AJSLP

Research Article

An Integrative Analysis of Spontaneous Storytelling Discourse in Aphasia: Relationship With Listeners' Rating and Prediction of Severity and Fluency Status of Aphasia

Anthony Pak-Hin Kong[a] and Cherie Wan-Yin Wong[b]

Purpose: This study investigated which of the three analytic approaches of oral discourse, including linguistically based measures, proposition-based measures, and story grammar, best correlated with aphasia severity and with naïve listeners' ratings on aphasic productions. The predictive power of these analytic approaches to aphasia severity and fluency status of people with aphasia (PWA) was examined. Finally, which approach best discriminated fluent versus nonfluent PWA was determined.
Method: Audio files and orthographic transcriptions of the storytelling task "The Boy Who Cried Wolf" from 68 PWA and 68 controls were extracted from the Cantonese AphasiaBank. Each transcript was analyzed using these 3 systems.
Results: The linguistic approach of discourse analysis best correlated with aphasia severity and naïve listeners' subjective ratings. Although both linguistically based and proposition-based measures significantly predicted aphasia severity, a subset of linguistic measures focusing on the quantity and efficiency of production were particularly useful for clinical estimation of the fluency status of aphasia.
Conclusions: The linguistically based measures appeared to be the most clinically effective and powerful in reflecting PWA's performance of spoken discourse.

Discourse production is one of the essential components in clinical assessment for people with aphasia (PWA), which contributes in making accurate diagnosis, planning treatments, and evaluating treatment outcomes (Armstrong, Brady, Mackenzie, & Norrie, 2007; Kong, 2016). Various approaches in quantifying multiple aspects of narrative production have been developed. For example, quantitative production analysis (QPA; Saffran, Berndt, & Schwartz, 1989) considers the lexical content and sentence structure of narratives, whereas conversational discourse analysis (Boles, 1997) quantifies verbal output of narratives and communication strategies demonstrated by PWA and their communication partners. These approaches yield an advantage of capturing the essential clinical characteristics of PWA through structured design and standardized computation of quantification measures. However, they usually contain complex and time-consuming procedures for language sample extraction and data processing, which limits their implementation in clinical situations. Comparatively, there are some clinically friendly approaches that allow clinicians to perform language analyses in an easier, but also objective and reliable, manner. The approaches of measuring main concepts (e.g., Kong, Whiteside, & Bargmann, 2016; Nicholas & Brookshire, 1995; Richardson & Dalton, 2016), linguistic characteristics (e.g., Hilger, Ramsberger, Gilley, Menn, & Kong, 2014; Menn, Ramsberger, & Helm-Estabrooks, 1994), and story grammar (e.g., Koo, 2006; Stein & Glenn, 1979) are some examples of clinically friendly methods. These systems allow clinicians to conduct objective analysis and obtain important diagnostic information in a less complex and time-consuming way (Gao & Benson, 1990; Kong, 2016; Richardson & Hudspeth, 2014b). Most of

[a]School of Communication Sciences and Disorders, University of Central Florida, Orlando
[b]Department of Special Education and Counselling, The Education University of Hong Kong
Correspondence to Anthony Pak-Hin Kong: antkong@ucf.edu

Editor-in-Chief: Julie Barkmeier-Kraemer
Editor: Anastasia Raymer
Received January 17, 2018
Revision received May 8, 2018
Accepted June 4, 2018
https://doi.org/10.1044/2018_AJSLP-18-0015

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

Downloaded From: https://ajslp.pubs.asha.org/ by a ReadCube User on 11/26/2018
Terms of Use: https://pubs.asha.org/ss/rights_and_permissions.aspx

the above-mentioned methods have been applied to native Cantonese-speaking PWA. Although they contain some degree of overlap in terms of their linguistic aspects of assessment, currently little evidence is available as to which approach of discourse measurement is the most useful in quantifying oral discourse. This study, therefore, aimed to explore which of these approaches would have a stronger relationship with Cantonese-speaking PWA's aphasia severity and with naïve listeners' subjective ratings of these productions. The predictive power of these measures to aphasia severity and the fluency status of PWA are also examined. Before reviewing these analytic tools in detail, an introduction to assessment of Cantonese aphasia with highlights of its similarities to English is in order.

Assessing Aphasia in Cantonese

The Cantonese version of the Western Aphasia Battery (CAB; Yiu, 1992) is the only standardized assessment of the overall severity of aphasia and published for native speakers of Cantonese. Following the blueprint of the original version, the CAB covers four main domains of language skills of PWA, including spontaneous speech, auditory comprehension, repetition, and naming. It provides users with a diagnostic label of the type of aphasic syndrome and an aphasia quotient (AQ), which can generally serve as an indicator of the severity of aphasia.

Apart from formally using standardized aphasia batteries, which usually do not provide a detailed evaluation of discourse-level performance of PWA, naïve listeners' perception on the abnormality of aphasic production is another dimension that is worth exploring (Blanken, Dittmann, Grimm, Marshall, & Wallesch, 2008). The intention and motivation for PWA to communicate or to engage in social activities is greatly affected by the perceptions of PWA's significant others and people in the social circle (Harmon, Jacks, Haley, & Faldowski, 2016). Previous research has explained that a listener's perceptions of spoken output encompass (a) the speaker's language behavior, (b) attribution of the speaker, and (c) the feelings that listeners experience in response to the language output (Croteau & Le Dorze, 2001; Hallé & Le Dorze, 2014). According to Harmon et al. (2016), spoken narratives by both fluent and nonfluent PWA have often been rated by listeners as less intelligible, less comfortable for listening, and less easy to understand. PWA have also been perceived as less clear and organized in their thought as well as less intelligent and competent, which might be attributed to not only the word finding difficulties, paraphasia, and pauses identified in their production but also the generally lower degree of ideas conveyed and less information transmitted in their speech (Duffy, Boyle, & Plattner, 1980; Harmon et al., 2016). These findings based on speakers of English were found to be consistent with a recent report by Kong, Linnik, Law, and Shum (2017), who suggested that Cantonese-speaking PWA's discourse were also rated with a lower degree of understandability, clarity, and overall coherence by naïve listeners.

The Linguistic Approach of Analyzing Discourse by Chinese PWA

The linguistic approach of analyzing oral discourse allows clinicians to quantify various linguistic components, such as the syntactic, morphological, and phonological aspects of impairments. The Cantonese adaptation of QPA (Law, 2001) is one of the most influential systems for measuring the linguistic complexity and accuracy of spoken narratives. With careful consideration of the Cantonese grammar, it was found to be able to highlight the linguistic characteristics of nonfluent PWA and their difference from controls and fluent PWA in a storytelling task. A major drawback of the Cantonese QPA was its high demand on users' extent of linguistic knowledge and time on data processing.

Another tool that is more widely used in clinical settings is the Cantonese Linguistic Communication Measure (CLCM; Kong & Law, 2004). The CLCM contains four single and colored picture stimuli that are designed to be culturally appropriate for eliciting narrative productions in the Cantonese-speaking population (Kong, 2006). The index system of CLCM captures important language aspects that are likely to be disrupted by brain damage. For example, PWA tend to produce a lower variety of informative words and utter shorter and less elaborated content at a slower rate than control participants. Among all CLCM indices, the production of errors (index of error [IER]), the ability and rate in producing key information (index of communication efficiency [ICE]), and the grammatical form of speech output (index of grammatical support) were found effective in discriminating between groups of controls and PWA (Kong, 2006). The remaining indices are able to differentiate between fluent and nonfluent PWA. In short, the CLCM indices were reported to be reliable and valid and significantly correlated with verbal language subtest scores in the CAB (Kong & Law, 2004).

