtypes of fluent aphasia, others have also more broadly compared fluent and non-fluent types across discourse measures. For instance, in a study of informativeness during picture description, Brookshire and Nicholas (1995) examine the types of violations made for productions that did not qualify as words or CIUs in persons with fluent and non-fluent aphasia and non-brain-damaged controls. Descriptively, all groups differed from one another on the number of words and CIUs produced; however, the fluent and non-fluent groups did not differ from one another on per cent CIUs produced. Regarding production violations, both aphasia groups significantly differed from controls on several deviation categories (e.g., inaccurate, false starts, unintelligible productions), yet only significantly differed from one another on non-word fillers such that the fluent group produced fewer of these deviations. In a study of lexical diversity in aphasia, Wright *et al.* (2003) note that persons with fluent aphasia produce significantly higher values on measures of productive vocabulary during a picture description discourse task than those with non-fluent aphasia. That is, even though word-retrieval deficits are pervasive in both fluent and non-fluent aphasia types, adults with fluent aphasia produce samples with a more diverse vocabulary, even when sample length is controlled.

In a more recent study, Manning and Franklin (2016) note differences between persons with fluent and non-fluent aphasia during narrative discourse on several macro- and microstructural features. They report that the individuals with non-fluent aphasia produced higher rates of subject/object omissions and article errors and omissions than those with fluent aphasia, although the latter finding was not statistically significant. Furthermore, persons with fluent aphasia produced a higher rate of errors related to information being presented in an illogical order as well as pronoun errors than the non-fluent group. Though these results were also not statistically significant, together with the other aforementioned results and studies, they provide preliminary evidence that various aphasia types may exhibit unique deficit manifestations across several variables during connected speech and highlights the need for establishing comparative norms for higher level language processing. Ascertaining the presence of differences and/or commonalities in performance across aphasia types during discourse is an important step to identifying group specific assessment and intervention needs.

The current study

A review of the literature reveals a vast number of connected speech analysis measures targeting various levels of discourse. Ideally, communication is informative, efficient and organized. As such, in the current study we have chosen to focus on measures that attempt to capture these characteristics. To capture the amount of information conveyed by a speaker, del Toro *et al.* (2008) constructed a measure, utterances with new information (UNIs), for proportionally assessing information content at the utterance level which is both relevant and coherent. Global coherence, mentioned above, is a measure of topic maintenance and considered to reflect the ease with which a listener can identify units of discourse as being semantically unified with the overall theme (e.g., Wright and Capilouto 2012). Difficulty with either measure could significantly impact communicative capability in adults with aphasia; however, it is not known whether these measures may differentiate fluent and non-fluent aphasia types. The purpose of the current study was to explore possible differences in discourse-level communication in persons with fluent and non-fluent aphasia during picture description. It addresses the following specific aims:

(1) To examine whether persons with fluent and non-fluent aphasia differ on the utterance-level discourse measure of information content. As others have noted, that groups did not proportionally differ on word-level measures of informativeness (Brookshire and Nicholas 1995), we predicted that groups would then also not differ at the utterance level.

(2a) To examine whether these groups differ on the utterance-level discourse measure of global coherence.

(2b) Subsequently to determine whether they differ in terms of the types of violations made that contribute to breakdowns in global coherence.

We predicted that groups would not differ on overall global coherence (Aim 2a) as informativeness and global coherence are highly interdependent (Marini and Urgesi 2012, Wright and Capilouto 2012); however, we did anticipate that groups would differ on the types of violations made (Aim 2b). While the analysis of error types was largely exploratory in nature, we expected that groups would differ on errors related to their language profile. For instance, we anticipated that persons with non-fluent aphasia would produce errors of utterance completeness or information gaps, while persons with fluent aphasia would produce errors related to specificity. We further expected that groups would be similar on other error types which may be more strategic in nature (e.g., use of repetition and task commentary). The nature of the error analysis also allowed the examination of how linguistic breakdowns (e.g., word-finding errors), compensations for linguistic breakdowns (e.g., commentary) and conventional global coherence breakdowns (e.g., off topic) impacted overall topic maintenance.

