Available online at www.sciencedirect.com
ScienceDirect
Journal of Communication Disorders 42 (2009) 442–464
Journal of Communication Disorders

The use of main concept analysis to measure discourse production in Cantonese-speaking persons with aphasia: A preliminary report

Anthony Pak-Hin Kong *

Department of Communication Sciences and Disorders, University of Central Florida, Orlando, FL, USA

Received 24 September 2008; received in revised form 20 June 2009; accepted 22 June 2009

Abstract

Discourse produced by speakers with aphasia contains rich and valuable information for researchers to understand the manifestation of aphasia as well as for clinicians to plan specific treatment components for their clients. Various approaches to investigate aphasic discourse have been proposed in the English literature. However, this is not the case in Chinese. As a result, clinical evaluations of aphasic discourse have not been a common practice. This problem is further compounded by the lack of validated stimuli that are culturally appropriate for language elicitation. The purpose of this study was twofold: (a) to develop and validate four sequential pictorial stimuli for elicitation of language samples in Cantonese speakers with aphasia, and (b) to investigate the use of a main concept measurement, a clinically oriented quantitative system, to analyze the elicited language samples. Twenty speakers with aphasia and ten normal speakers were invited to participate in this study. The aphasic group produced significantly less key information than the normal group. More importantly, a strong relationship was also found between aphasia severity and production of main concepts. While the results of the inter-rater and intra-rater reliability suggested the scoring system to be reliable, the test–retest results yielded strong and significant correlations across two testing sessions one to three weeks apart.

Learning outcomes: Readers will demonstrate better understanding of (1) the development and validation of newly devised sequential pictorial stimuli to elicit oral language production, and (2) the use of a main concept measurement to quantify aphasic connected speech in Cantonese Chinese.

Published by Elsevier Inc.

1. Introduction

A discourse can be defined as ‘‘a language unit whose organization supersedes any single word or sentence’’ (Olness, 2006, p. 176) and is composed of a series of connected sentences for conveying a message (Cherney, 1998). Various approaches have been proposed to quantify discourse production among speakers with aphasia. For example, the Quantitative Production Analysis (QPA) devised by Saffran, Berndt, and Schwartz (1989) focuses on quantifying the lexical contents and sentence structures of narratives. This analytic system was able to capture in detail the essential characteristics of grammatical, morphological, and structural disruptions of aphasic output. A discourse analysis that quantifies sentential and discourse grammars within a narrative was proposed by Ulatowska, Freedman-Stern, Doyel, Macaluso-Haynes, and North (1983a) and Ulatowska, North, and Macaluso-Haynes (1981). While the

* Tel.: +1 407 823 4791; fax: +1 407 823 4816.
E-mail addresses: pkong@mail.ucf.edu, antkong@graduate.hku.hk.

0021-9924/$ – see front matter. Published by Elsevier Inc.
doi:10.1016/j.jcomdis.2009.06.002

sentential grammar is analyzed in terms of (1) the number of words and morphemes, and (2) complexity and types of clauses, etc., the discourse grammar is measured by (1) presence of narrative abstract, settings, or morals, (2) length of modifying clauses, (3) number of story episodes, and (4) number of linkages between episodes, etc. It was found that the narratives elicited from speakers with aphasia tended to be shorter, less elaborated, and less informative. Another system that measures conversational abilities of speakers and the interactions between conversational partners, referred to as the Conversational Discourse Analysis (CDA), was developed by Boles (1997). In this system, the quantity of verbal output (e.g., the total number of words per utterance and utterances per conversation) and communicative strategies demonstrated by speakers (e.g., self-repairs and requests for clarifications) are considered. The above-mentioned systems contain detailed procedures for eliciting language samples and for undertaking data analysis that capture the essential clinical characteristics of speakers with aphasia. However, as these systems involve complex and time-consuming data transcription and analyses, they are unlikely to be widely used in clinical situations.

Some methods that are more clinically friendly were proposed in the literature. One of them is the Linguistic Communication Measure (LCM; Menn, Ramsberger, & Helm-Estabrooks, 1994) which was designed with the clear purpose of supplementing, rather than replacing, standard tests. Using the Cookie Theft picture of the Boston Diagnostic Aphasia Examination (Goodglass & Kaplan, 1993) as the speech elicitation stimulus, its aim was to reflect progress or deterioration in oral narratives over time. It contained a small number of indices to quantify (1) the amount of information conveyed (e.g., number of content units), (2) extent of word finding difficulty and emptiness of speech (e.g., Index of Lexical Efficiency), and (3) occurrence of morphological errors and omissions (e.g., Index of Grammatical Support). The computation of these indices required relatively minimal clinical experience and linguistic training. Although these performance indices were not intended to differentiate between aphasia types, it was argued that they reflected clinically important differences among aphasic individuals.

Another clinical approach employs rule-based procedures to score key episodes or scenarios expressed in connected speech. According to Van Dijk (1980), key concepts are the building components of discourse macrostructure. Organized sequences of key concepts thus form the skeletal outline of a discourse. Nicholas and Brookshire (1995) first reported the use of this scoring method to compare the amount of main information conveyed about a topic among 20 English speaking speakers with aphasia and 20 non-brain-damaged adults. Connected speech was elicited from each participant using four single pictures, two sets of sequential pictures, two conversations related to participants' personal information, and two procedural descriptions. A main concept was defined as ''statements (that) provide an outline of the gist or essential information portrayed in the stimulus pictures or an outline of the essential steps in the procedures... and should contain one and only one main verb'' (p. 148). The participants' language samples were evaluated on whether all potential main concepts were present, and if so whether the essential information in each main concept mentioned was accurate and complete. Each main concept identified was then classified as one of the following four categories: (1) Accurate and Complete (AC), (2) Accurate but Incomplete (AI), (3) Inaccurate (IN), and (4) Absent (AB). It was found that the non-brain-damaged group produced significantly more AC main concepts and significantly fewer AI, IN, and AB main concepts than the aphasia group. Moreover, out of the four categories, the AI and IN were found to be more discriminating between the two speaker groups than the AC and AB. The authors also suggested that the number of AB and AC was strongly related to estimated severity of aphasia. The point-to-point inter-rater and intra-rater reliability was 86% and 90% on average for the four scores, with the categories AI and AB causing the majority of disagreements. A mean correlation coefficient of 0.77 for the four scores in test–retest reliability also indicated acceptably stable performance of description on two occasions seven to ten days apart. It was concluded that the measure of main concept production could reflect the overall communication success and could supplement existing measures of communicative informativeness and efficiency.

About a decade after the first report of using main concept measurement in speakers with aphasia, Capilouto, Wright, and Wagovich (2006) compared the performance of eight healthy adults and eight adults with aphasia on their ability to convey main events in two single pictures and two sets of sequential picture stimuli. In this study, a main event was defined as ''an event that was of sufficient importance to the story as a whole and independent from the other events in the story'' (p. 207). This definition differed from the main concepts in Nicholas and Brookshire (1995) in that a main event might contain more than one main verb. As claimed by the authors, this could allow the capture of a speaker's ability to convey relationships between characters and events. At the same time, how well a speaker extracted critical ideas and inter-relationships from irrelevant detail could be indirectly evaluated (Capilouto, Wright, & Wagovich, 2005). A binary scoring system was used to calculate the proportion of main events produced by each participant; only those main events that contained all necessary information were scored as present. The results

indicated that the aphasic group produced significantly fewer main events than the healthy counterparts, thus highlighting the usefulness of the main event measurement to characterize their impaired narrative discourse abilities. Test–retest reliability of the main events measured at seven to 21 days apart was calculated, with significant correlations of 0.88 and 0.73 for sequential picture and single picture stimuli, respectively.

In summary, the various analytic systems for aphasic narratives mentioned above have looked at different aspects of discourse production. For the comprehensive systems, while the QPA (Saffran et al., 1989) focused primarily on the grammatical disruptions, the CDA (Boles, 1997) looked at the pragmatics of speech. These systems are different from the approach of quantifying speech contents and efficiency in the LCM described by Menn et al. (1994) or the proposition-counting approach suggested by Nicholas and Brookshire (1995) and Capilouto et al. (2006). Moreover, the measurement of proposition, i.e., key concept, production has been found to be a reliable means of quantifying discourse abilities in speakers with aphasia. Its critical clinical advantages lie in the fact that it allows clinicians to quantify aphasic discourse abilities in an easy, objective and reliable manner. It was also found to be a useful supplement to existing analytic tools of connected speech in English.