Detailed measures of the aphasic language form can inform the extent of an expressive deficit. In addition, clinicians may infer the pragmatic and cognitive abilities of PWA. For example, carefully selected linguistic measures were found to reflect on how well high-functioning PWA performed in daily life conversations (Kong & Law, 2009). In the study of Yu et al. (2013), it was also suggested that the linguistic functions of poststroke Chinese PWA were closely related to their cognitive orientation, spatial perception, visual perception, and thinking operation. See Kong (2017) for a more detailed review of aphasia assessment in Chinese.

The Proposition-Based Approach of Analyzing Discourse

This approach of discourse analysis focuses on how well a speaker can provide the outline of the gist or essential information portrayed in a stimulus picture or an outline of the essential steps in a procedure (Van Dijk, 1980). One example is the main concept analysis (MCA; Kong, 2009,

Downloaded From: https://ajslp.pubs.asha.org/ by a ReadCube User on 11/26/2018
Terms of Use: https://pubs.asha.org/ss/rights_and_permissions.aspx

2011) that captures the presence, accuracy, and completeness of concepts in a discourse. This method has been claimed to be a reliable means of quantifying discourse abilities through a sequential picture description task among speakers with aphasia (Kong, 2009) and dementia (Kong et al., 2016). Task-specific normative performance (and data) of unimpaired speakers are often used as a basis for comparing how far behind a PWA is. Hence, the MCA involves the construction of a "main concept checklist" drawn from a large control sample prior to conducting subsequent analysis of PWA production (Dalton & Richardson, 2015). According to Kong (2009), MCA was effective in differentiating people with or without aphasia by comparing the output of proposition. It has also been proved to be sensitive to discriminate fluent from nonfluent PWA by its accurate/complete code, absent code, and the composite score (Richardson & Hudspeth, 2014a). MCA studies in Cantonese Chinese (e.g., Kong, 2009) and English (e.g., Kong et al., 2016) also revealed significant correlation between the MCA performance and AQs of CAB (Yiu, 1992) and its original version. Similar results in Mandarin Chinese have also been reported (Kong & Yeh, 2015). MCA has been concluded as an easy-to-perform, informative, and reliable measure of discourse adequacy (Dalton & Richardson, 2015). This approach has also been illustrated for its stability of measuring oral discourse in chronic aphasia (Kong, 2011).

Using Story Grammar to Analyze Spoken Discourse

Story grammar looks at the organizational structures of stories (Stein & Glenn, 1979), which allow the formation of a cognitively based framework (or schema) that helps a speaker generate a story and guides a listener to comprehend (Mandler & Johnson, 1977). The major components of story grammar include (a) setting and episode, (b) initiating event, (c) attempt, (d) consequences/reactions, (e) abstract/internal responses, and (f) coda/plan (Richardson & Hudspeth, 2014b). This analytic approach focuses on the superstructure of a narrative production (Koo, 2006) and forms the basis for differentiation between PWA and unimpaired speakers (Richardson & Hudspeth, 2014b), as well as individuals with and without brain injuries (Lê, Coelho, Mozeiko, & Grafman, 2011). Specifically, story length and completeness of story components were suggested to be sensitive to discriminate between language-impaired and control groups (Koo, 2006; Richardson & Hudspeth, 2014b). Note that this approach has also been applied to distinguish discourse patterns in terms of cohesion and coherence (Zasler, Katz, & Zafonte, 2007). Koo (2006) speculated that word finding difficulties could lead to the reduced usage of story components, and microlinguistic deficits further contributed to impaired use of story grammar.

Aims

The three analytic approaches of oral discourse using linguistically based measures, proposition-based measures, and story grammar have been found to be sensitive in differentiating people without aphasia from PWA. They quantify narrative productions from different perspectives and, therefore, are supplementary to each other. Clinically when a language sample is obtained from PWA, given the extensive amount of time involved in data processing, it is not realistic to conduct a full-scale analysis using all the proposed variables of each approach. In addition, there is currently no consensus regarding which measures can best distinguish narratives produced by PWA and unimpaired individuals because different studies have used these measures for different spoken genres. Application of these methods to discourse samples elicited from the same task (storytelling task in our case) will allow a better investigation of the psychometric properties of these quantification variables. There are three aims in this study:

1. To investigate which of the three analytic approaches of discourse analysis best correlated with aphasia severity, as reflected by the CAB (Yiu, 1992), and with naïve listeners' ratings on PWA's productions. It was hypothesized that the proposition-based approach would best correlate with aphasia severity because Kong (2009) has reported the highest degree of relationships of MCA (four out of six parameters on a sequential picture description tasks) with CAB AQ. It was also hypothesized that the linguistic approach would best correlate with naïve listener's ratings. This is because the characteristics of language deficits in PWA (e.g., word finding difficulties, paraphasias, reduced informativeness of content) are mostly reflected under the linguistic approach, and these features might be perceived as less clear and less organized in subjective ratings (Duffy et al., 1980; Harmon et al., 2016).
2. To examine the predictive power of the three analytic approaches to aphasia severity and fluency status of PWA. It was hypothesized that the significant predictive power of the linguistically based and proposition-based approaches based on picture description tasks (Kong, 2009, 2016) would extend to the storytelling task in this study.
3. To determine which of the three analytic approaches best discriminated fluent versus nonfluent PWA based on a cluster analysis.

Method

Data Set

This study used data extracted from the Cantonese AphasiaBank (Kong & Law, 2018). In particular, audio files and orthographic transcriptions of the storytelling task "The Boy Who Cried Wolf" from 68 PWA (out of the pool of 105) and 68 controls (out of a total of 149) were used. Each PWA–control pair was age- and education-matched, based on five age groups and three education

Downloaded From: https://ajslp.pubs.asha.org/ by a ReadCube User on 11/26/2018
Terms of Use: https://pubs.asha.org/ss/rights_and_permissions.aspx

Table 1. Distribution of different age and education subgroups.

Education level	Age				
	30–39 years	40–49 years	50–59 years	60–69 years	70 years or older
0–6 years	0	0	4 (4 M)	7 (1 M + 6 F)	2 (1 M + 1 F)
7–13 years	3 (1 M + 2 F)	15 (13 M + 2 F)	22 (17 M + 5 F)	8 (6 M + 2 F)	0
> 13 years	2 (1 M + 1 F)	2 (2 M)	3 (3 M)	0	0

Note. M = male; F = female.

levels (see Table 1).[1] For the PWA, the assessment results of their language deficits as reflected in the CAB (Yiu, 1992), including the AQ and aphasic syndrome, were also used. There were in total 51 fluent (46 PWA with anomic aphasia, four PWA with transcortical sensory aphasia, one with Wernicke's aphasia) and 17 nonfluent PWA (eight PWA with Broca's aphasia, nine PWA with transcortical motor aphasia).

Stage 1—Setup of Normative Data

Previous studies reporting the analysis of storytelling of "The Boy Who Cried Wolf" in Cantonese have not applied the MCA (Kong, 2009) or CLCM (Kong & Law, 2004). Twenty-four language samples (i.e., 35%) were randomly selected from the control group to establish a normative data set, which will be the basis of quantification of PWA and control samples in Stage 2. With reference to the definition of "main concept, i.e., a statement referring to a gist of a story, with only one main verb, that is important but independent from other concepts in the same story" (Kong, 2009, p. 446), all potential main concepts were first identified from the story. A total of 13 main concepts, each having appeared at least 60% of the time in the 24 language samples, were included as the finalized list of target main concepts (see Appendix A). For each main concept, all correct and appropriate alternative lexical items that had appeared in one or more of the 24 samples were considered as acceptable replacements (Law, Kong, Lai, & Lai, 2015); they are listed in Appendix B.