Methods

Participants

A total of 31 individuals were included in this study (15 male, 16 female), including 13 persons with fluent and 18 persons with non-fluent aphasia. Data were obtained from the AphasiaBank project (MacWhinney *et al.* 2011). Individuals were between 25 and 80 years of age; had a single, left-hemisphere stroke; and were monolingual speakers of English. Individuals with identified apraxia of speech, concomitant neurological conditions (e.g., tumour, seizure disorder), undetermined aphasia classification, or an anomic or global classification, as determined by the Western Aphasia Battery—Revised (WAB-R; Kertesz 2006), were excluded from the study. Individuals classified as anomic or global in their aphasia type were excluded due to the potential for skewing severity of the two groups. There was no significant difference in aphasia quotient scores between individuals with fluent (mean = 65.84, SD = 11.46) and non-fluent (mean = 59.32, SD = 10.15) aphasia, $t(29) = -1.67$, $p = .105$. Time post-onset was also not significantly different between persons with fluent (mean = 4.27, SD = 3.58) and non-fluent (mean = 5.05, SD = 3.48) aphasia, $t(28) = 0.599$, $p = .554$. The mean ages of persons with fluent and non-fluent aphasia were 65.32 years (SD = 11.77) and 57.32 years (SD = 13.72), respectively. Mean age was not significantly different between groups, $t(29) = -1.70$, $p = .100$. Finally, Levene's test was significant for the demographic variable of years of education ($F = 6.224$, $p = .019$), indicating unequal variances. As such, degrees of freedom (d.f.) were adjusted from 27 to 15.742. The adjusted test indicated that years of education was not significantly different between groups, $t(15.74) =$

Table 1. Participant demographics

AphasiaBank ID	Age (years)	Months post-onset	Education	WAB-AQ	Aphasia type
ACWT10a	48.4	118	12	57.4	Broca
ACWT11a	61.7	26	16	48.9	Wernicke
Adler05a	68.1	18	25	65.5	Conduction
Adler14a	71.4	135	13	83	Conduction
BU02a	66.3	14	12	63.5	Transcortical motor
BU03a	58.7	18	16	77.9	Conduction
BU07a	52.4	50	16	51.5	Broca
BU08a	64.6	110	12	39.7	Broca
BU09a	78.5	70	20	65.3	Conduction
Elman07a	65.5	59	16	63.4	Broca
Elman12a	57.4	42	20	74.4	Wernicke
Elman14a	76.3	56	17	65.7	Conduction
Kansas10a	77	19	12	61.4	Conduction
Kansas13a	43.7	41	11	70.4	Conduction
Kansas17a	54.6	12	13	71.9	Broca
Kansas21a	60.9	144	21	77.2	Conduction
Kansas22a	45.5	31	15	67.5	Conduction
Kansas23a	75.6	10	12	54.7	Conduction
Kempler04a	60.3	40	16	54.6	Broca
Kurkland18a	74.3	56	16	44	Wernicke
Kurkland19a	70.5	106	12	67.2	Broca
MSU07a	25.6	15	16	61.4	Broca
Scale31a	64.5	22	n.a.	51.5	Broca
Scale33a	57.4	104	n.a.	71.1	Broca
Thompson09a	74	48	14	79.3	Transcortical motor
Tucson22a	57	60	16	72	Broca
Whiteside03a	76.5	48	12	47.2	Broca
Whiteside08a	37.8	12	16	54.7	Broca
Whiteside12a	70.5	70	14	54.3	Broca
Whiteside16a	46.8	n.a.	12	53.4	Broca
Wright206a	39	143	14	53.7	Broca

Note: WAB-AQ, Western Aphasia Battery-Aphasia Quotient.

Table 2. Group comparison of demographic variables

Demographic measures	Fluent (n = 13) Mean	Fluent (n = 13) SD	Non-fluent (n = 18) Mean	Non-fluent (n = 18) SD	t	(d.f.)	p	d
WAB-AQ	65.84	11.45	59.32	10.15	–1.67	(29)	.105	.60
Age	65.32	11.77	57.32	13.72	–1.70	(29)	.100	.63
Education	16.46	4.12	13.94	1.81	–2.06	(15.74)	.057	.79
Time post-onset	4.27	3.58	5.05	3.48	0.60	(28)	.554	.22

Note: WAB-AQ, Western Aphasia Battery-Aphasia Quotient; SD, standard deviation.