1.1. Analytic systems of discourse production in Cantonese

Chinese is a language spoken by around one-sixth of the world's population. Cantonese Chinese, which has over 110 million speakers worldwide, is the second most widely spoken dialect after Mandarin Chinese (Fung, 2009). It is widely spoken in the southern provinces in China, Hong Kong, Macau, and Chinese communities in Southeast Asia, North and South America, Australia, and Europe (Bauer & Benedict, 1997). Before discussing the analytic tools for Cantonese aphasic narratives, an introduction to Chinese grammar with highlights of its major differences from English grammar is in order. Unlike English and other European languages, there is a paucity of inflectional morphology in Chinese. The great majority of Chinese syllables in isolation (morphemes) have their own meaning. By the means of affixation (the attachment of bound morphemes to word stems), reduplication (the doubling of syllables or words), and compounding (the combination of independent morphemes), Chinese words are formed, including nouns, verbs, and adjectives. Cantonese, like Mandarin Chinese, is a Subject-Verb-Object (SVO) language with sentential modifiers between the subject and verb position (Packard, 1993). The language permits extensive omission of subjects and objects. This contrasts with English where the absence of the subject in a sentence, i.e., an elliptical utterance, is grammatically incorrect. In addition, the typical SVO word order can be violated in spoken colloquial Chinese language and still be considered grammatical. For example, a unique phenomenon in Cantonese is the right dislocation of nouns or pronouns in a sentence (VOS) (e.g., 搵番粒鈕喇, 媽媽 *wan2*[1] *faan1 lap1 nau2 laa3*, *maa4 maa1* ''found the button, mother''). Topicalization also occurs when a word or phrase in the object position is placed in the sentence or clause beginning (OSV), emphasizing the object and making it the sentence topic (e.g., 啲錢, 佢用晒喇 *di1 cin2*, *keoi2 jung6 saai3 laa3* ''all the money, he spent''). For a more in-depth discussion of the Chinese grammatical system, see Chao (1968), Kong (2006), or Li and Thompson (1981).

Compared to the English literature, relatively little research has been carried out on the discourse deficits among Cantonese-speaking individuals with aphasia. Earlier work was based on the QPA (Saffran et al., 1989). In particular, by taking into considerations the characteristics of Cantonese grammar, Yiu (1995) and Law (2001) have put forth a modified QPA for Cantonese to highlight in detail the linguistic characteristics of aphasic narratives, such as reduced amount and content, lower degree of sentence complexity and elaboration, impaired use of grammatical morphemes and lexical items of different form classes. As such, it is not suitable for clinical evaluation of aphasic narratives due to time-consuming transcript analysis. It was only in recent years that a clinically oriented examination system of impaired narratives in Cantonese was developed. Kong (2006) and Kong and Law (2004) have introduced a Cantonese Linguistic Communication Measure (CLCM) for analyzing aphasic narratives of a picture description task. The CLCM, developed based on the LCM (Menn et al., 1994), was designed to be a clinically feasible tool to reflect discourse performance using indices that were meant to evaluate the (1) contents (e.g., lexical diversity, counts of informative words and errors), (2) grammatical support, (3) degree of elaboration, and (4) efficiency (e.g., rate of

[1] In this paper, phonetic transcriptions are given in *jyutping*, a system developed by the Linguistic Society of Hong Kong to transcribe Chinese characters into roman letters. The tone of a syllable is represented by the number in the transcription.

production of informative words) of verbal outputs. While the Index of Errors was most useful in distinguishing speakers with aphasia from controls, the Index of Grammatical Support and Index of Communication Efficiency were most discriminative between individuals with fluent and non-fluent aphasia (Kong & Law, 2009). Four new single pictorial stimuli were developed in the CLCM. This was motivated by the lack of culturally appropriate stimuli to elicit language samples in the Chinese population.

1.2. Aims of the present study

The first aim of this study was to develop and validate four sets of newly devised sequential pictorial stimuli for elicitation of language samples among Cantonese speakers with aphasia. The use of four sets of stimuli was motivated by the fact that decisions made about an individual's connected speech based on one short sample are highly unreliable (Brookshire & Nicholas, 1994a) and the analysis of connected language samples elicited from four or more stimuli can ensure adequate test–retest stability (Brookshire & Nicholas, 1994b). In particular, Brookshire and Nicholas (1994b) found that the consistency of connected speech production increased substantially as the number of elicitation stimuli was raised from one to four or five. According to Potechin, Nicholas, and Brookshire (1987), the advantage of using sequential pictures to elicit verbal description from speakers with aphasia was that language samples were generally longer and, therefore, suited better to sentence level analyses. Moreover, the samples tended to consist of descriptions on a related sequence of activities; this contrasts with the tendency to list contents without the consideration of underlying relationships among characters or items in single picture stimuli. While it is commonly agreed that language and culture specific variables should be considered when evaluating speakers with aphasia, Kong (2006) has specifically highlighted the limitations of using picture stimuli with western cultural characteristics to elicit language samples from the Chinese population. The employment of the Cookie Theft picture (Goodglass & Kaplan, 1993) as the picture stimulus in a pilot study of the CLCM (Kong & Law, 2003) was found to be problematic. Not only did the speakers (elderly aphasic subjects as well as some of the controls) wrongly identify the kitchen scene as a living room, they were also unable to recognize the theme of stealing cookies from the cookie jar and were less capable of capturing the key elements in the picture. This seriously affected the production of target information and therefore reduced the validity of the test. The contents of stimuli in this study should therefore be culturally appropriate to the Chinese-speaking population.

A second purpose of the study was to evaluate the use of the main concept measurement, a clinically oriented analytic approach proposed by Nicholas and Brookshire (1995), for aphasic connected speech in Cantonese elicited through the picture sets. More specifically, how well this analytic method can (1) differentiate between speakers with and without aphasia, and (2) reflect the extent of the performance level of main concept production associated with the severity of aphasia would be investigated. The use of main concept measurement was motivated by (1) its straightforward computation method, (2) its reasonable stability across time to quantify content in aphasic discourse, and (3) its ability to act as a supplementary measure for diagnosis as well as intervention planning and monitoring.

2. Method

2.1. Development of stimulus materials

Four sets of sequential pictures were developed in the current study. Each picture set contains four detached single black-and-white line drawings with dimensions of 15 cm by 21 cm. While two of them contain only one character, the other two sets contain three characters and, therefore, should potentially contain more numbers of main concepts (than the first two sets) with reference to the inter-relationships among characters. The lexical items contained in the pictures were controlled in a way that the same item does not appear in more than one picture set. In picture set 1 (Cooking in a kitchen), an old lady is cutting some carrots in a kitchen. After cutting her finger accidently, she covers her wound with a band-aid that she took out from a first-aid kit. Picture set 2 (Waking up late for work) depicts a man waking up late. He brushes his teeth and combs his hair at the same time before he rushes to change his clothes. Finally, he finds that he is wearing a pair of non-matched socks. Picture set 3 (Buying ice-cream) shows a mother and her daughter visiting a store to buy some ice-cream. The mother is paying while the salesman is scooping an ice-cream cone. After that, the daughter drops her ice-cream and the salesman is kind enough to give her a new cone. In picture set 4 (Helping an old man), an old man carrying a grocery bag is walking in an opposite direction of a father and his son. The oranges from

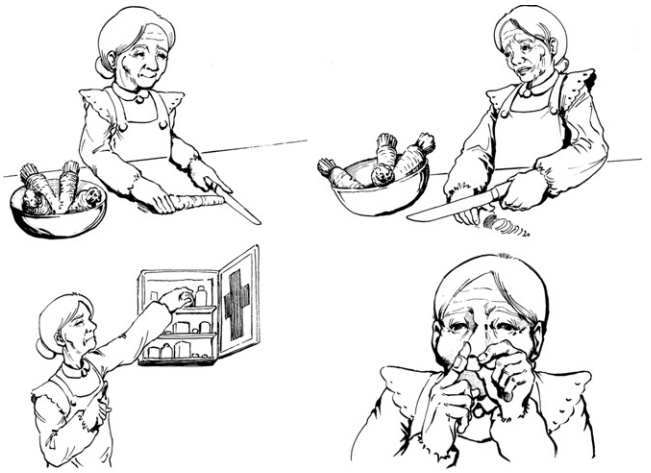

Fig. 1. An example of the pictures, set 1 (Cooking in a kitchen), used in this study.

the old man's bag fall on the ground. The boy notices the incident and helps the old man. At the end, the old man praises the boy. Fig. 1 shows an example of the pictures, set 1, used in this study. The other three sets of pictures can be obtained from the author.