Next, with reference to the definition of "informative-word" (or i-word), that is, a unit or piece of information in the form of a lexeme correctly produced for describing a (key) element of the stimulus material (Kong & Law, 2004), all potential informative-words were identified. In other words, an i-word here is the same as a content unit in the Linguistic Communication Measure in Menn et al. (1994), but different from a correct information unit, which was defined as a group of words that are "intelligible in context, accurate in relation to the picture(s) or topic, and relevant to and informative about the content of the picture(s) or the topic" (Nicholas & Brookshire, 1993, p. 348). The story was first divided into different scenarios. Scenarios that did not appear at least 60% of the time in the 24 language samples were excluded; a total of 12 scenarios were found. The key lexical items used in each final scenario were categorized into "person," "action," "object," "place," and "others" as in Kong (2006). The final list of i-words can be found in Appendix C.

Finally, a list of essential components of story grammar was compiled following the framework of Koo (2006). Any story grammar components that appeared in 60% or more of the 24 language samples were counted as essential components. The finalized list included one setting, three episodes, and one coda, which was the same as in Koo (2006). A total number of 12 essential story grammar components (see Appendix C) were drawn from these samples.

Stage 2—Processing of PWA Data

Each language sample was independently analyzed by all three previously mentioned analytic approaches:

1. Proposition-based approach: The presence, accuracy, and completeness of each identified main concept was scored following the criteria in Table 2, which were established by Kong (2009), Nicholas and Brookshire (1995), and Richardson and Hudspeth (2014a). A final main concept score (or MC score) was computed for each participant by summing the total.
2. Linguistically based approach: Each sample was analyzed using the seven indices suggested in Kong and Law (2004). First, each sample was counted for its (a) total number of words (TW) and (b) total number of i-words (IW). This was followed by computing the following variables: (c) index of lexical efficiency, computed by dividing TW by IW; (d) ICE, calculated by dividing IW by the duration of recording in minutes; (e) index of grammatical support,

[1]We acknowledge the imbalance makeup of the fluent versus nonfluent PWA groups and a more ideal case would be an even distribution of aphasia types. Note that the majority of participants in the Cantonese AphasiaBank (Kong & Law, 2018) were those with fluent aphasia, and there was a high proportion of fluent PWA with anomic aphasia. According to Kong and Law (2018), this distribution pattern was representative of the makeup of daily clinical caseload for chronic aphasia. In other words, the imbalanced aphasia type of PWA (and therefore corresponding data extracted) in the present study was primarily a result of the imbalanced pool of data for use to begin with. In addition, our careful matching of PWA–control pairs (based on age and education levels) may have also contributed to the current uneven distribution of participants because the normative data in the corpus contained similar number of young, mid-age, and older unimpaired individuals, but the PWA tended to be older. As a result, there were insufficient older controls for pairing the PWA.

Downloaded From: https://ajslp.pubs.asha.org/ by a ReadCube User on 11/26/2018
Terms of Use: https://pubs.asha.org/ss/rights_and_permissions.aspx

Table 2. Scoring criteria for analysis of main concept.

Code	Representation	Numeric value
Accurate/Complete	All essential elements of the concept were produced correctly.	3
Accurate/Incomplete	Some essential elements were produced correctly, but at least one essential element was omitted.	2
Inaccurate	One or more pieces of essential elements in the main concept were inaccurate.	1
Absent	No essential elements of the concepts were produced.	0

computed by dividing the sum of correct closed-class words and affixes by IW; (f) index of elaboration, computed by dividing the total number of correct stem morphemes by IW; and (g) IER, computed by dividing the sum of jargons, neologisms, phonemic paraphasias, and semantic paraphasias by IW.

3. Story grammar approach: Each sample was rated in the well-formedness of story grammar (Koo, 2006), with criteria specified in Table 3.

Seven (or 10% of) PWA samples were randomly selected for establishing intrarater and interrater reliability of the three sets of analyses. Specifically, the samples were reanalyzed by the second author to determine the within-rater consistency. As for across-rater consistency, we invited two independent raters (i.e., Interrater A and Interrater B) to independently analyze the 10% randomly selected samples. The two raters' scores were compared with the original analysis previously completed by the second author. Cronbach's alpha and point-to-point reliability analyses were then carried out.

Stage 3—Ratings of Spoken Samples From Naïve Listeners

A total of 25 naïve Cantonese speakers were invited to perform perceptual ratings of all the 136 extracted audio files. These listeners, including 12 men and 13 women, were all 18 years or older and did not have any background or knowledge in speech and language pathology. They were divided into four listening groups (three groups of six people and one of seven). The 136 audio files had been randomized into four different sets of sequences, one for each listening group, before the participants were asked to listen and to independently rate each file. In other words, the listeners did not communicate with each other while rating the samples.

A questionnaire was used for naïve listeners' perceptual rating of the PWA production in three aspects (Harmon et al., 2016), including (a) behavioral (pertaining to the speech), (b) cognitive (pertaining to thoughts about the speaker), and (c) affective (pertaining to feelings). Specifically, the questionnaire consisted of five questions (see Table 4), with two statements concerning speech output, two statements regarding a speaker's attributes, and one statement related to a listener's feeling. A 5-point Likert scale (strongly agree, agree, neutral, disagree, and strongly disagree) was adopted for the rating. Points were allocated in a descending order, with 5 points for *strongly agree* and 1 point for *strongly disagree*. The five statements were presented in both English and Chinese to the participants, who were required to provide ratings immediately after listening to each audio recording.

The listening session lasted for an average of 3.5 hr, including a 15-min introduction and a 5-min break. The participants were introduced to the questionnaire at the beginning of the session. The second author then demonstrated the scoring procedures (with reference to the five statements) using three different audio files selected from the Cantonese AphasiaBank. This allowed the participants to be familiarized with the procedure and to establish a common rating standard. Participants could also raise any inquiries in these practice trials.

Stage 4—Statistical Analysis

A Pearson's *r* correlation was conducted to examine the relationships of the three analytic approaches with the AQ of CAB and naïve listeners' ratings. A multiple regression analysis was then performed to determine the predictive power of the three approaches to AQ of CAB and fluent and nonfluent type of aphasia. To be specific, the variables of IW and IER were computed separately with total score of main concept approach and story grammar

Table 3. Scoring criteria for analysis of story grammar.

Code	Representation	Numeric value
Well-formed	The story appeared with all the essential components.	3
Adequately formed	The story appeared with setting and > 50% of essential story grammar components in each episode.	2
Poorly formed	The story appeared without setting and/or < 50% of essential story grammar components in each episode.	1
Undefined	The story appeared with none of the essential components.	0

Downloaded From: https://ajslp.pubs.asha.org/ by a ReadCube User on 11/26/2018
Terms of Use: https://pubs.asha.org/ss/rights_and_permissions.aspx

Table 4. Questionnaire statements and corresponding categories for naïve listeners' perceptual rating of audio files.

Number	Statement (English/Chinese)	Category
1	This person told the story completely. 這個人講出了一個完整故事。	Speech output
2	I would feel comfortable listening to stories told by this person. 聽這個人講故事令我感到舒服。	Listener's feeling
3	I think this person has a clear mind and organization. 我覺得這個人有清晰的思維和組織力。	Speaker's attributes
4	This person's story contained a lot of errors/inaccurate information. 這個人的故事內容有許多錯誤資料。	Speech output
5	This person's speech did not sound like a story. 這個人的講話不像一個故事。	Speaker's attributes

approach in regressions. Finally, a cluster analysis was conducted to examine the distribution of the fluent and nonfluent aphasia groups. All statistical analyses were run through SPSS, referring to procedures and interpretation suggested in Landau and Everitt (2004).