–2.056, p = .057. The mean years of education were 16.46 (SD = 4.12) for fluent individuals and 13.94 (SD = 1.81) for non-fluent individuals. See table 1 for descriptive information by participant and table 2 for demographic comparisons by group.

Discourse samples

Language samples consisted of responses to picture-description tasks obtained from AphasiaBank (MacWhinney *et al.* 2011). All participants provided responses to the following pictorial stimuli as part of the project protocol: cat rescue (Nicholas and Brookshire 1993); and broken window and refused umbrella (MacWhinney *et al.* 2011). All discourse samples within the AphasiaBank database are transcribed in the CHAT format (MacWhinney 2000) and utterances segmented in accordance with Quantitative Production Analysis standards (Berndt *et al.* 2000).

Discourse analyses

Utterances were coded on the linguistic outcome measures of UNIs (del Toro *et al.* 2008) and overall global coherence score. Utterances classified as having low global coherence scores were then coded with one of seven identified global coherence error types, which are described below.

Utterances with new information (UNIs)

The UNIs provided a measure of information content. An utterance was identified as a UNIs if it provided new, semantically coherent information which could be restated by the reader (del Toro *et al.* 2008). For example, if a participant produced the following utterance when describing the cat rescue picture: 'And there was a tree there that they could get higher and get the bird.' This utterance would not be considered a UNI because the information is not semantically coherent with the theme of the picture (i.e., the firefighters are not identified and are not there to get the bird). The utterance 'The bird is happy singing along' would be considered a UNI because it is consistent with the pictured event.

Global coherence

Global coherence provided a measure of topic maintenance and was evaluated using a four-point scale (Wright *et al.* 2010, Wright and Capilouto 2012) adapted for use with AphasiaBank discourse transcripts. Each utterance was evaluated based on the degree to which it related to the global discourse topic. Rating scale scores were defined as follows with examples from the cat rescue picture (Nicholas and Brookshire 1993): 4 = definite relations between utterance and main detail of the topic (e.g., There's a cat in this tree); 3 = utterance is related to the topic but may include tangential information or is related to the topic but is missing information that must be inferred (e.g., He was stuck because this fell); 2 = utterance is remotely related to the topic or references an unimportant/non-critical component of the stimulus (e.g., He can't get himself); and 1 = no relationship between utterance and topic (e.g., Those are great) (Wright *et al.* 2013).

Global coherence errors

To analyse whether fluent and non-fluent subtypes demonstrated different error patterns in global coherence, the entire discourse corpus was considered to identify different types of violations. Utterances with low global coherence scores (i.e., 1 or 2) were then coded for seven identified coherence violation types. *Commentary* category errors stemmed from utterances that included comments on the task or task performance. *Non-specific* error violations stemmed from utterances that were ambiguous due to an overreliance on vague/non-specific words. *Not complete* category errors were defined as single words or utterances which were abandoned before conveying all of the required information. Utterances that repeated information without adding additional material were classified as *repeated* error violations. *Incorrect* category errors were defined as utterances containing erroneous information. *Detail* errors included utterances that conveyed unimportant information. While these utterances were correct in regard to content, the information was not important to the discourse topic. *Off-topic* category errors stemmed from utterances that were not related to the topic due to being egocentric or incorrect, but did not fit into the aforementioned categories. See table 3 for a list of global coherence error violations and examples.

Coding reliability

Point-to-point interrater reliability was completed. The second author was the primary coder and the first author independently coded 33 transcripts for reliability. In total, coding reliability was completed on 35.5% of transcripts for UNIs, global coherence and error type. To calculate point-to-point reliability, the total number of correct codes (i.e., agreed upon codes) was divided by the total number of codes (e.g., UNIs, global coherence ratings, global coherence errors). The calculation produced a proportion reliable for each type of discourse code. Reliability was 93.76% for UNIs, 89.38% for global coherence ratings and 87.01% for global coherence error types. Disagreements were resolved via discussion between the first and second authors. Owing to the potentially subjective nature of global coherence error coding, point-by-point intra-rater reliability was also completed for this measure on 35.5% of transcripts with agreement of 86.49%.

Results

Research questions 1 and 2a: Do persons with fluent and non-fluent aphasia differ on utterance-level discourse measures of information content (UNIs) and global coherence?