To validate the main concepts depicted in the newly devised pictorial stimuli, procedures in the first stage of validation described by Nicholas and Brookshire (1995) were adopted. Eight speech-language pathologists in Hong Kong were invited to take part in the present study. The author explained to the SLPs the definition of a main concept, which is a statement that is important but independent from other concepts in the same story, i.e., a gist of the story. They were told that (1) each main concept can contain only one main verb or action, and (2) each main concept should be composed of two or more pieces of essential information, which is defined as the word(s) that is accurate in relation to the picture and informative about the content of the concept (Nicholas & Brookshire, 1995). For example, ''cut'' is the main verb for the main concept ''An old lady cuts carrots'' in picture set 1, and ''lady'', ''cut'' and ''carrots'' are the three pieces of essential information. Following this guideline, each SLP was asked to write a list of main concepts for each of the four picture sets. Main concepts that were listed by at least six out of the eight SLPs were included in the final main concept list. There are five and six main concepts for picture sets 1 and 2, respectively. As for set 3 and 4, there are nine and six main concepts, respectively. For each main concept, the essential information had to be listed by at least six out of the eight SLPs to be included in the final list. Appendix A shows the main concepts for picture set 1 to 4.

2.2. Subjects

Twenty participants (15 male and five female) with aphasia were recruited through two stroke support groups in Hong Kong, including The Self Help Group for the Brain Damaged and The Hong Kong Stroke Association. All of the participants had suffered a single unilateral cerebrovascular accident no less than six months post-onset at the time of the first test session. According to the results of the Cantonese version of the Western Aphasia Battery (CAB; Yiu, 1992), ten of them had a fluent aphasia (including eight Anomia, one Conduction, and one Wernicke's aphasia) and the other ten exhibited a non-fluent aphasia (including three Broca's, five Transcortical Motor, and two Isolation aphasia). A pre-test was carried out on the aphasic participants by asking them to describe four single picture stimuli in the Cantonese Linguistic Communication Measure (CLCM; Kong, 2006; Kong & Law, 2004). Based on their performance in the picture description tasks in the CAB and CLCM, none of them showed any form of visual deficits that might affect their scanning of the stimuli. In addition, all of them were judged to have adequate comprehension ability to complete the task in the present study. Subjects were divided into two fluency groups, fluent and non-fluent. Independent-sample t-tests were conducted to evaluate if the two fluency groups were significantly different in terms of aphasia severity (as estimated by the average Aphasia Quotient, AQ), age, and education level. It was found that the non-fluent speakers had significantly more severe aphasia than the fluent

Table 1
Background information on participants with Aphasia.

Subject	Gender	Age	Years of education	Aphasia type	Aphasia quotient	Occupation	Time post-onset (month)
Fluent							
A-FLC	F	49	2	Anomic	90.2	Tailor	48
A-CML	M	42	11	Anomic	89.8	Auto racer	15
A-TSK	M	49	9	Anomic	95.4	Driver	42
A-CYH	M	63	9	Anomic	93.8	Factory worker	47
A-HCH	M	65	10	Anomic	87.3	Cableman	17
A-HSF	M	50	8	Anomic	94.5	Cargo driver	58
A-TSO	F	52	13	Anomic	97.8	Teacher	71
A-STC	M	57	6	Anomic	80.7	Top of Form Decoration worker	108
C-TKY	M	64	4	Conduction	71	Top of Form Electrician	24
W-WSC	F	60	3	Wernicke's	62.6	Top of Form Tailor	61
Non-fluent							
TM-TKF	M	52	6	Transcortical motor	65.2	Cargo driver	37
TM-YKH	M	52	6	Transcortical motor	65.7	Chef	46
TM-TKS	M	76	3	Transcortical motor	64.7	Top of Form Decoration worker	93
TM-NCW	M	60	3	Transcortical motor	63.6	Top of Form Decoration worker	87
TM-MCT	F	46	9	Transcortical motor	66.7	Top of Form Tailor	42
B-LSK	M	46	9	Broca's	39.4	Jewelry maker	99
B-CSC	F	39	11	Broca's	63.1	Clerk	37
B-CCY	M	52	9	Broca's	60.7	Driver	56
I-YCM	M	48	13	Isolation	36.6	Top of Form Property manager	36
I-CCH	M	53	13	Isolation	30.5	Chef	9

Note: F = female, M = male; etiology of all participants is cerebrovascular accident.

group, $t(18) = 5.321$, $p = 0.000$. However, the two groups were not statistically different in terms of age and education level. In particular, the fluent speakers ranged in age from 42 to 65 years (mean = 55.1 years; standard deviation = 7.81 years) and had an average group AQ of 86.3 (standard deviation = 11.52). As for the non-fluent group, the age ranged between 39 and 76 years (mean = 52.4 years; standard deviation = 9.98 years) and the average group AQ was 55.6 (standard deviation = 4.47). The background information on the participants with aphasia is provided in Table 1.

Ten non-brain-damaged adults (five male and five female) were invited to take part in this study and act as controls. They ranged in age from 44 to 66 years (mean = 54.0 years; standard deviation = 6.77 years). All control speakers reported a negative history of neurological disease(s), head injury(ies), or other medical conditions that may affect their language performance. They were administered a short version of the CAB and a picture description task in the CLCM to rule out any expressive language deficits. All participants had signed a consent form for participation. The speakers with aphasia also signed a Health Insurance Portability and Accountability Act (HIPAA) authorization form for disclosure of their protected health information.

3. Procedures

3.1. Data collection

All testing procedures and data collection were conducted in a quiet, distraction-free room. The participants were presented with the four sets of sequential pictures in a random order. For each picture set, the test administrator would present individual pictures in the correct sequence, i.e., the participants were not required to arrange the sequence before they were prompted to describe them. The test administrator would then point to each main concept depicted in each of the four pictures. The participants were subsequently instructed to tell everything they saw happening in the pictures. Only general verbal cues or probing questions, including 哩度呢? *lei1 dou6 ne1* "What about here?", 仲有冇呀? *zung6 jau3 mou3 aa5* "Any more?", or 哩度發生咩事呀? *lei1 dou6 faat8 sang1 me1 si6 aa5* "What's happening here?" were given. If a participant showed a tendency to name individual items only, he/she was

instructed to use longer utterances or narratives and to perform the description task one more time. Language samples were recorded by a digital voice recorder, once the instruction had been given, until the participant indicated that he/she had finished or when 30 seconds of silence had passed. Ten speakers in the aphasic groups, who could return to the testing site no more than three weeks after the first testing, were asked to describe the picture stimuli again to obtain test–retest reliability. Each participant's language samples were then transcribed orthographically for further analysis.

3.2. Data analysis

The orthographically transcribed samples were scored on the basis of whether the main concepts were present and if the information in each main concept mentioned was accurate and complete. Six measures were employed to quantify each participant's performance. While the first four measures were directly adopted from Nicholas and Brookshire (1995), the other two were newly devised:

1. Number of Accurate and Complete concepts (AC): In order for a concept to be considered as accurate and complete, the speaker must provide all essential information in the main concept. The lexical items used to refer to the essential information must also be correct.
2. Number of Accurate but Incomplete concepts (AI): This is counted when a speaker misses one or more pieces of essential information in the main concept. The given essential information must all be accurate.
3. Number of Inaccurate concepts (IN): This is counted when one or more pieces of essential information in the main concept given are inaccurate. The fact that a speaker provides incomplete essential information would not affect the scoring as this measure focuses on only the accuracy of the description.
4. Number of Absent concepts (AB): None of the essential information in a main concept is given. That is, a speaker does not mention a particular main concept in the oral description.
5. Main concept score (MC): This is a score summarizing the first four measures and is computed by the formula $(3 \times AC + 2 \times AI + 1 \times IN)$. The calculation method proposed is based on the considerations of the following three areas of discourse production: the presence of essential information (independent to the degree of correctness) in a description, the accuracy in providing essential information, and the completeness of essential information given. For each of the 26 main concepts, one point is credited to the final score if each of above-mentioned discourse skills is evident in the language sample.
6. Number of AC per minute: This is an efficiency measure and is adapted from the Index of Communication Efficiency in the CLCM (Kong, 2006; Kong & Law, 2004), an index that was found to be useful in discriminating speakers with fluent and non-fluent aphasia. The length of a sample is measured and converted to minutes (e.g., 150 s = 2.50 min). Timing begins after the instruction has been given and ends with the termination of the last word of the speaker. The measure is then calculated by dividing the AC by the duration of the recording in minutes.