Results

Concerning Research Question 1, the results of Pearson's *r* showed that all three analytic approaches were significantly correlated with PWA's aphasia severity (see Table 5). Among all factors in the three approaches, the IER of the linguistically based approach yielded the highest correlation with aphasia severity, $r(62) = .67$, $p \leq .01$, this was followed by the IW, $r(67) = .65$, $p \leq .01$. In addition, the three approaches were significantly correlated to the perceptual ratings of naïve listeners. Concerning the completeness of story (Question 1 of questionnaire), IW yielded the highest correlation, $r(67) = .86$, $p \leq .01$, and the total MC score came with the second highest correlation, $r(68) = .78$, $p \leq .01$. With regard to the subjective judgment of comfortableness (Question 2 of questionnaire), the ICE yielded the highest correlation, $r(67) = .85$, $p \leq .01$, and this result was comparatively higher than other factors. As for making inferential judgment on the speaker's clarity and organization (Question 3 of questionnaire), the IW again yielded the highest correlation, $r(67) = .81$, $p \leq .01$. Concerning the inaccuracy of information delivered in the story from the listeners' perspective (Question 4 of questionnaire), the more the IW in a story, $r(67) = -.82$, $p \leq .01$, and the higher the communication efficiency (ICE) of a story was delivered, $r(67) = -.76$, $p \leq .01$, the lower the score on inaccuracy was rated by the listeners. Finally, judging a speech's likeness of a story (Question 5 of questionnaire), the IW, $r(67) = -.83$, $p \leq .01$, and ICE, $r(67) = -.80$, $p \leq .01$, were also the strongest correlation factors determining the listeners' rating. In other words, listeners tended to feel that output sounded like a story when there were more i-words produced and when the communication efficiency was higher.

Concerning the predictive powers of the three approaches to aphasia severity (Research Question 2), the results of multiple regression for predicting aphasia severity was significant (see Tables 6 and 7). Most predictors, including total MC score, $\beta = .25$, $t(58) = 2.07$, $p < .005$, IER, $\beta = -.54$, $t(58) = -6.10$, $p < .01$, and IW, $\beta = -.61$, $t(63) =$

Table 5. Correlations between the three analytic approaches and aphasia quotient of the Cantonese version of Western Aphasia Battery as well as naïve listeners' perceptual ratings.

	Pearson's *r*								
		CLCM indices							
	MC_Total	TW	IW	ILE	ICE	IGS	IEL	IER	SG_Total
Cantonese version of Western Aphasia Battery									
Aphasia quotient	.576**	.434**	.645**	−.598**	.558**	.580**	.421**	−.669**	.538**
Naïve listeners' perceptual ratings									
Q1	.784**	.574**	.864**	−.441**	.779**	.670**	.309*	−.328**	.742**
Q2	.703**	.517**	.810**	−.449**	.845**	.645**	.284*	−.335**	.690**
Q3	.696**	.530**	.814**	−.427**	.790**	.646**	.339**	−.333**	.679**
Q4	−.748***	−.530**	−.821**	.439**	−.764**	−.708**	−.336**	.327**	−.693**
Q5	−.725**	−.558**	−.826**	.419**	−.803**	−.658**	−.313*	.334**	−.686**

Note. MC_Total = total score in main concept measurement; CLCM = Cantonese Linguistic Communication Measure; TW = total number of words; IW = total number of i-words; ILE = index of lexical efficiency; ICE = index of communication efficiency; IGS = index of grammatical support; IEL = index of elaboration; IER = index of error; SG_Total = total score in story grammar measurement.

*$p \leq .05$. **$p \leq .01$. ***$p \leq .001$.

Downloaded From: https://ajslp.pubs.asha.org/ by a ReadCube User on 11/26/2018
Terms of Use: https://pubs.asha.org/ss/rights_and_permissions.aspx

Table 6. Results of multiple regression of MC_Total, IER, and SG_Total.

Independent variables	Dependent variables	
	AQ of CAB (% of variances)	Fluent or nonfluent type of aphasia
MC_Total	0.043* (33.2%)	0.058 (23.6%)
IER	0.000** (44.7%)	0.019* (17.1%)
SG_Total	0.122 (28.9%)	0.276 (22.2%)
r^2	.60 (60%)	.34 (34%)

Note. MC_Total = total score in main concept measurement; IER = index of error; SG_Total = total score in story grammar measurement; AQ = aphasia quotient; CAB = Cantonese version of Western Aphasia Battery.

*$p \leq .05$. **$p \leq .01$.

-2.60, $p < .05$, were significant, except for story grammar, $\beta = -.19$, $t(58) = 1.57$, $p > .05$. The predictive power of IER, $\beta = .27$, $t(58) = 2.42$, $p < .05$, and IW, $\beta = -.66$, $t(63) = -2.74$, $p < .01$, was significant for fluency type of PWA.

A cluster analysis was conducted for each parameter of the three analytic approaches with aphasia severity and fluency status of aphasia (Research Question 3). At the cutoff point of 95% confident interval, clusters were found for two indices of the linguistically based approach, IER and index of lexical efficiency (see Figure 1). In particular, the cluster of fluent PWA formed in the lower error and better lexical efficiency range of performance.

Interrater and Intrarater Reliability

The results of Cronbach's alpha (see Table 8) indicated a good intrarater (0.82) and interrater reliability (0.81 and 0.82 for two different raters) as well as internal consistency among the scoring of parameters in these three approaches. The point-to-point intrarater agreement for the proposition-based, linguistically based, and story grammar approach was 91%, 85%, and 93%, respectively. The point-to-point interrater agreement was slightly lower: 88%, 82%, and 90%, respectively. Post hoc review of the raw data major discrepancies in both the proposition-based and linguistically based approaches occurred in two types of utterances: (a) relatively fluent sentences that contained a target word but at the same time a lot of repeated or unrelated information and (b) nonfluent productions that contained excessive pausing, struggling, and fillers.

Discussion

Consistent with previous reports in the literature, the proposition-based (Kong, 2009;), linguistically based (Kong & Law, 2004), and story grammar (Koo, 2006) approaches of discourse analysis were found to be significantly correlated to PWA's aphasia severity. Moreover, as shown in Table 5, measures of these three approaches were highly correlated to listeners' ratings in terms of PWA's speech output (e.g., total MC score and listeners' perception of story completeness or the association of IER with listeners' identification of erroneous/inaccurate information), PWA's attributes of production (e.g., ICE and listeners' perception of story clarity/organization or the association of Total score in story grammar with listeners' perception of story-like production), and listeners' perceived comfortability of listening.

The clinical values of these approaches have been addressed in previous analyses of PWA's oral discourse, but typically they were independently investigated (Kong, 2016, 2017). The CAB (Yiu, 1992) utilizes a language sample elicited from a single picture description task and a PWA's responses to six personal questions as the basis for clinicians' subjective rating of overall fluency and information content of spoken output. Specifically, the only criterion for dichotomizing fluency status using the CAB rating

Table 7. Results of multiple regression of MC_Total, IW, and SG_Total.