Owing to the asymmetric distribution of the data, as determined by skewness values as well as visual inspection of histogram data, non-parametric statistics were conducted for the following results. UNIs and global coherence were evaluated across three discourse samples for 13 people with fluent aphasia and 18 people with non-fluent aphasia. The average proportion of UNIs was 0.519 (SD = 0.179) for people with fluent aphasia and 0.314 (SD = 0.213) for people with non-fluent aphasia. Comparisons using a Mann–Whitney U-test were completed to compare the average proportion of UNIs in the discourse across groups. Results indicated that the proportion of UNIs was significantly higher

Table 3. Global coherence error violations

Error type	Definition of error type	Example from the cat-in-tree picture description from Nicholas and Brookshire (1993)
Commentary	Utterances that include comments on the task or task performance	I don't know what it is
Non-specific	Utterances that are ambiguous due to an overreliance on vague/non-specific words	She had been in her little ride there
Not complete	Utterances abandoned before conveying all the required information	Firemen
Repeated	Utterances that repeat previously provided information without adding any additional information	The firemen arrive. Firemen arrive[a]
Incorrect	Utterances with erroneous information	And there was a book there that they could get higher and get the bird
Detail	Utterances that convey unimportant information	The bird is happy
Off topic	Utterances not related to the topic due to being egocentric or incorrect and do not fit into the other categories	The firemen, my wife's cousin is one

Note: [a]Second utterance would be scored as repeated.

Table 4. Group comparison of discourse measures

Discourse measures	Group	*N*	Mean	SD	d.f.	Mean rank	Sum of ranks	Mann–Whitney *U*-test	*Z*	Significance (two-tailed)
Total utterances	Fluent	13	16.83	10.35	1	16.38	213.00			
	Non-fluent	18	13.47	5.55	1	15.72	283.00	112.00	–0.200	.841
Utterance with new information (UNIs)	Fluent	13	51.86	17.87	1	21.46	279.00			
	Non-fluent	18	31.37	21.35	1	12.06	217.00	46.00	–2.84	.004*
Global coherence	Fluent	13	2.15	0.39	1	21.69	282.00			
	Non-fluent	18	1.67	0.56	1	11.89	214.00	43.00	–2.96	.002*

Notes: SD, standard deviation; d.f., degrees of freedom.
*$p < .05$.

for people with fluent aphasia than non-fluent aphasia ($U = 46.00$, $p = .004$). Regarding average global coherence scores, people with fluent aphasia produced an average of 2.15 (SD = 0.392) out of a possible score of 4, while the average for people with non-fluent aphasia was 1.67 (SD = 0.564). A Mann–Whitney *U*-test was again completed to compare groups and revealed that global coherence was significantly higher for people with fluent aphasia ($U = 43.00$, $p = .002$). See table 4 for group comparisons for UNIs and global coherence.

A partial correlation controlling for aphasia severity (i.e., WAB-R AQ) was used to evaluate the relationship between the average proportion of UNIs and the average global coherence scores. The analysis revealed a significant high positive correlation between the two measures, $r(0) = .753$, $p = .000$.

Research question 2b: Do people with fluent and non-fluent aphasia differ on the types of violations made which contribute to breakdowns in global coherence?

Utterances were evaluated for global coherence on a scale of 1–4 (Wright and Capilouto 2012), with 4 being completely on topic and 1 being not on topic (see the description in the Methods section). Utterances with scores of 1–2 were evaluated and coded for one of seven possible error types (table 3). The proportion of seven types of error violations in people with fluent and non-fluent aphasia were evaluated using Pearson chi square tests. A Bonferroni correction was completed ($\alpha = .05/7 = .007$) to reduce the risk of type one error due to multiple comparisons. Therefore, $p < .007$ were considered significant. The group with fluent aphasia produced a total of 417 utterances with a 1 or 2 global coherence score and the group with non-fluent aphasia produced 637 utterances with a 1 or 2 score. See table 5 for Pearson chi square results.

Commentary

People with fluent aphasia produced a total of 108 utterances described as *commentary*, which made-up 25.90% of their errors. People with non-fluent aphasia produced 186 utterances classified as *commentary*, which was 29.20% of their errors. This difference was not significant, $\chi^2(1) = 1.365$, $p = .243$.