Appendix B provides a detailed description of the scoring criteria of the main concept measures. A sample list of main concepts for picture set 1 (Cooking in a kitchen) and scoring examples for these main concepts is given in Appendix C.

3.3. Statistical analysis

In order to examine the comparability of the four picture sets among aphasic and normal speakers, one-way ANOVAs were carried out for the performance of the four main concept measures (AC, AI, IN, and AB). To determine whether speakers with aphasia were different from controls, the mean, standard deviation, and range of the six measures were obtained for three speaker groups, namely the fluent aphasic, non-fluent aphasic, and control groups. The performances between the aphasic and control participants were then compared using one-way ANOVAs. The Bonferroni method was employed (Silva & Stam, 1995) to control for the occurrence of Type I errors. In particular, the p-value of each ANOVA was adjusted to 0.0167 (dividing 0.05 by 3).

Moreover, to evaluate the strength of the relationship between the estimated aphasia severity of the aphasia groups and their main concept performance, Pearson's correlation coefficients were calculated between the participants'

performance subtest scores of the CAB (Yiu, 1992), including the (1) spontaneous speech subtest score, (2) fluency rating score, (3) naming score, and (4) overall AQ, and the main concept performances. In particular, while positive correlations are expected between the CAB scores and AC, MC score, and AC per minute, negative correlations are expected for the remaining main concept variables. This will be related to the establishment of concurrent validity for the existing main concept measures.

Three reliability measures, including the inter-rater, intra-rater, and test–retest reliabilities, were calculated. For the number of AC per minute, which is a continuous measurement, the Pearson's correlation was computed. As for the remaining five main concept measures, which are ordinal in nature, the correlations were obtained using the Kendall's Tau correlation coefficient. Ten transcripts (50% of the data), including five samples from the fluent aphasic group and five from the non-fluent aphasic group, were randomly selected and analyzed by two speech-language pathologists, who were previously involved in the validation of the main concepts. The agreement between the author's analyses and those from each of the two raters were calculated to obtain the inter-rater reliability. Both raters went through a 45-min training session provided by the author, which included the introduction of the main concept measures and their calculation method, and a demonstration of the calculation of two scripts. The ten transcripts were also re-analyzed by the author to assess the intra-rater reliability. In addition, a point-to-point inter-rater and intra-rater reliability (Hula, McNeil, Doyle, Rubinsky, & Fossett, 2003; Nicholas & Brookshire, 1995) was carried out using the formula [Total agreements/(Total agreements + Total disagreements) × 100]. For the test–retest reliability, ten aphasic speakers who were able to return for retesting no more than three weeks after the first testing were asked to describe the picture stimuli again.

4. Results

The performance of the two groups of speakers with aphasia and the controls on each picture set is given in Table 2. Concerning how comparable the picture sets are in terms of eliciting main concepts from speakers with and without aphasia, the results of the one-way ANOVAs showed that for normal speakers, the number of Accurate and Complete concepts (AC) elicited was significantly different across the stimuli [$F(3.36) = 3.531$, $p = 0.024$]. Tukey HSDs revealed that picture set 3 (with nine potential main concepts) could significantly elicit more AC than set 1 (five concepts). As for the speakers with aphasia as a group, significant difference was found in the number of Absent concepts (AB) [$F(3.36) = 8.327$, $p = 0.000$]. Tukey HSDs further revealed that speakers with aphasia missed significantly more main concepts in the description of picture set 3 than the remaining stimuli. However, no significant differences were found in the production of AC concepts between sets 1 and 3 (as in the normal group).

The performance of the two aphasic groups and controls is given in Table 3. In general, the control group performed better than the aphasic participants on all measures, with the performance of the fluent aphasic group being better than that of the non-fluent aphasic group. The results of the one-way ANOVA shown in Table 4 indicate significant differences across the three groups on all six measures. Post hoc analyses further indicated that the controls are significantly better than the fluent as well as non-fluent aphasic participants on all measures. Moreover, significant differences on four out of the six measures, except for the AI and IN, were found between the two aphasic groups. A detailed review of the raw data revealed that while the non-fluent aphasic speakers tended to omit one or more pieces of essential information in the AI concepts, those in the fluent aphasic group tended to produce non-specific words or category names in place of essential information.

Table 5 displays the results of the correlations between the performance scores of the CAB (Yiu, 1992) and the main concept performance among participants with aphasia. All of the predicted relationships followed the correct direction, except for the correlation between the CAB naming subtest and the AI main concept scores. The Pearson's r coefficients yielded were statistically significant for AC, AB, MC score, and AC per minute. The coefficients between the CAB performances and AI and IN yielded were relatively low, suggesting that estimated aphasia severity had no important relation to the number of incomplete and inaccurate main concepts produced.

Table 6 displays the results of Kendall tau and Pearson's r coefficients for the main concept measures in the three types of reliability measures. All coefficients are significant at the 0.05 level or lower. The coefficients of the intra-rater reliability were comparable to those of the inter-rater reliability. The point-to-point intra-rater reliability was 80%. Considering the inter-rater reliability, the coefficients for the AI and IN were comparatively lower for both raters. The

Table 2
Raw scores of speakers with aphasia and controls.

	Set 1			Set 2			Set 3			Set 4		
	F	NF	Con	F	NF	Con	F	NF	Con	F	NF	Con
AC	2.60 (1.51), 0–4	0.30 (0.95), 0–3	4.40 (0.97), 2–5	2.10 (1.79), 0–5	0.00 (0.00), 0–0	5.60 (0.70), 4–6	1.80 (1.14), 0–4	0.00 (0.00), 0–0	6.50 (2.55), 4–9	2.50 (1.58), 0–5	0.20 (0.42), 0–1	5.40 (0.70), 4–6
AI	0.60 (0.70), 0–2	1.40 (0.97), 0–3	0.60 (0.97), 0–3	1.30 (1.25), 0–4	1.50 (1.43), 0–4	0.30 (0.48), 0–1	1.70 (0.82), 0–3	1.60 (0.97), 0–3	0.90 (1.20), 0–3	1.20 (1.03), 0–3	1.40 (0.84), 0–2	0.50 (0.53), 0–1
IN	0.60 (0.70), 0–2	0.50 (0.53), 0–1	0.00 (0.00), 0–0	0.90 (0.74), 0–2	1.00 (0.82), 0–2	0.00 (0.00), 0–0	0.60 (0.97), 0–2	0.70 (0.82), 0–2	0.10 (0.32), 0–1	0.60 (0.52), 0–1	0.60 (0.52), 0–1	0.00 (0.00), 0–0
AB	1.20 (0.42), 1–2	2.80 (0.92), 1–4	0.00 (0.00), 0–0	1.70 (0.95), 1–4	3.50 (1.78), 1–6	0.10 (0.32), 0–1	4.90 (0.57), 4–6	6.70 (1.16), 5–9	1.50 (1.51), 0–5	1.70 (0.67), 1–3	3.80 (1.03), 3–6	0.10 (0.32), 0–1

Note: AC = Number of Accurate and Complete concepts, AI = Number of Accurate but Incomplete concepts, IN = Number of Inaccurate concepts, AB = Number of Absent concepts; F = Fluent aphasic group, NF = Non-fluent aphasic group, Con = Control group; the values are listed in the order ''mean, (standard deviation), range''.

Table 3
Descriptive summary of participant performance.

Participant group	AC	AI	IN	AB	MC score	AC per min
Fluent						
Mean	9.00	4.80	2.70	9.50	39.30	3.68
Standard deviation	4.67	1.81	1.95	1.78	9.72	2.40
Range	0–16	2–8	0–7	7–13	19–53	0.00–6.92
Non-fluent						
Mean	0.50	5.90	2.80	16.80	16.10	0.09
Standard deviation	0.97	2.69	1.69	3.65	6.97	0.15
Range	0–3	1–9	1–6	11–24	3–27	0.00–0.39
Control						
Mean	21.90	2.30	0.10	1.70	70.40	9.41
Standard deviation	4.33	2.26	0.32	2.06	8.60	3.33
Range	11–26	0–7	0–1	0–7	48–78	4.00–14.02

Note: AC = Number of Accurate and Complete concepts, AI = Number of Accurate but Incomplete concepts, IN = Number of Inaccurate concepts, AB = Number of Absent concepts, MC Score = Main concept score.

Table 4
Statistical comparisons (ANOVAs) between performances of aphasic and control groups.