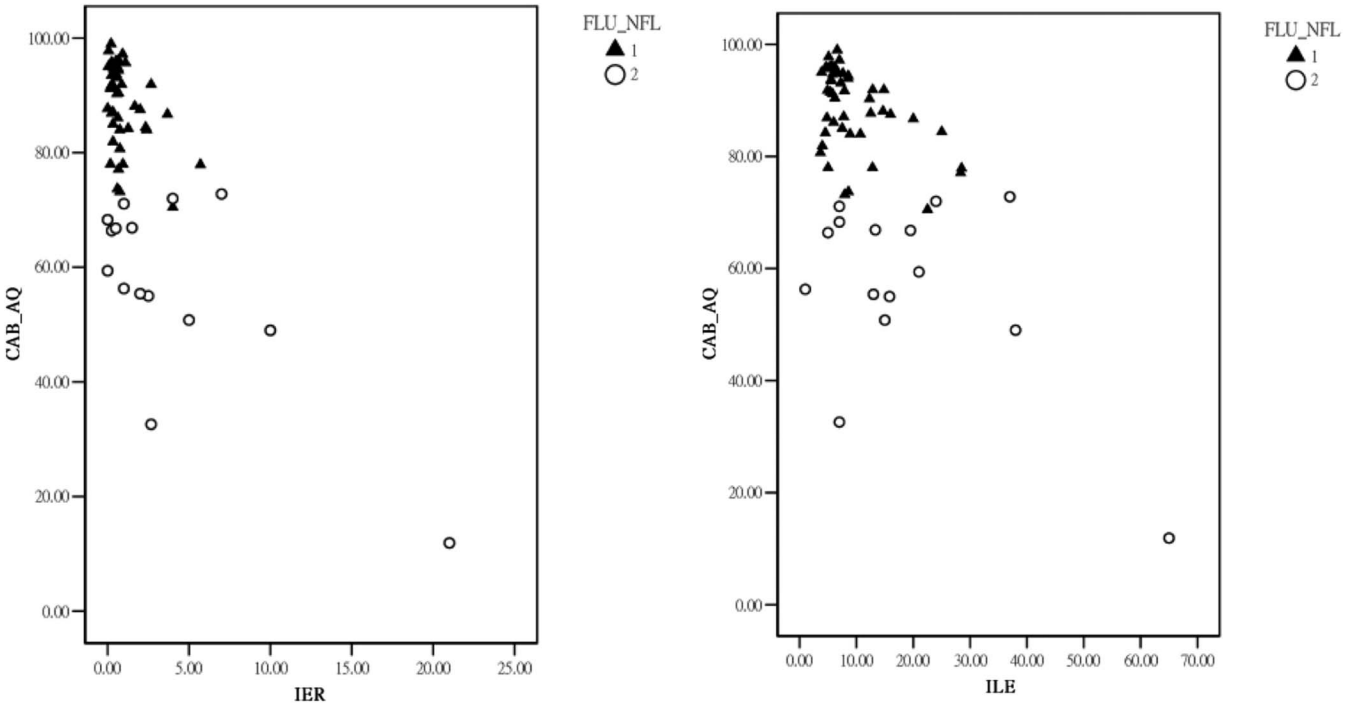
Independent variables	Dependent variables	
	AQ of CAB (% of variances)	Fluent or nonfluent type of aphasia
MC_Total	0.527 (33.2%)	0.787 (23.6%)
IW	0.011* (41.6%)	0.008** (34.9%)
SG_Total	0.719 (28.9%)	0.454 (22.2%)
r^2	.42 (42%)	.35 (35%)

Note. MC_Total = total score in main concept measurement; IW = total number of i-words; SG_Total = total score in story grammar measurement; AQ = aphasia quotient; CAB = Cantonese version of Western Aphasia Battery.

*$p \leq .05$. **$p \leq .01$.

Downloaded From: https://ajslp.pubs.asha.org/ by a ReadCube User on 11/26/2018
Terms of Use: https://pubs.asha.org/ss/rights_and_permissions.aspx

Figure 1. Cluster analysis of fluency status of participants with aphasia, with ≥ 95% confident interval. The triangle dots (1) represent fluent aphasia and the circle dots (2) represent nonfluent aphasia. CAB_AQ = aphasia quotient of Cantonese version of Western Aphasia Battery; IER = index of error; ILE = index of lexical efficiency; FLU_NFL = fluent vs. nonfluent.



scale was whether a speaker had produced at least one complete sentence; this method of making a clinical diagnosis is crude and insufficient for informing the discourse-specific symptoms of PWA. The AQ is also an overall estimation of aphasia severity based on the PWA's multiple aspects of performance such as spontaneous speech, naming, repetition, and auditory comprehension. We argue here that the results from the three discourse-specific analyses based on storytelling tasks can better, more objectively, and sensitively reflect a PWA's narrative skills. The integrative nature of this study in examining these various variables in concert was clinically innovative. Specifically, two linguistic measures, namely IER and IW, were found to best correlate with the AQ of CAB. In particular, the IER was the strongest parameter in delineating aphasia severity at the narrative discourse level, which provided further support to Kong's (2006) report of error production being the most discriminative index between the people with and without aphasia. In addition, current findings suggested that agrammatism and errors might be more manifested in narrative productions, as compared to elicited sentence production in description tasks (Coppens & Patterson, 2017; Kok, van Doorn, & Kolk, 2007). Furthermore, the fact that we found IW to be highly correlated to PWA's severity was clinically relevant. According to Kong (2006), an i-word served as a foundation in determining the grammatical support and elaboration of a narrative production because it must be used with correct closed and/or open class morphemes to form a proposition of a discourse. This property is also applicable to quantifying i-words in telling a story in this study. The MC score we used for the proposition-based analysis of storytelling did not yield the high level of correlation as in Kong (2009), as expected originally, that utilized a picture description task. It has been suggested that the cognitive–linguistic demand for narrative production was higher than that in picture description (Longacre, 1996). The story grammar approach, on the other hand, was shown to be the least related to AQ of CAB, possibly owing to its sole focus on superstructure, such as setting, episodes, or coda (Koo, 2006), which did not directly account for the aphasic language characteristics in its analysis. All in all, out of the three approaches, the linguistically based analysis seemed to be the most powerful in measuring discourse-level performance in PWA. This comprehensive information should be treated as important supplementary (or additional) clinical data to guide the clinician's diagnosis of aphasia.

Table 8. Intrarater and interrater reliability.

Independent variables	Cronbach's α	Correlation[a]
Intrarater	.82	Good
Interrater A	.82	Good
Interrater B	.81	Good

[a]Based on Landau and Everitt (2004).

Downloaded From: https://ajslp.pubs.asha.org/ by a ReadCube User on 11/26/2018
Terms of Use: https://pubs.asha.org/ss/rights_and_permissions.aspx

To deepen our understanding of the relationship between PWA's discourse deficits and naïve listeners' perception, this study included a large number of laymen raters to make judgments on aspects such as completeness of production, ease and degree of understanding the content, and accuracy of output. Of the three discourse analytic approaches, the linguistically based approach was found to correlate the best with naïve listeners' subjective ratings. In particular, Statements 1 and 4 of the questionnaire were related to the speakers' output on completeness and accuracy. From the perspective of naïve listeners, it appeared that a higher number of i-words produced and total MC score could lead to a better perception of PWA's completeness and accuracy of production. In addition to these two parameters, the rate of how quickly information is provided (ICE) was also important because it affected how listeners perceived the naturalness of the story being delivered. This result provided further insights as to how a "long but with little content" production, commonly found in PWA's discourse due to paraphasias or neologisms, might reduce listeners' overall judgment of the content accuracy. In short, our study is consistent with Andreetta, Cantagallo, and Marini (2012), who claimed the close connection of PWA's slow speech rate, reduced mean length of utterance, and poorer retrieval of precise lexical units in spoken narratives.