Table 5. Comparison of global coherence error types

Global coherence error type	Utterance level information by group: Fluent TU = 620 N	Fluent TE = 417 %	Non-fluent TU = 722 N	Non-fluent TE = 637 %	Analyses χ^2	Analyses p
Commentary	108	25.90%	186	29.20%	1.365	0.243
Non-specific	90	21.58%	74	11.62%	19.05	0.000*
Not complete	75	17.99%	246	38.62%	50.417	0.000*
Repeated	22	5.28%	39	6.12%	0.331	0.565
Incorrect	81	19.42%	71	11.15%	13.995	0.000*
Off topic	26	6.23%	12	1.88%	13.729	0.000*
Detail	15	3.60%	9	1.41%	5.403	0.020

Notes: TU, total utterances; TE, total errors.
*$p < .007$, adjusted for multiple comparisons.

Non-specific

People with fluent aphasia produced 90 *non-specific* utterances, totalling 22.58% of the errors they produced, while people with non-fluent aphasia produced 74 *non-specific* errors, which was 11.62% of their low global coherence utterances. This difference was significant based on a Pearson chi square analysis, $\chi^2(1) = 19.50$, $p = .000$.

Not complete

A total of 17.99% of low global coherence utterances were classified as *not complete* for people with fluent aphasia (75 utterances). People with non-fluent aphasia produced 246 utterances that were *not complete*, which was 38.62% of their global coherence errors, making this the most common error type for people with non-fluent aphasia. Pearson chi square analysis revealed a significant difference, $\chi^2(1) = 50.417$, $p = .000$.

Repeated

Repeated utterances made up 5.28% of errors produced by people with fluent aphasia (22 utterances) and 6.12% of the errors produced by people with non-fluent aphasia (39 utterances). The difference was not significant, $\chi^2(1) = .331$, $p = .565$.

Incorrect

People with fluent aphasia produced 81 *incorrect* utterances which, was 19.42% of their global coherence errors. People with non-fluent aphasia produced 71 *incorrect* utterances which was 11.15% of their errors. This resulted in a significant difference, $\chi^2(1) = 13.995$, $p = .000$.

Off topic

Off-topic errors were more prevalent in people with fluent aphasia making up 6.23% of their global coherence discourse errors, with 26 off-topic utterances produced. People with non-fluent aphasia produced 12 *off-topic* utterances, which was 1.88% of their total error utterances. Pearson chi square analysis indicated that this difference was significant, $\chi^2(1) = 13.729$, $p = .000$.

Detail

Detail utterances were the least frequent error type. People with fluent aphasia produced 15 *detail* utterances, which was 3.60% of their errors and people with non-fluent aphasia produced nine *detail* utterances (1.41% of errors). The percentage of *detail* errors was not significantly different for people with fluent versus non-fluent aphasia, $\chi^2(1) = 5.403$, $p = .020$.

Discussion

The study examined differences in utterance-level discourse measures of informativeness and global coherence in persons with fluent and non-fluent aphasia during picture description. Moreover, similarities and differences in the types of violations made that contributed to breakdowns in global coherence was also investigated. We discuss the findings below.

Inconsistent with our first prediction, groups differed on the informativeness of their discourse with the fluent group producing a higher average proportion of UNIs than the non-fluent group. This is in contrast to Brookshire and Nicholas (1995) who found that while groups differed on the overall number of CIUs produced, they did not differ on per cent CIUs produced. Several variables (e.g., sample size, group makeup of aphasia type) may contribute to differences between the present findings and those of Brookshire and Nicholas and it may be that word-level measures are better able to capture the abilities of persons with non-fluent aphasia given their limited ability to construct an utterance. However, differences perhaps also highlight important issues regarding the selection of discourse outcomes and scoring discourse informativeness at the word level. It has been argued that while individual words may be relevant to a topic, they may lack coherence (Hasan 1985, del Toro *et al.* 2008). Accordingly, while word-level units provide some quantification about the amount of information conveyed, they may provide insufficient structure for the production to be coherent. Measures such as the UNIs, which consider both relevancy and coherence and assess content proportionally, provide an alternative method of assessing information content.