	One-way ANOVA $F(2,27)$	Post hoc (Tukey HSD)		
		F vs. NF	*F* vs. Con	NF vs. Con
AC	83.95^{***}	***	***	***
AI	6.54^{**}		*	**
IN	10.44^{***}		**	***
AB	82.69^{***}	***	***	***
MC score	102.60^{***}	***	***	***
AC per minute	39.21^{***}	**	***	***

Note: AC = Number of Accurate and Complete concepts, AI = Number of Accurate but Incomplete concepts, IN = Number of Inaccurate concepts, AB = Number of Absent concepts, MC Score = Main concept score; F = Fluent aphasic group, NF = Non-fluent aphasic group, Con = Control group; $^{*}p \leq 0.05$; $^{**}p \leq 0.0167$; $^{***}p \leq 0.001$.

Table 5
Correlations between the pperformance scores of the CAB and main concept performance in the participants with aphasia.

CAB performance scores	Pearson's *r*					
	AC	AI	IN	AB	MC Score	AC per min
Spontaneous speech subtest score	0.91^{*}	−0.19	−0.21	-0.89^{*}	0.94^{*}	0.86^{*}
Fluency score	0.95^{*}	−0.39	−0.19	-0.84^{*}	0.92^{*}	0.90^{*}
Naming subtest score	0.76^{*}	0.03	−0.06	-0.88^{*}	0.86^{*}	0.72^{*}
Aphasia quotient	0.83^{*}	−0.07	−0.18	-0.87^{*}	0.90^{*}	0.79^{*}

Note: AC = Number of Accurate and Complete concepts, AI = Number of Accurate but Incomplete concepts, IN = Number of Inaccurate concepts, AB = Number of Absent concepts, MC Score = Main concept score; $^{*}p \leq 0.001$.

point-to-point inter-rater reliability was 70% for both raters. As for the test–retest reliability, the coefficients ranged from 0.52 to 0.97, with the coefficients for the AI and IN being lower.

5. Discussion

The present study has developed and validated four new sets of sequential pictorial stimuli geared to elicit language samples from adults with aphasia in the Cantonese-speaking community. This is the first report of its kind that is

Table 6
Reliability measures of aphasic data.

Measure	Inter-rater reliability		Intra-rater reliability	Test–retest reliability
	Rater 1 vs. author	Rater 2 vs. author		
Kendall tau coefficient				
AC	0.83***	0.84***	0.89***	0.93***
AI	0.57*	0.48*	0.46*	0.52*
IN	0.66*	0.56*	0.66**	0.78*
AB	0.86***	0.77***	0.78***	0.75***
MC score	0.87***	0.81***	0.82***	0.85***
Pearson's *r*				
AC per minute	0.97***	0.97***	0.89***	0.97***

Note: AC = Number of Accurate and Complete concepts, AI = Number of Accurate but Incomplete concepts, IN = Number of Inaccurate concepts, AB = Number of Absent concepts, MC Score = Main concept score; $^{*}p \leq 0.05$; $^{**}p \leq 0.01$; $^{***}p \leq 0.001$.

related to Chinese. More importantly, given that most pictorial stimuli reported or published in the literature of Chinese aphasiology are single pictures, the products of this study have significant impact on the profession of speech and language pathology by broadening the range of testing materials. It is worth mentioning that the contents (e.g., characters and items) in these pictures are culturally appropriate, but not culturally specific, for the Chinese speaking population. This can ensure the potential use of these stimuli for other populations and subsequently allow the possibilities of future cross-linguistic comparative studies.[2]

A main concept measurement (Nicholas & Brookshire, 1995) was adopted for analyzing discourse production of Cantonese speakers with aphasia. Important modifications of the computation procedures with reference to the distinctive characteristics of Cantonese grammar were made. In particular, Chinese permits omissions of subjects and objects in an utterance (Lu, Tzeng, Bates, & Wulfeck, 1998). Use of elliptical utterances in a discourse is considered grammatically acceptable in Chinese, given that the subject component is previously given (Yiu & Worrall, 1996). This criterion was included in judging the completeness of a concept in the existing scoring system, i.e., an elliptical utterance or an utterance with missing objects was considered grammatical, accurate, and complete as long as the corresponding pronoun (or subject) or object was given in previous main concepts. This was usually the case when a speaker was describing two main concepts involving the same character or object in the picture set. This differed from the original scoring method in English in which the subject of an elliptical utterance or the object in an ill-formed utterance would be counted as a necessary piece of essential information of the main concept; a lack of the subject and/or object would therefore be counted as Accurate but Incomplete (AI) even if the remaining essential information was produced accurately. In other words, one major advantage of this system of main concept measurement lies in the fact that it is linguistically suitable for Cantonese Chinese.

The inclusion of the two new measures, MC score and AC/minute, has broadened the scope of rating discourse abilities without the need of time-consuming measure computation. While the former measure provides an overall summary of the degree of accuracy, completeness, and relevancy of discourse output; the measure of AC per minute can indicate how efficiently the information is conveyed. Both measures were found to have significant correlation to the subtest scores and the AQ of the CAB (Yiu, 1992). They are, therefore, believed to be useful indicators of aphasia severity and/or the overall performance of a speaker with aphasia. The inclusion of these measures in the main concept measurement may also have important clinical implications in supplementing the guiding of treatment focus and in gauging changes of connected speech during intervention.

The fact that picture set 3, with nine potential main concepts as opposed to five or six in the remaining sets, could elicit more Accurate and Complete concepts (AC) from controls is not unexpected. However, this is not the case for the aphasic group. More specifically, instead of producing more AC concepts, it was found that the aphasic group tended

[2] A pilot study is now being carried out in Cork, Ireland, to examine the use of these picture stimuli in English speakers with aphasia (C. Pettigrew, personal communication, 3 October 2008).

to omit most of the target information in set 3, and hence had a higher number of Absent concepts (AB); the number of AC concepts was not different from the other three pictures sets. This raises the issue of how well speakers with aphasia can handle procedural narrative tasks of varying difficulty (in terms of the quantity of content required to convey). Although Ulatowska, Freedman-Stern, Doyel, Macaluso-Haynes, & North (1983b) has described a clinical feature in aphasia for which speakers tended to produce only a certain proportion of essential contents in a procedural narrative task, whether there would be different patterns of manifestation (such as the patterns among speakers with varying types or severity of aphasia) have not been systematically investigated. This may be a further extension of the current study.

Consistent with the literature in English, the findings in this study suggested speakers with aphasia produced a significantly lower proportion of main concepts than normal speakers. A strong relationship between aphasia fluency and productions of main concepts was also found. However, one should also note that the fluency of the subjects in the present study was confounded with severity. Therefore, the group differences found in performance of the sequential picture description could be due to their type of aphasia or their level of aphasia. Future investigation should try to examine the effects of aphasia types and severity in isolation on discourse production, which may be reflected in the main concept analysis. For example, a larger numbers of speakers with aphasia can be recruited, so that there would be comparable numbers of subjects for each aphasia type and severity level. This would then allow the establishment of specific profiles of main concept variables for classifying various aphasia types. This will be of significant clinical implication by allowing speech therapists to provide a clinical diagnosis of the specific type of aphasia as in the CAB.

The results also suggest that the main concept measures in this study can be scored with acceptable inter-rater, intra-rater, and test–retest reliability for speakers with aphasia. In fact, the mean test–retest in the present study of 0.80 is comparable to the values of 0.77 and 0.88 reported in Nicholas and Brookshire (1995) and Capilouto et al. (2006), respectively. The results of the present point-to-point reliability are also comparable to those reported in these two studies. However, the inter-rater reliability in both raters and the intra-rater reliability are relatively lower for incomplete (AI) and inaccurate (IN) main concepts. A closer look at the original data revealed that the discrepancy between the author and each of the raters occurred mainly when the concepts were produced as broken words and/or phrases. In particular, for the IN main concepts originally scored by the author, both inter-raters had a tendency to consider them as Absent (AB) concepts; as for the AI concepts, they were either classified as Accurate and Complete (AC) or AB by the raters. When the author was re-analyzing the intra-rater data a tendency of rating IN and AI concepts as AB concepts was also found. The discrepancy of rating could be due to the fact that both inter-raters and the author carried out the analyses based solely on the orthographically transcribed language samples. Without listening to the participants' original oral production, disagreements in segmenting sentences for each main concept occurred. Finally, the test–retest reliability for the AI concepts is also found to be lower, possibly due to the variable production of details and errors from the participants with aphasia over the two interviews. This was especially the case for speakers with non-fluent aphasia when struggling for the contents.