It is worth mentioning that Statements 3 and 5 of the questionnaire addressed PWA speakers' attributes from the listeners' perspective. The number of i-word and ICE were, again, the strongest parameters that correlated to naïve listeners' judgment on the speakers' clarity and organization, as well as the likeness of story. These findings paralleled previous conclusions that linguistic measures could best reflect the deficits of PWA's language output in terms of word finding difficulties, paraphasias, and reduced informativeness of content (Duffy et al., 1980; Harmon et al., 2016). The subjective ratings of PWA's less clear and unorganized production could also be attributed to the interruptions after frequent pauses, lexical fillers, and repetitions owing to lexical retrieval difficulties (Andreetta et al., 2012). Although the completeness, accuracy, cohesion, as well as coherence may also affect the degree of clarity and organization of a narrative production (Kong et al., 2017), the linguistic measures of IW and efficiency of producing the narrative seemed to play a greater role for naïve listeners' subjective judgment on and feeling of comfortability toward output quality, outbidding the MC scores or story grammar approach. In short, our findings suggested that error production in storytelling captured by an objective (linguistic) analysis could better delineate aphasia severity at the discourse level. From the perspective of subjective or perceptual judgment on narrative productions, listeners tended to rate an output based on the common understanding of what content should appear in the story (i.e., IW) and how naturally the story was delivered (i.e., ICE), rather than digging into the errors identified.

The linguistically based approach yielded the best predictive power among the three analytic systems. Specifically, the IER, a variable that accounted for the number of jargons, neologisms, and paraphasias in PWA's discourse, was the most powerful in predicting aphasia severity. These characteristics have historically been suggested to discriminate PWA from those without aphasia (e.g., MacWhinney, Fromm, Holland, Forbes, & Wright, 2010; Ulatowska, Freedman-Stern, Doyel, Macaluso-Haynes, & North, 1983). Somewhat unexpected was the relatively lower (but still statistically significant) correlations between IER and naïve listeners' perceptual ratings. Given that errors and referential ambiguity could be traced from people with mild and moderate aphasia and the reduction of sentential language and increase of errors were more prominent in people with more severe aphasia (Light, 1993), it is argued that IER when combined with measures such as IW or MC score that are connected with production content is clinically useful for predicting PWA severity. Although it still remains open relative to the best or the most clinically useful measures for evaluating aphasic oral discourse, with reference to our current findings (i.e., the linguistic approach to analyzing discourse demonstrated the best association with aphasia severity and fluency), clinicians are recommended to pay attention to linguistic parameters such as the efficiency and quantity of content as well as production of errors, despite the extra time and effort on clinical application. For decades, practicing clinicians have been reluctant to use discourse analysis for clinical purposes such as making diagnoses or measuring treatment outcomes. This limited application has always been related to the efficiency of analysis. Until the method of discourse quantification is further refined, this work may be destined to remain in the realm of interesting research. We argue here that the pioneering work in discourse analysis presented here can be used as the first step to directly address and explore the clinical efficiency of implementing discourse analysis.

As for the predictive power of PWA's fluency status (fluent vs. nonfluent), our results revealed that the IW produced was the most powerful parameter. Although the MCA (Kong, 2009) has been reported to be useful in differentiating fluent and nonfluent PWA, the current investigation did not find a significant predictive power of the MC score. A possible reason could be the difference in discourse genre type—a picture description in Kong (2009) versus a storytelling task here. Interestingly, measurement of story grammar was the least powerful in predicting the severity and fluency status of aphasia, although its evaluation correlated with naïve listeners' perceptual judgment of discourse impairments. Note that the fluent PWA group also demonstrated higher AQ scores, which may be considered as a potential confounding factor; however, the cluster pattern of nonfluent PWA as shown in Figure 1 indicated a more heterogeneous distribution. Further investigations are warranted to examine how discourse symptoms in aphasia are manifested across different degrees of severity and/or different aphasic syndromes.

There are two major limitations in this study. First, we currently only focused on one storytelling task, that is, "The Boy Who Cried Wolf." It is unclear if the existing findings would be equally applicable to other genres, such

Downloaded From: https://ajslp.pubs.asha.org/ by a ReadCube User on 11/26/2018
Terms of Use: https://pubs.asha.org/ss/rights_and_permissions.aspx

as procedural discourse or sequential picture descriptions. Knowing that all these narrative tasks are commonly used by clinicians for formal as well as informal clinical evaluations, a follow-up study is in progress, and hopefully the end results can provide additional insights regarding the psychometric properties of discourse analysis. Second, the number of participants for each type of aphasia was uneven in this research. Our participants were dominated by the anomic type, and other aphasic syndromes (such as Broca's, transcortical sensory, transcortical motor, isolation, Wernicke's, or global aphasia) occupied a very small proportion. Limited variety of aphasia types and uneven fluent and nonfluent PWA might have adversely affected the grouping in cluster analysis. This problem may be avoided in the future by ensuring more balanced distribution of PWA participants.

Conclusion

In summary, the linguistically based approach of discourse analysis was found to best correlate with aphasia severity and naïve listeners' perceptual ratings of the production. The linguistically based and proposition-based approaches were also promising in predicting aphasia severity, but a subset of linguistic measures focusing on the quantity and efficiency of production were particularly useful for clinical estimation of the fluency status of aphasia.

Acknowledgments

This study was supported by a grant funded by the National Institutes of Health to Anthony Pak-Hin Kong (PI) and Sam-Po Law (Co-I; Project NIH-R01-DC010398). Special thanks to the staff members in the following organizations (in alphabetical order) for their help in participant recruitment: Christian Family Service Center (Kwun Tong Community Rehabilitation Day Center), Community Rehabilitation Network of The Hong Kong Society for Rehabilitation, Internal Aphasia Clinic at the University of Hong Kong, Hong Kong Stroke Association, Lee Quo Wei Day Rehabilitation and Care Centre of The Hong Kong Society for Rehabilitation, and Self-Help Group for the Brain Damaged. The authors would also like to acknowledge the invaluable contribution of the people living with aphasia and all listeners who participated in this study.

References

Andreetta, S., Cantagallo, A., & Marini, A. (2012). Narrative discourse in anomic aphasia. *Neuropsychologia, 50*(8), 1787–1793.

Armstrong, L., Brady, M., Mackenzie, C., & Norrie, J. (2007). Transcription-less analysis of aphasic discourse: A clinician's dream or a possibility? *Aphasiology, 21*(3/4), 355–374.

Blanken, G., Dittmann, J., Grimm, H., Marshall, J. C., & Wallesch, C.-W. (2008). *Linguistic disorders and pathologies: An international handbook*. Berlin, Germany: De Gruyter Mouton.

Boles, L. (1997). Conversational analysis as a dependent measure in communication therapy with four individuals with aphasia. *Asia Pacific Journal of Speech, Language, and Hearing, 2,* 43–61.

Coppens, P., & Patterson, J. (2017). *Aphasia rehabilitation: Clinical challenges.* Burlington, MA: Jones & Bartlett Learning.

Croteau, C., & Le Dorze, G. (2001). Spouses' perceptions of persons with aphasia. *Aphasiology, 15,* 811–825. https://doi.org/10.1080/02687040143000221

Dalton, S. G., & Richardson, J. D. (2015). Core-lexicon and main-concept production during picture sequence description in non-brain-damaged adults and adults with aphasia. *American Journal of Speech-Language Pathology, 24,* 923–938.

Duffy, J. R., Boyle, M., & Plattner, L. (1980). Listener reactions to personal characteristics of fluent and nonfluent aphasic speakers. In R. H. Brookshire (Ed.), *Clinical aphasiology: Proceedings of the conference* (pp. 117–126). Minneapolis, MN: BRK Publishers.

Gao, S., & Benson, F. (1990). Aphasia after stroke in native Chinese speakers. *Aphasiology, 4,* 31–43.