Regarding global coherence, we predicted that groups would not differ on overall scores but rather group differences would emerge in the types of errors which contributed to low global coherence. Contrary to our prediction, groups significantly differed in their overall global coherence scores such that persons with fluent aphasia averaged higher scores than the non-fluent group. Thus, persons with non-fluent aphasia may have greater difficulty conveying substantive units which maintain an overall theme/topic due largely to their difficulty producing complete utterances. That is not to say that impairments in maintaining discourse theme were not apparent in the fluent group. While we did not compare groups with non-brain damaged controls, the mean global coherence score for the fluent group was a 2.15 out of a possible score of 4.0. This is consistent with literature that documents deficits in global coherence in people with fluent aphasia (e.g., Christiansen 1995). Additionally, the fluent group produced more error types (e.g., commentary, non-specific, off topic, incorrect) than the non-fluent group, who primarily produced not complete and commentary errors. Thus, while the non-fluent group may have produced lower global coherence scores, both groups likely exhibit an impaired ability to maintain global coherence which may impact overall communicative competence.

Consistent with our prediction regarding global coherence violations, groups demonstrated qualitatively different patterns of errors which contributed to low global coherence scores. While the fluent group produced more utterances that contained non-specific words, erroneous information or were unrelated to the topic, the non-fluent group produced more utterances that were single words or abandoned before conveying all the required information. Thus, while the fluent group produced more substitutive errors, the non-fluent group produced errors related to omission. These results suggest possible distinct targets for intervention for improving discourse coherence in these groups such that target specificity within the context of discourse may be emphasized in persons with fluent aphasia while methods to ensure utterance completion may be emphasized in persons with non-fluent aphasia. Another important facet of these findings is that the majority of errors were not classified as 'off topic', yet overall global coherence was low. This is consistent with prior research identifying how breakdowns at the microlinguistic level (e.g., word-finding errors, syntactic errors) can negatively impact discourse macrostructure (Christiansen 1995, Andreetta and Marini 2015), and lends support to using error analysis in conjunction with other measures to construct a more complete picture of discourse functionality and identify potential targets for clinical intervention.

While the groups had some differentiating characteristics, they produced similar proportions of utterances containing commentary, repeated information and unimportant details. These results somewhat mirror those of Christiansen (1995) who found that persons with conduction and Wernicke's aphasia produced more comments during narrative discourse and more coherence violations related to repetition of propositions and irrelevant information than persons with anomic aphasia and healthy controls. We expound upon those results as we have demonstrated similar patterns of findings in persons with non-fluent aphasia as well. These violations may reflect individual responses to linguistic breakdown and/or strategies to compensate for underlying impairments by buying the speaker time (e.g., by commenting on the task or discussing factors not central to the task) or reiterating points to ensure listener comprehension. Further, commentary on task performance and the task itself was one of the most common error types in both groups, and had a substantial impact on global coherence making up 25.9% of errors for the fluent group and 29.2% in the non-fluent group. This finding indicates that the response persons with aphasia have to their language difficulties, not just the linguistic breakdowns, substantially impact their ability to construct coherent discourse.

Interestingly, despite the greater presence of multiple error types in utterances rated low in global coherence in persons with fluent aphasia, their discourse was rated as more informative and higher in overall global coherence. We demonstrated a strong, positive relationship between our utterance level measures, which supports claims by Wright and Capilouto (2012) of a relationship between lexical informativeness and coherence. The present results suggest that the ability to produce a complete utterance may be an important facet of discourse ability and that emphasis on improving discourse output could be a target for intervention in non-fluent aphasia. For instance, interventions that emphasize subject–verb–object constructions may prove more beneficial for improving some discourse abilities over those which target isolated word retrieval.

Finally, with further regard to the above-mentioned finding of the relationship between the UNIs and global coherence ratings, the result is not unexpected given there is a degree of overlap between these two measures. Indeed, the UNIs is meant to capture utterances that are both relevant and coherent and the global coherence scale evaluates a similar construct (i.e., relevance of each utterance to an overall semantic topic). These results may assist in clinical decision-making regarding their use. For instance, the global coherence scale may be used to provide more detailed analysis of the relevance of each utterance as it is a 1–4 scale, and may facilitate error analysis as low scoring utterances can be easily identified. However, the UNIs may be a better choice in clinical situations in which time is limited and a simple binary code for each utterance would be more efficient.