One limitation of the present analytic system is that it did not examine in detail the relations between main concepts that were expressed in the aphasic discourse. In addition, as each main concept within a story is considered to be of equal degree of importance, the system cannot assess how well a speaker with aphasia is able to express the most critical aspects of a discourse in connected speech. This is a consequence of the operational definition of main concept measurement, which primarily focused on the quantification of presence, accuracy, and completeness in conveying ideas. Clinicians who would like to address other aspects of aphasic discourse production should employ other analyses, such as the CLCM (Kong, 2006; Kong & Law, 2004) for details on linguistic complexity or a main event analysis (Capilouto et al., 2006) for evaluating speakers' ability to express interrelationships between discrete ideas or actions. In addition, Wright, Capilouto, Wagovich, Cranfill, and Davis (2005) have previously found that younger healthy adults conveyed significantly more main concepts than older adults. Although an evaluation of the effect of age, as well as other factors such as education, is beyond the scope of the present study, a future development of this study should collect normative data with reference to factors that may affect language performance. This will be particularly important for clinicians when making judgments about the amount of relevant content on given language production tasks (Kong, 2006).

Analysis of discourse production among speakers with aphasia is an important and a necessary clinical procedure. While objective measurement of aphasic discourse as a part of a comprehensive assessment has recently been

promoted in western countries, such as the United States (American Speech-Language-Hearing Association, 2004) and the United Kingdom (RCSLT, 2005), it is often not carried out in clinical settings, but instead is limited to use for research purposes due to the time-consuming process of data transcription and analysis. This is believed to be the same case in Hong Kong. Armstrong, Brady, Mackenzie, and Norrie (2007) have recently addressed the use of a transcription-less analytic procedure of aphasic discourse. They were able to establish a high degree of concurrent validity for the concept production measures, adopted from the main concept measures in Nicholas and Brookshire (1995), between the agreements of a transcription-based analyses and their transcription-less analyses. Furthermore, it was suggested that the transcription-less analyses were promising in capturing a range of discourse features among speakers with aphasia. Based on the experiences of the author[3] and the comments of the speech-language pathologists[4] who used the main concept measurement systems during the establishment of the inter-rater reliability, the use of a transcription-less approach can be readily applied to the existing analytic procedures with appropriate training and practice. Such an outcome warrants further investigations and collection of aphasic data to guide clinical judgment of discourse production in aphasia.

6. Conclusion

The ability to highlight key concepts through narratives has been considered as a reliable indicator of the discourse abilities among speakers with aphasia. This study has reported a validation of four sets of sequential pictures to elicit discourse production from speakers with aphasia in Cantonese. Moreover, the present investigation has demonstrated the applicability of a main concept measurement to analyze discourse production elicited from these picture stimuli among Cantonese speakers with aphasia. In summary, there are three major advantages of this measurement, including (1) that the system utilizes sequential pictures, which are culturally appropriate to Chinese speakers, instead of a single picture, (2) that the revised analytic system is linguistically appropriate to Cantonese Chinese, and (3) that it is a relatively simpler and quicker system than other existing quantification systems in Chinese.

Acknowledgments

The author is grateful to all the participants for taking part in this study as well as Mr. Siu-Lam Yuen, Chairman of the Self Help Group for the Brain Damaged, and Mr. Kwok-Fan Poon, President of The Hong Kong Stroke Association, for their help in participant recruitment. Special thanks to speech-language pathologists Ms. Maggie Chu, Ms. Elaine Eramela, Mr. Jason Kan, Mr. Dustin Lau, Dr. Alice Lee, Mr. Wilson Leung, and Ms. Kathy Woo for their help in this study. Thanks also go to Ms. Lorinda Kwan, Clinical Supervisor in the Division of Speech and Hearing Sciences at the University of Hong Kong, for arranging student clinicians (Rita Chan, Jessi Chan, Regine Cheng, Emily Hui, Lydia Lai, Winnie Lam, Winsome Lam, Claudy Lau, Flora Lee, Rubi Ng, Jacky Sin, Chryso Wong, Sally Wong, Vivian Wong, Camille Yu, and Julia Yuen) to assist in the running of this study.

[3] For the author, the time required to carry out a full-scale calculation of main concept variables (fewer than 20 min for the four picture sets) is shorter than computing the CLCM variables (about 20 min for each of the four language samples).

[4] Comments from the speech-language pathologists include: (1) that the system is clinically oriented, (2) that the efficiency and accuracy of rating main concept production improved with practice, and (3) that orthographic transcriptions can become optional for data analyses with practice in the future.

Appendix A. Main concepts for picture set 1–4

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

Set 1 (Cooking in a kitchen)	
1	The old lady is **cutting** carrots 婆婆/女人/嬸嬸 **切** 紅蘿蔔 *po4 po2/neoi3 jan4/sam2 sam2* ***cit3*** *hung4 lo4 baak6*
2	The old lady **cuts** her finger 婆婆 **切**到 手指 *po4 po2* ***cit3*** *dou2 sau2 zi2*
3	The old lady's finger is **bleeding** 婆婆隻 手指 **流血** *po4 po2 zek3 sau2 zi2* ***lau4 hyut3***
4	The old lady is **looking for** something in a first-aid box 婆婆 去 藥箱度 **攞/搵** 膠布/藥水/嘢 *po4 po2 heoi3 joek6 soeng1 dou6* ***lo2/wan2*** *gaau1 bou3/joek6 seoi2/je3*
5	The old lady is **sticking** a band-aid 婆婆 **貼/稔** 膠布 *po4 po2* ***tip3/lim4*** *gaau1 bou3*

Set 2 (Waking up late for work)	
1	The man **wakes up** 哥哥/男仔/男人 **起身** *go4 go1/naam4 zai2/naam4 jan4* ***hei2 san1***

2	The man is **shocked/late** 哥哥 **好驚/遲**咗 *go4 go1* ***hou2 geng1/ci4*** *zo2*
3	The man is **brushing** teeth 哥哥 **刷牙** *go4 go1* ***caat8 ngaa4***
4	The man is **combing** hair 哥哥 **梳頭** *go4 go1* ***so1 tau4***
5	The man is **putting on** his pants 哥哥 **著** **褲** *go4 go1* ***zoek8*** *fu3*
6	The man is **wearing** a pair of two socks that are different/wrong in color 哥哥 **著**咗 **鴛鴦/錯** **襪** *go4 go1* ***zoek8*** *zo2 jyun1 joeng1/co3 mat9*

Set 3 (Buying ice-cream)	
1	The mother and the girl **approach** an ice-cream store 媽媽 同 妹妹 **去** 雪糕店 *maa4 maa1 tung4 mui4 mui2* ***heoi3*** *syut3 gou1 dim3*

2	The girl **asks for/wants** an ice-cream 妹妹 **要** 雪糕 *mui4 mui2* ***jiu3*** *sik9 syut3 gou1*
3	The mother is **paying** 媽媽 **俾錢** *maa4 maa1* ***bei2 cin2***
4	The man is **scooping** an ice-cream cone 售貨員 **畢** 雪糕 *sau6 fo3 jyun4* ***bat1*** *syut3 gou1*
5	The girl **dropped** the ice-cream on the floor 妹妹 個雪糕球 **跌**咗 落 地下 *mui4 mui2 go3 syut3 gou1* ***dit8*** *zo2 lok9 dei6 haa2*
6	The girl is **crying/unhappy** 妹妹 **喊/好傷心** *mui4 mui2* ***haam3/hou2 soeng1 sam1***
7	The man **looks at** the girl from behind 售貨員 **見**到 經過 *sau6 fo3 jyun4* ***gin3*** *dou2 ging1 gwo3*

8	The man **gives** the girl another ice-cream cone 售貨員 **俾** 另一杯 雪糕 妹妹 *sau6 fo3 jyun4* ***bei2*** *ling6 jat1 bui1 syut3 gou1 mui4 mui2*
9	The girl is **smiling** 妹妹 **笑** *mui4 mui2* ***siu3***

Set 4 (Helping an old man)	
1	The father and the son are **walking** on the street 爸爸 同 弟弟 **行街** *baa5 baa1 tung4 dai4 dai2* ***haang4 gaai1***
2	The old man is **carrying** a grocery bag 老伯 **拎**住 一袋嘢 *lou3 baak9* ***lik1*** *zyu6 jat1 doi6 je3*
3	The oranges **fall** on the floor 啲橙 **跌**晒 落 地下 *di1 caang2* ***dit8*** *saai3 lok9 dei6 haa2*
4	The boy **sees** the incident 弟弟 **見**到 經過 *dai4 dai2* ***gin3*** *dou2 ging1 gwo3*
5	The boy is **helping** the old man 弟弟 **幫** 伯伯 *dai4 dai2* ***bong1*** *baak9 baak9*
6	The old man is **praising** the boy 伯伯 **稱讚** 弟弟 *baak9 baak9* ***cing1zaan3*** *dai4 dai2*