Hallé, M. C., & Le Dorze, G. (2014). Understanding significant others' experience of aphasia and rehabilitation following stroke. *Disabilities and Rehabilitation, 36*(21), 1774–1782. https://doi.org/10.3109/09638288.2013.870608

Harmon, T. G., Jacks, A., Haley, K. L., & Faldowski, R. A. (2016). Listener perceptions of simulated fluent speech in nonfluent aphasia. *Aphasiology, 30*(8), 922–942. https://doi.org/10.1080/02687038.2015.1077925

Hilger, A., Ramsberger, G., Gilley, P., Menn, L., & Kong, A. P. H. (2014). Analysing speech problems in a longitudinal case study of logopenic variant PPA. *Aphasiology, 28*(7), 840–861. https://doi.org/10.1080/02687038.2014.895974

Kok, P., van Doorn, A., & Kolk, H. (2007). Inflection and computational load in agrammatic speech. *Brain and Language, 102*(3), 273–283.

Kong, A. P. H. (2006). *A Cantonese Linguistic Communication Measure for evaluating aphasic narrative production* (Unpublished doctoral dissertation). The University of Hong Kong, Hong Kong SAR, China.

Kong, A. P. H. (2009). The use of main concept analysis to measure discourse production in Cantonese-speaking persons with aphasia: A preliminary report. *Journal of Communication Disorders, 42,* 442–464.

Kong, A. P. H. (2011). Family members' report on speech-language pathology and community services for persons with aphasia in Hong Kong. *Disability and Rehabilitation, 33,* 2633–2645.

Kong, A. P. H. (2016). *Analysis of neurogenic disordered discourse production: From theory to practice*. New York, NY: Psychology Press.

Kong, A. P. H. (2017). Speech-language services for Chinese-speaking people with aphasia (C-PWA): Considerations for assessment and intervention. *Perspectives of the ASHA Special Interest Groups, 2*(3), 100–109.

Kong, A. P. H., & Law, S. P. (2004). A Cantonese Linguistic Communication Measure for evaluating aphasic narrative production: Normative and preliminary aphasic data. *Journal of Multilingual Communication Disorders, 2,* 124–146.

Kong, A. P. H., & Law, S. P. (2009). A linguistic communication measure for monitoring changes in Chinese aphasic narrative production. *Clinical Linguistics & Phonetics, 23*(4), 255–269.

Kong, A. P. H., & Law, S. P. (2018). Cantonese AphasiaBank: An annotated database of spoken discourse and co-verbal gestures by healthy and language-impaired native Cantonese speakers. *Behavior Research Methods*. Advance online publication. https://doi.org/10.3758/s13428-018-1043-6

Kong, A. P. H., Linnik, A., Law, S., & Shum, W. (2017, March). Measuring discourse coherence in anomic aphasia using Rhetorical Structure Theory. *International Journal of Speech-Language Pathology, 17,* 1–16. https://doi.org/10.1080/17549507.2017.1293158

Downloaded From: https://ajslp.pubs.asha.org/ by a ReadCube User on 11/26/2018
Terms of Use: https://pubs.asha.org/ss/rights_and_permissions.aspx

Kong, A. P. H., Whiteside, J., & Bargmann, P. (2016). The main concept analysis: Validation and sensitivity in differentiating discourse produced by unimpaired English speakers from individuals with aphasia and dementia of Alzheimer type. *Logopedics Phoniatrics Vocology, 41*(3), 129–141. https://doi.org/10.3109/14015439.2015.1041551

Kong, A. P. H., & Yeh, C. C. (2015). A Taiwanese Mandarin Main Concept Analysis (TM-MCA) for quantification of aphasic oral discourse. *International Journal of Language & Communication Disorders, 50*(5), 580–592. https://doi.org/10.1111/1460-6984.12157

Koo, C. Y. (2006). *Narrative production of Cantonese aphasic patients* (Unpublished doctoral dissertation). The University of Hong Kong, Hong Kong SAR, China.

Landau, S., & Everitt, B. S. (2004). *A handbook of statistical analyses using SPSS*. Boca Raton, FL: CRC Press.

Law, S. P. (2001). A quantitative analysis of Cantonese aphasic production. *Journal of Psychology in Chinese Society, 2,* 211–237.

Law, S. P., Kong, A. P. H., Lai, L. W. S., & Lai, C. (2015). Effects of context and word class on lexical retrieval in Chinese speakers with anomic aphasia. *Aphasiology, 29*(1), 81–100.

Lê, K., Coelho, C., Mozeiko, J., & Grafman, J. (2011). Measuring goodness of story narratives: Implications for traumatic brain injury. *Aphasiology, 25*(6–7), 748–760.

Light, L. L. (1993). Language changes in old age. In G. Blanken, J. Dittmann, H. Grimm, J. C. Marshall, & C.-W. Wallesch (Eds.), *Linguistic disorders and pathologies* (pp. 900–918). Berlin, Germany/New York, NY: de Gruyter.

Longacre, R. E. (1996). *The grammar of discourse: Topics in language and linguistics*. New York, NY: Plenum.

MacWhinney, B., Fromm, D., Holland, A., Forbes, M., & Wright, H. (2010). Automated analysis of the Cinderella story. *Aphasiology, 24*(6–8), 856–868. https://doi.org/10.1080/02687030903452632

Mandler, J. M., & Johnson, N. S. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology, 9,* 111–151.

Menn, L., Ramsberger, G., & Helm-Estabrooks, N. (1994). A linguistic communication measure for aphasic narratives. *Aphasiology, 8,* 343–359.

Nicholas, L. E., & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *Journal of Speech and Hearing Research, 36,* 338–350.

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

Richardson, J. D., & Hudspeth, S. G. (2014a). *Main concept production in persons with aphasia: A comparison of subtypes*. Paper presented at Clinical Aphasiology Conference: Clinical Aphasiology Conference (44th), St. Simons Island, GA.

Richardson, J. D., & Hudspeth, S. G. (2014b). *Story grammar analysis in persons with mild aphasia*. Paper presented at Clinical Aphasiology Conference: Clinical Aphasiology Conference (44th), St. Simons Island, GA.

Richardson, J. D., & Dalton, S. C. (2016). Main concepts for three different discourse tasks in a large non-clinical sample. *Aphasiology, 30*(1), 45–73.

Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (1989). The quantitative analysis of agrammatic production: Procedure and data. *Brain and Language, 37,* 440–479.

Stein, N. L., & Glenn, C. G. (1979). An analysis of story comprehension in elementary school children. In R. O. Freedle (Ed.), *New directions in discourse processing* (pp. 53–120). Norwood, NJ: Ablex.

Ulatowska, H. K., Freedman-Stern, R., Doyel, A. W., Macaluso-Haynes, S., & North, A. J. (1983). Production of narrative discourse in aphasia. *Brain and Language, 19*(2), 317–334.

Van Dijk, T. A. (1980). *Macrostructures: An interdisciplinary study of global structures in discourse, interaction, and cognition*. Hillsdale, NJ: Lawrence Erlbaum Associates Inc.

Yiu, E. M. L. (1992). Linguistic assessment of Chinese-speaking aphasics: Development of a Cantonese aphasia battery. *Journal of Neurolinguistics, 7,* 379–424.

Yu, Z., Jiang, S., Bi, S., Li, J., Lei, D., & Sun, L. (2013). Relationship between linguistic functions and cognitive functions in a clinical study of Chinese patients with post-stroke aphasia. *Chinese Medical Journal, 126*(7), 1252–1256.