Results of the current study provide a foundation for understanding group differences in discourse production in fluent and non-fluent aphasia; however, there are limitations of note. An obvious limitation of this study is the inherent heterogeneous nature of aphasia. Although aphasia severity was controlled for between groups using the WAB-AQ, this single measure does not truly capture the heterogeneous nature of language breakdowns due to aphasia. Additionally, age and education were controlled for, but several other factors may have influenced performance (e.g., general naming abilities, unidentified apraxia of speech). Likewise research documents impairments in persons with aphasia in cognitive skills such as attention, memory and executive functions (Murray 2012), which may influence various aspects of discourse production in this population. For instance, Cahana-Amitay and Jenkins (2018) note that working memory impairments negatively impact global and local coherence during narrative production in persons with aphasia. Thus, while the present study did not explore the possible influence of non-linguistic cognitive abilities on discourse production, future studies should attempt to elucidate possible relationships. With regard to the statistical analyses, the use of the Mann–Whitney *U*-test, which examines differences in the ranked positions of scores in different groups, limits one's ability to interpret the magnitude of difference between groups on our outcome measures of informativeness and global coherence. However, the non-parametric distribution of the data favoured the use of the non-parametric test. Finally, the current study examined discourse in response to constrained picture-description tasks. Research demonstrates that topic, emotionality and discourse elicitation techniques may impact discourse production (e.g., Borod *et al.* 2000, Armstrong *et al.* 2016). As such, future research, such as that currently underway in our laboratory, should emphasize exploring group differences during less structured tasks (e.g., personal narratives) that are more representative of everyday talk.

Conclusions

The majority of studies on discourse in stroke aphasia have emphasized examining impairments in higher level language processing by comparing client performance with neurologically healthy controls. There is little understanding about how various aphasia types differ from one another in their discourse-level communication breakdowns, and thus limited understanding of specific assessment and intervention needs. The findings of this study suggest that individuals with fluent aphasia were rated higher in their ability to produce informative and coherent discourse in response to picture-description tasks than individuals with non-fluent aphasia. Furthermore, it demonstrates group commonalities as well as differences in their patterns of errors that contribute to low-discourse coherence. We suggest that the ability to produce a complete utterance may be an important facet of discourse ability and provide possible unique targets for assessment and intervention to improve communicative competency in these groups.

Acknowledgements

The authors thank the clients, researchers and coordinators who contribute to, and make possible, AphasiaBank. ***Declaration of interest:*** The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

References

Andreetta, S., Cantagallo, A. and Marini, A., 2012, Narrative discourse in anomic aphasia. *Neuropsychologia*, **50**(**8**), 1787–1793. https://doi.org/10.1016/j.neuropsychologia.2012.04.003

Andreetta, S. and Marini, A., 2015, The effect of lexical deficits on narrative disturbances in fluent aphasia. *Aphasiology*, **29**(**6**), 705–723. https://doi.org/10.1080/02687038.2014.979394

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

Armstrong, E., Bryant, L., Ferguson, A. and Simmons-Mackie, N., 2016, Approaches to assessment and treatment of everyday talk in aphasia. In I. Papathanasiou and P. Coppens (eds), *Aphasia and Related Neurogenic Communication Disorders*, 2nd edn (Burlington, MA: Jones & Bartlett Learning), pp. 269–285.

Berndt, R., Wayland, S., Rochon, E., Saffran, E. and Schwartz, M., 2000, *Quantitative Production Analysis: A Training Manual for the Analysis of Aphasic Sentence Production* (Hove: Psychology Press).

Borod, J. C., Rorie, K. D., Pick, L. H., Bloom, R. L., Andelman, F., Campbell, A. L., . . . Sliwinski, M., 2000, Verbal pragmatics following unilateral stroke: Emotional content and valence. *Neuropsychology*, **14**(**1**), 112–124. https://doi.org/10.1037/0894-4105.14.1.112

Brookshire, R. H. and Nicholas, L. E., 1995, Performance deviations in the connected speech of adults with no brain damage and adults with aphasia. *American Journal of Speech–Language Pathology*, **4**(**4**), 118–123. https://doi.org/10.1044/1058-0360.0404.118