Appendix B. Rules for scoring main concepts

There are altogether four sets of sequential pictures. Clinicians will be required to decide (1) if a main concept is present in a speaker's description, and (2) if the information in each main concept mentioned is accurate and complete. For each concept, one of the following four scores will be given:

B.1. AC—Accurate and Complete

1. All essential information in a main concept is accurate and complete.
2. Wordings:
   (a) All the lexical items in the essential information must be accurate in order for the main concept to be considered as AC.
   (b) If a speaker uses the category name to refer to a specific item in a main concept, such as 菜 *coi3* "vegetable" for 紅蘿蔔 *hung4 lo4 baak6* "carrot" or 生果 *saang1 gwo2* "fruit" for 橙 *caang2* "orange", consider it incomplete.
   (c) Speakers are not restricted to the use of the specific lexical items listed in Appendix A. Alternate wordings are allowed as long as the general meaning is the same. For example, acceptable alternatives for 女人 *neoi3 jan4* "woman" may include 媽媽 *maa4 maa1* "mother", 太太 *taai3 taa*i2 "wife", 師奶 *si1 laai1* "housewife", or 嬸嬸 *sam2 sam2* "aunt" and so on.
   (d) If the essential information in a main concept is accurate, but inaccurate words are also included in the statement, consider the main concept accurate unless the inaccurate words alter essential information to make it inaccurate.
   (e) If a speaker self-corrects or restates the essential information in a main concept, only the final version spoken will be scored. For example, if a speaker said 弟弟. . . 唔係. . .伯伯稱讚弟弟 *dai4 dai2. . . m4 hai6. . . baak3 baak3 cing1 zaan3 dai4 dai2* "The boy . . . no . . . the man praised the boy" in picture set 4, consider it accurate. In contrast, if the description is 伯伯. . . 唔係. . .弟弟稱讚伯伯 *baak3 baak3 . . . m4 hai6. . . dai4 dai2 cing1 zaan3 baak3 baak3* "The man . . . no . . . the boy praised the man", consider it inaccurate.
3. Pronouns:
   (a) Correctly used personal pronouns are allowed as long as the antecedent/original referent for the pronoun is provided previously in the production. For example, if a speaker said 哥哥起身。佢發覺佢 遲到呀 *go4 go4 hei2 san1. . . keoi3 faat3 gok3 keoi3 ci4 dou3 aa3* "The brother woke up. He found out he was late." in picture set 2, both main concepts will be accurate.
   (b) If a speaker only uses personal pronouns to refer to the referent(s) in the pictures, consider the first main concept incomplete. For the remaining main concepts with correct/acceptable pronouns, consider them accurate.
   (c) If the pronoun for a referent is ambiguous, consider the main concept incomplete. For example, if a speaker said 佢俾錢買雪糕 *keoi3 bei2 cin2 maai 3 syut3 gou1* "He/she paid for the ice-cream." in picture set 3 without stating whether it was the mother, the girl or the salesman paying, consider it AI.
   (d) If the pronoun for a referent is inaccurate, consider the main concept inaccurate. For example, if a speaker said 佢哋切緊紅蘿蔔 *keoi3 dei6 cit3 gan 2 hung4 lo4 baak6* "They are cutting carrots" in picture set 1, consider it IN.
   (e) If a referent or pronoun is correctly given in the first of a series of main concepts with the remaining pronoun missing, consider them accurate. For example, if a speaker said 個男人好快嘅刷牙, 之後就 梳頭 *go3 naam4 jan4 hou2 faai3 gam2 caat3 ngaa4. . . zi1 hau6 zau6 so1 tau4* "The man brushed his teeth quickly, then combed his hair" in picture set 2, the second main concept is considered AC as the referent 男人 *naam4 jan4* "man" is provided previously and is correct for the description.
   (f) Elliptical statements are considered grammatical, accurate, and complete as long as the corresponding pronoun (or the subject) is given in previous main concept(s). This is usually the case when a speaker is describing two main concepts involving the same character in the picture set. This also happens when more than one main concept is given in one statement by a speaker.
4. Form: Essential information does not have to be spoken in standard grammatical form or standard word order, as long as deviations would not lead to miscomprehension of the essential meaning of the main concept. For example, misuse of a classifier in the main concept 售貨員畢件雪糕 sau6 fo3 jyun4 bat1 gin6 syut3 gou1 "the salesman

scoops ice-cream'' or an aspect marker in 弟弟拖咗爸爸行街 *dai4 dai2 to1 zo2 baa4 baa1 haang4 gaai1* ''the boy held his father's hand and walk'' will still be considered as accurate as long as they contain all the essential information.

5. Articulation: The production does not have to be correctly articulated to be considered accurate, as long as it is intelligible to a listener as the target words in the context of what the speaker is saying.

B.2. AI—Accurate but Incomplete

1. One or more pieces of essential information in the main concept are missing, but all given essential information must be accurate.
2. Wordings: If non-specific words are spoken in place of essential information, consider the main concept incomplete. For example, if a speaker said 阿婆切緊啲嘢 *aa3 po4 cit3 gan 2 di1 je3* ''The granny is cutting something'' in picture set 1, consider it incomplete.

B.3. IN—Inaccurate

1. One or more pieces of essential information in the main concept given are inaccurate. Examples include paraphasias of action words or nouns, neologisms, and jargons or irrelevant lexical items that do not match with the content of the picture stimuli.
2. If a main concept containing inaccurate essential information is also missing some other essential information, it is still considered to be IN.

B.4. AB—Absent

1. None of the essential information in a main concept is given, i.e., the speaker does not provide any essential information that appears to be an attempt to communicate a particular main concept.

Note that if a main concept is self corrected, score the final production. For example, if a speaker said 弟弟稱讚伯伯. . . 唔係. . .伯伯稱讚弟弟呀 *dai4 dai2 cing1 zaan3 baak3 baak3 . . . m4 hai6. . . baak3 baak3 cing1 zaan3 dai4 dai2 aa3* ''The boy praised the man . . . no . . . the man praised the boy'' in picture set 4, consider it accurate. If a main concept is accurate and is repeated or restated, score the best version. For example, if a speaker said 阿婆切緊啲嘢。阿婆切緊啲紅蘿蔔呀 *aa3 po4 cit3 gan 2 di1 je3 ∘ aa3 po4 cit3 gan 2 di1 hung4 lo4 baak6 aa3* ''The granny is cutting something. She is cutting some carrots'' in Set 1, the final main concept produced will be used for scoring.

Appendix C. Main concepts and scoring examples for picture set 1

Examples of statements that would receive AC, Al, or IN scores are given for each of the five main concepts. For each AC main concept, essential information (defined as the word(s) that is accurate in relation to the picture and informative about the content of the concept) is underlined. The main verb is **bolded**. Common alternate words for the essential information in each concept are given. Note that it is not required for the main concepts to be given in the transcript in the same order they are shown in the list of main concepts. If a main concept is spoken several times, only the final version is scored in the case of self correction.

1	**AC**	The old lady is **cutting** carrots 婆婆/女人/嬸嬸 **切** 紅蘿蔔 *po4 po2/neoi3 jan4/sam2 sam2* ***cit3*** *hung4 lo4 baak6*
	AI	The mother cuts something 媽媽切嘢 *maa4 maa1 cit3 je3*
	IN	There is an old lady using a knife to cut potatoes 有個婆婆攞住把刀切緊啲馬鈴薯 *jau5 go3 po4 po2 lo2 zyu6 baa2 dou1 cit3 gan2 di1 maa3 ling4 syu4*
		The mother cuts fruits 媽媽切生果 *maa4 maa1 cit3 saang1 gwo2*
2	**AC**	The old lady **cuts** her finger 婆婆 **切**到 手指 *po4 po2* ***cit3*** *dou2 sau2 zi2*
	AI	She is careless and hurt herself 佢自己唔小心整到 keoi3 zi6 gei2 m4 siu2 sam1 zing2 dou2