Zasler, N., Katz, D., & Zafonte, R. D. (2007). *Brain injury medicine: Principles and practice*. New York, NY: Demos Medical Publishing.

Downloaded From: https://ajslp.pubs.asha.org/ by a ReadCube User on 11/26/2018
Terms of Use: https://pubs.asha.org/ss/rights_and_permissions.aspx

Appendix A

Finalized List of Target Main Concepts for Storytelling of "The Boy Who Cried Wolf" in Cantonese

MC1	A shepherd boy was **keeping** some sheep 有 個 牧童 **放** 羊。
MC2	The shepherd boy **felt** bored 牧童 **覺得** 好悶。
MC3	The shepherd boy **cried** "Wolf!" 牧童就 **嗌**:「狼來了!」
MC4	The villagers **ran up** the hill 村民就 **上** 山。
MC5	The villagers **saved** the shepherd boy 村民 **救** 牧童。 The villagers **drove away** the wolf 村民 **趕** 狼。
MC6	But (the villagers) **found** no wolves 點知 **發現** 無 狼。 But the shepherd boy was **telling** a lie 點知 牧童 **講** 大話。
MC7	The shepherd boy **laughed** 牧童 哈哈**笑**。 The villagers were **tricked** by the shepherd boy 村民 被牧童 **整蠱**。
MC8	The villagers were angry and **left** 村民 好嬲嘅 **走**。
MC9	The wolf really **came** 真係 有 狼 **來** 了。
MC10	The shepherd boy **cried** "Wolf!" 牧童 **嗌**:「狼來了!」
MC11	The villagers did not **help** the shepherd boy 村民 無 **幫** 牧童。
MC12	The wolf **killed** all the shepherd boy's sheep 狼 **食**晒 牧童 啲 羊。 All the shepherd boy's sheep were **killed** by the wolf 牧童 啲 羊就 俾狼 **食**晒。
MC13	The shepherd boy was sad and **returned** home 牧童 好傷心嘅 **返** 屋企。 The shepherd boy was sad and **went down** the hill 牧童 好慘嘅 **落** 山。

Note. The main verb for each main concept is bolded. All the essential information is underlined.

Downloaded From: https://ajslp.pubs.asha.org/ by a ReadCube User on 11/26/2018
Terms of Use: https://pubs.asha.org/ss/rights_and_permissions.aspx

Appendix B

Lexical Items That Are Commonly Accepted as Alternatives for Storytelling of "The Boy Who Cried Wolf" in Cantonese

Main concept	Target lexical item	Acceptable alternatives
MC1, 2, 3, 5, 6, 7, 10, 11, 12, 13	牧童	放羊嘅小朋友
	Shepherd boy	The child who herded sheep
MC1	放	看/睇
	Keeping	Herding/Looking after
MC3, 10	嗌	話/大叫
	Cried	Said/Shouted
MC4	上	去
	Ran up	Went to
MC5	趕	扑/打
	Drove away	Knocked down/Beat
MC5	救	幫
	Saved	Helped
MC6	點知	殊不知/原來
	But	It never came to my mind that…/It turned out…
MC8	走	返去
	Left	Went back
MC8	嬲	唔開心/失望
	Angry	Unhappy/Disappointed
MC10	嗌	話/大叫
	Cried	Said/Shouted
MC10	狼來了	救命
	"Wolf!"	"Help!"
MC11	幫	信/理/救/睬
	Help	Believe/Answer/Save/Attend to
MC12	食	搶/捉
	Killed	Took away/Caught
MC13	返/落	喊
	Returned/Went down	Wept
MC13	屋企	村
	Home	Village
MC13	傷心/慘	後悔
	Sad	Regret

Downloaded From: https://ajslp.pubs.asha.org/ by a ReadCube User on 11/26/2018
Terms of Use: https://pubs.asha.org/ss/rights_and_permissions.aspx

Appendix C (p. 1 of 2)

Finalized List of Story Grammar and Informative-Words for Storytelling of "The Boy Who Cried Wolf" in Cantonese

Story grammar	Informative-words					
	Scenario	Person	Action	Object	Place	Others
Setting	1. Introducing the main character	牧童/牧羊嘅小朋友 Shepherd boy/The child who herded sheep	放/看/睇 Keeping/Herding/ Looking after	羊 Sheep	山上 On the mountain	
Episode 1—Initiating event	2. Inferring why the boy cried wolf					
	Version 1	牧童 Shepherd boy	覺得 Felt			好悶/無聊 Bored
	Version 2	牧童 Shepherd boy 村民 Villagers	玩 Tricked			
Episode 1—Attempt	3. The boy lied					
	Version 1	牧童 Shepherd boy	話/嗌 Cried/Shouted			"狼來了" "Wolf"!
	Version 2	牧童 Shepherd boy 狼 Wolf	話/嗌 Cried/Shouted 黎 Came to		山上 On the mountain	
Episode 1—Reaction	4. Reaction of villagers					
	Version 1	村民/農夫 Villagers/Farmers 狼 Wolf	幫忙 Help 趕/打 Drove away/Beat			
	Version 2	村民/農夫 Villagers/Farmers 牧童 Shepherd boy	跑/衝/黎 Ran to/Rushed to/ Came to 救 Saved		山 Mountain	
Episode 2—Initiating event	5. The truth	狼 Wolf 牧童 Shepherd boy 村民/農夫 Villagers/Farmers	無 Was not 講大話 Told lies 被欺騙/呃 Were deceived 被整蠱 Were tricked			點知/原來 But/It turned out……
Episode 2—Attempt	6. The reaction of the boy	牧童 Shepherd boy 村民 Villagers	笑 Laughed 呃到 Lied to			蠢 Stupid

Downloaded From: https://ajslp.pubs.asha.org/ by a ReadCube User on 11/26/2018
Terms of Use: https://pubs.asha.org/ss/rights_and_permissions.aspx

Appendix C (p. 2 of 2)

Finalized List of Story Grammar and Informative-Words for Storytelling of "The Boy Who Cried Wolf" in Cantonese

	Informative-words					
Story grammar	**Scenario**	**Person**	**Action**	**Object**	**Place**	**Others**
Episode 2—Reaction	7. Reaction of villagers	村民/農夫 Villagers/Farmers	落 Went down 走 Left		山/屋企 Mountain/Home	好激氣/好嬲/好唔開心 Furiously/Angrily/ Unhappily
Episode 3—Initiating event	8. Real wolf came	狼 Wolf	黎 Came			真係 Really
Episode 3—Attempt	9. Reaction of the boy on real wolf coming	牧童 Shepherd boy	嗌/叫 Cried/ Shouted			"狼來" "Wolf!" "救命呀" "Help!"
Episode 3—Reaction	10. Reaction of villagers on real wolf coming	村民/農夫/山下面嘅人 Villagers/Famers/ People living at the foot of mountain 牧童/嗰個小朋友 Shepherd boy/The child who herded sheep	唔相信/救/理/幫 Did not believe/save/ attend to/help			
Coda	11. The consequence of the event Version 1	牧童啲羊 Shepherd boy's sheep 狼 Wolf	被食晒/咗 Were eaten			
	Version 2	狼 Wolf 牧童啲羊 Shepherd boy's sheep	捉/食晒 Were caught/eaten			
	12. Reaction of the boy toward the incident	牧童/嗰個小朋友 Shepherd boy/The child who herded sheep	落 Came back 喊 Cried		山/屋企/村 Mountain/Home/ Village	好慘/好傷心/好後悔 Poorly/In sorrow/ With regret

Downloaded From: https://ajslp.pubs.asha.org/ by a ReadCube User on 11/26/2018
Terms of Use: https://pubs.asha.org/ss/rights_and_permissions.aspx