Cahana-Amitay, D. and Jenkins, T., 2018, Working memory and discourse production in people with aphasia. *Journal of Neurolinguistics*, **48**, 90–103. https://doi.org/10.1016/j.jneuroling.2018.04.007

Christiansen, J. A., 1995, Coherence violations and propositional usage in the narratives of fluent aphasics. *Brain and Language*, **51**(**2**), 291–317. https://doi.org/10.1006/brln.1995.1062

Del Toro, C. M., Altmann, L. J., Raymer, A. M., Leon, S., Blonder, L. X. and Gonzalex Rothi, L. J., 2008, Changes in aphasic discourse after contrasting treatments for anomia. *Aphasiology*, **22**(**7–8**). https://doi.org/10.1080/02687030701844204

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

Ellis, C., Henderson, A., Wright, H. H. and Rogalski, Y., 2016, Global coherence during discourse production in adults: A review of the literature. *International Journal of Language and Communication Disorders*, **51**(**4**), 359–367. https://doi.org/10.1111/1460-6984.12213

Frederiksen, C. H., Bracewell, R. J., Breuleux, A. and Renaud, A., 1990, The cognitive representation and processing of discourse: Function and dysfunction. In YVES JOANETTE and HIRAM H. BROWNELL (eds), *Discourse Ability and Brain Damage* (New York, NY: Springer), pp. 69–110.

Glosser, G. and 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

Hasan, R., 1985, The texture of a text. In M. A. K. Halliday and R. Hasan (eds), *Language, Context, and Text: Aspects of Language in a Social–Semiotic Perspective* (Geelong, VIC: Deakin University Press), pp. 70–96.

Kertesz, A., 2006, *Western Aphasia Battery—Revised* (San Antonio, TX: Pearson).

Linnik, A., Bastiaanse, R. and Höhle, B., 2016, Discourse production in aphasia: A current review of theoretical and methodological challenges. *Aphasiology*, **30**(7), 765–800. https://doi.org/10.1080/02687038.2015.1113489

MacWhinney, B., 2000, *The CHILDES Project: Tools for Analyzing Talk*. 3 (Mahwah, NJ: Lawrence Erlbaum Associates).

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

Marini, A. and Urgesi, C., 2012, Please get to the point! A cortical correlate of linguistic informativeness. *Journal of Cognitive Neuroscience*, **24**(**11**), 2211–2222. https://doi.org/10.1162/jocn_a_00283

Manning, M. and Franklin, S., 2016, Cognitive grammar and aphasic discourse. *Clinical Linguistics and Phonetics*, **30**(**6**), 417–432. https://doi.org/10.3109/02699206.2015.1128981

Murray, L. L., 2012, Attention and other cognitive deficits in aphasia: Presence and relation to language and communication measures. *American Journal of Speech–Language Pathology*, **21**(**2**), S51–S64. https://doi.org/10.1044/1058-0360(2012/11-0067)

Nicholas, L. E. and Brookshire, R. H., 1993, A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *Journal of Speech, Language, and Hearing Research*, **36**(**2**), 338–350. https://doi.org/10.1044/jshr.3602.338

Olness, G. S. and Ulatowska, H. K., 2011, Personal narratives in aphasia: Coherence in the context of use. *Aphasiology*, **25**(**11**), 1393–1413. https://doi.org/10.1080/02687038.2011.599365

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

Sherratt, S., 2007, Multi-level discourse analysis: A feasible approach. *Aphasiology*, **21**(**3–4**), 375–393. https://doi.org/10.1080/02687030600911435

Wright, H. H. and Capilouto G. J., 2012, Considering a multi-level approach to understanding maintenance of global coherence in adults with aphasia. *Aphasiology*, **26**(**5**), 656–672. https://doi.org/10.1080/02687038.2012.676855

Wright, H. H., Fergadiotis G., Koutsoftas A. and Capilouto G. J., 2010, Coherence in stories told by adults with aphasia. *Procedia Social and Behavioral Sciences*, **6**, 111–112. https://doi.org/10.1016/j.sbspro.2010.08.056

Wright, H. H., Silverman, S. and Newhoff, M., 2003, Measures of lexical diversity in aphasia. *Aphasiology*, **17**(**5**), 443–452. https://doi.org/10.1080/02687030344000166