	IN	The old lady sawed herself 婆婆踞親 *po4 po2 goe3 can1*
3	**AC**	The old lady's finger is **bleeding** 婆婆隻 手指 **流血** *po4 po2 zek3 sau2 zi2* ***lau4 hyut3***
	AI	The old lady's finger is painful 女人隻手指痛 neoi2 jan4 *zek3 sau2 zi2 tung3*
	IN	The old lady's feet is bleeding 婆婆隻腳流血 *po4 po2 zek3 goek3 lau4 hyut3*
4	**AC**	The old lady is **looking for** something in a first-aid box 婆婆 去 藥箱度 **攞/搵** 膠布/藥水/嘢 *po4 po2 heoi3 joek6 soeng1 dou6* ***lo2/wan2*** *gaau1 bou3/joek6 seoi2/je3*
	AI	The old lady goes to the first-aid box 婆婆去藥箱度 *po4 po2 heoi3 joek6 soeng1 dou6*
	IN	The boy is getting a band-aid from a first-aid box 哥哥去藥箱度搵膠布 *go5 go1 heoi3 joek6 soeng1 dou6 wan2 gaau1 bou3*
5	**AC**	The old lady is **sticking** a band-aid 婆婆 **貼/稔** 膠布 去手指度 *po4 po2* ***tip3/lim4*** *gaau1 bou3 heoi3 sau2 zi2 dou6*
	AI	The band-aid is used 膠布呢就係黐翻好 *gaau1 bou3 ne1 zau6 hai6 ci1 faan1 hou2*
		Put something on the finger 搵啲藥黎敷住嗰隻手 *wan2 di1 joek6 lai4 fu1 zyu6 go2 zek3 sau2*
	IN	The old lady is tying a band-aid 嗰啲膠布綁住隻手指 *go2 di1 gaau1 bou3 bong2 zyu6 zek3 sau2 zi2*
		Apply a band-aid 搽膠布 *caa4 gaau1 bou3*

Appendix D. CEU questions

1. Which of the following statements about the main concept measurement is the most accurate?
   1. It is a rule-based scoring method to quantify the completeness and accuracy of the content in oral discourse.
   2. The computation of the measures for this system is straightforward.
   3. It will allow users to assess the degree of lexical diversity and extent of elaboration in aphasic discourse.

   (a) 1 only.
   (b) 2 only.
   (c) 3 only.
   (d) 1 and 2.
   (e) All of the above.

2. A tool with good _____ should give the same measurement over time.
   (a) Test-retest reliability.
   (b) Face validity.
   (c) Intra-rater reliability.
   (d) External validity.
   (e) Inter-rater reliability.

3. ''A language unit whose organization supersedes single words and is composed of connected sentences for conveying a message'' is a definition of:
   (a) Story.
   (b) Episode.
   (c) Discourse.
   (d) Conversation.
   (d) None of the above.

4. Which of the following statements about the grammatical systems in Chinese and English are accurate?
   1. Same as the English grammatical system, there is a rich inflectional morphology in Chinese.
   2. The typical ''agent-action-recipient'' word order can be violated in colloquial spoken Chinese.
   3. Chinese vocabularies can be formed by combination of individual morphemes.
   4. Chinese allows omission of subjects in a sentence. In contrast, elliptical sentences are considered ill-formed in English.

   (a) 1, 2, and 3.
   (b) 1, 2, and 4.
   (c) 1, 3, and 4.
   (d) 2, 3, and 4.
   (e) All of the above.

5. Which of the following statements about the rules for scoring main concepts is the most accurate?
   (a) For a concept to be considered as ''accurate and complete'', all target essential information must be present in the description.
   (b) If oral description is not in standard grammatical form, it will automatically be counted as ''inaccurate''.
   (c) If a main concept is self corrected, the final production will be used for scoring.
   (d) (a) and (c).
   (e) All of the above.

References

American Speech-Language-Hearing Association, (2004). Ad Hoc Committee on Admission/Discharge Criteria in Speech-Language Pathology. (2004, April). Guildelines: Admission/Discharge criteria in speech-language pathology. *ASHA, 24(supplement)*, 65–70.

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

Bauer, R. S., & Benedict, P. K. (1997). *Modern cantonese phonology.* Berlin/New York: Mouton de Gruyter.

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.

Brookshire, R. H., & Nicholas, L. E. (1994a). Test-retest stability of measures of connected speech in aphasia. *Clinical Aphasiology, 22*, 119–133.

Brookshire, R. H., & Nicholas, L. E. (1994b). Speech sample size and test-retest substainity of connected speech measures for adults with aphasia. *Journal of Speech and Hearing Research, 37*, 399–407.

Capilouto, G., Wright, H. H., & Wagovich, S. A. (2005). CIU and main event analyses of the structured discourse of older and younger adults. *Journal of Communication Disorders, 38*, 431–444.

Capilouto, G. J., Wright, H. H., & Wagovich, S. A. (2006). Reliability of main event measurement in the discourse of individuals with aphasia. *Aphasiology, 20*, 205–216.

Chao, Y. R. (1968). *A grammar of spoken Chinese*. University of California Press: Berkeley.

Cherney, L. R. (1998). Pragmatics and discourses: An introduction. In L. R. Cherney, B. B. Shadden, & C. A. Coelho (Eds.), *Analyzing discourse in communicatively impaired adults* (pp. 1–7). Gaithersburg, MD: Aspen Publishers Inc.

Fung, R. S. Y. (2009). Characteristics of Chinese in relation to language disorders. In S. P. Law, B. Weekes, & A. Wong (Eds.), *Language disorders in speakers of Chinese* (pp. 1–18). Multilingual Matters: Clevedon, UK.

Goodglass, H., & Kaplan, E. (1993). *The boston diagnostic aphasia examination*. Philadelphia: Lee & Feliger.

Hula, W. D., McNeil, M. R., Doyle, P. J., Rubinsky, H. J., & Fossett, T. R. D. (2003). The inter-rater reliability of the story retell procedure. *Aphasiology, 17*, 523–528.

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., & Law, S. P. (2003). A Cantonese linguistic communication measure. *Asia Pacific Journal of Speech, Language and Hearing, 8*(*3*), 229–234.

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). Cantonese Linguistic Communication Measure (CLCM): A clinical tool for assessing aphasic narrative production. In S. P. Law, B. Weekes, & A. Wong (Eds.), *Language disorders in speakers of Chinese Clevedon* (pp. 255–272). UK: Multilingual Matters.

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

Li, C., & Thompson, S. (1981). *A functional reference of grammar of mandarin Chinese*. Berkeley: University of California Press.

Lu, C. C., Tzeng, O., Bates, E., & Wulfeck, B. (1998). Sensitivity to grammaticality in Chinese aphasics. *Brain and Language, 65*, 62–64.

Menn, L., Ramsberger, G., & Helm-Estabrooks, N. (1994). A linguistic communication measures for aphasic narratives. *Aphasiology, 8*(*4*), 343–359.

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*, 145–156.

Olness, G. S. (2006). Genre, verb, abd coherence in picture-elicited discourse of adults with aphasia. *Aphasiology, 20*, 175–187.

Packard, J. L. (1993). *A linguistic investigation of aphasic Chinese speech*. Dordrecht: Kluwer Academic Publishers.

Potechin, G. G., Nicholas, L. E., & Brookshire, R. H. (1987). Effects of picture stimuli on discourse production. *Aphasiology, 17*, 216–220.

Royal College of Speech and Language Therapists. (2005). *RCSLT Clinical Guidelines*. Bicester, Oxon: Speechmark Publishing.

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.

Silva, A. P. D., & Stam, A. (1995). Discriminant analysis. In L. G. Grimm & P. R. Yarnold (Eds.), *Reading and understanding multivariate statistics* (pp. 277–318). Washington, DC: American Psychological Association.

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

Ulatowska, H. K., Freedman-Stern, R., Doyel, A. W., Macaluso-Haynes, S., & North, A. J. (1983b). Production of procedural discourse in aphasia. *Brain and Language, 18*, 315–341.

Ulatowska, H. K., North, A. J., & Macaluso-Haynes, S. (1981). Production of narrative and procedural discourse in aphasia. *Brain and Language, 13*, 345–371.

Van Dijk, T. A. (1980). *Macrostructures*. Hillsdale, NJ: Lawrence Erlbaum.

Wright, H. H., Capilouto, G. J., Wagovich, S. A., Cranfill, T. B., & Davis, J. E. (2005). Development and reliability of a quantitative measure of adults' narratives. *Aphasiology, 19*, 263–273.

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

Yiu, E. M. L. (1995). *Sentence production in aphasic subjects: A cross-language comparison*. Unpublished doctoral dissertation. Brisbane, Australia: The University of Queensland.

Yiu, E., & Worrall, L. (1996). Patterns of grammatical disruption in Cantonese aphasic subjects. *Asia Pacific Journal of Speech, Language and Hearing, 1*, 105–126.