Linguistic analysis of discourse in aphasia: A review of the literature

Lucy Bryant, Alison Ferguson & Elizabeth Spencer

Published online: 22 Mar 2016.

Submit your article to this journal ↗

Article views: 394

View related articles ↗

View Crossmark data ↗

Taylor & Francis
Taylor & Francis Group

Linguistic analysis of discourse in aphasia: A review of the literature

Lucy Bryant, Alison Ferguson, and Elizabeth Spencer

Speech Pathology Discipline, School of Humanities and Social Sciences, Faculty of Education and Arts, University of Newcastle, New South Wales, Australia

ABSTRACT

This review examined previous research applications of linguistic discourse analysis to assess the language of adults with aphasia. A comprehensive literature search of seven databases identified 165 studies that applied linguistic measures to samples of discourse collected from people with aphasia. Analysis of methodological applications revealed an increase in published research using linguistic discourse analysis over the past 40 years, particularly to measure the generalisation of therapy outcomes to language in use. Narrative language samples were most frequently subject to analysis though all language genres were observed across included studies. A total of 536 different linguistic measures were applied to examine language behaviours. Growth in the research use of linguistic discourse analysis and suggestions that this growth may be reflected in clinical practice requires further investigation. Future research directions are discussed to investigate clinical use of discourse analysis and examine the differences that exist between research and clinical practice.

Assessment of language impairment in aphasia is regularly conducted within clinical speech pathology practice using psycholinguistic assessment tools. The standardisation of such tools lends efficiency and reliability to the assessment process by providing a strict set of instructions for administration, interpretation of results and classification of impairment. However, such assessments may be limited in the results they provide by their failure to consider the interaction between the structural, linguistic components of language – phonology, morphology, syntax and semantics (Prins & Bastiaanse, 2004; Armstrong, Brady, Mackenzie, & Norrie, 2007). The focus on the nature of language deficits and how these may reflect regions of neurological damage captures the domain of impairment within the World Health Organization's International Classification of Functioning, Disability and Health (ICF) (World Health Organization, 2011). Further, by controlling the administration environment and limiting potential distractions, the cognitive demands that are present during typical language use and conversation are not adequately represented in psycholinguistic testing contexts. Such outcomes have been illustrated in research that demonstrated that persons with aphasia performed differently, and in many cases better, in communicative contexts than in controlled testing conditions (Mayer & Murray, 2003; Ulatowska et al., 2003; Kemper, McDowd, Pohl, Herman, &

Jackson, 2006; Beeke, Maxim & Wilkinson, 2008; Herbert et al., 2008). With some people with aphasia performing deceptively well on these standardised tasks, other means of assessment should be considered.

In acknowledgement of these limitations, assessment methods have been developed with a greater focus on the functional domains of the ICF–social participation and activity. Termed functional assessments for their focus on communication activities and situations experienced in the course of everyday life, these methods often require observation and subjective rating of the ability to perform communicative activities (Armstrong, Ferguson, & Simmons-Mackie, 2013). Another approach has been applied to discourse, or language in use. The definition and scope of what is meant by discourse varies across disciplines, but within speech pathology the term is typically applied to describe the way language in use is structured above the sentence level (Armstrong, 2000). The functional approach to defining discourse also recognises the role that structures such as words and phrases serve in meaningful conversational contributions (Ulatowska, Allard, & Chapman, 1990; Cherney, 1998). For the purposes of this review, both the structural and functional definitions have been considered in identifying discourse as connected language samples representing any of the four discourse genres: exposition, procedural, narrative or conversational. While such a definition takes a very general approach to discourse, it is necessary in the context of reviewing clinical and research approaches to such assessment in order to consider all instances of discourse analysis use. Both the structural and functional theoretical underpinnings of discourse utilise similar processes of language sample collection and analysis to achieve an understanding of the way individuals with aphasia use language to communicate. When considering the analysis of discourse, definitions of the term as it is used in speech pathology generally describe the assessment of language productions of only one individual, even in contexts where another may contribute. Therefore, discourse analysis within speech pathology, and within this review, assesses the language of only one speaker, excluding any contributions made by others. When approaching this analysis, the researcher or clinician may consider linguistic, non-verbal or pragmatic components of language. Where the language of more than one individual is included in the analysis, the term 'conversational analysis' is applied. Through conversational analysis, interactive and pragmatic features of language beyond the linguistic structures are also examined (Prins & Bastiaanse, 2004).[1] This study focuses on the analysis of linguistic structures within discourse contexts, referred to hereafter as 'linguistic discourse analysis'.

Linguistic discourse analysis provides a supplementary form of assessment to psycho-linguistic tools which allow the identification of isolated impairments within single linguistic domains by detecting additional difficulties and adaptive strengths that are apparent when these domains interact (Huber, 1990; Coelho, Grela, Corso, Gamble, & Feinn, 2005; Marini, Andreetta, del Tin, & Carlomagno, 2011). Linguistic discourse analysis may also provide a means of identifying clinically significant residual impairments which standard measures may overlook when examining linguistic domains in isolation.

---

[1]'Conversational analysis' is a term not uncommon in the speech pathology literature to describe a general and less formal approach to the analysis of interaction and all interlocutors' utterances. It differs from Conversation Analysis (CA), the specific procedure for transcription and analysis defined by researchers such as Sacks, Schegloff & Jefferson (1974).

This is particularly evident in studies that have applied the cognitive demands of dual task conditions to the collection of connected language samples. Language performance was adversely affected by the concurrent completion of other activities, as were other activities performed under conditions of divided attention (Murray, Holland, & Beeson, 1998; Kemper et al., 2006). For example, in a group of persons with aphasia deemed 'recovered' by standardised testing requirements, Kemper et al., (2006) found that having the person perform a motor task or ignore background noise while producing a language sample for analysis resulted in reduced lexical diversity, grammatical complexity and idea density in the resulting language samples. This additional sensitivity to language deficits that can be achieved using linguistic discourse analysis may reveal strengths and weaknesses that are important in guiding interventions with a focus on functional language outcomes (Coelho et al., 2005; Marini et al., 2011).

However, linguistic discourse analysis for the assessment of aphasia has been limited within clinical speech pathology practice for a number of reasons which have been discussed in the aphasia literature. Brady and colleagues (2012) suggested that the lack of a clinically acceptable tool has led to the use of surrogate assessments in the form of functional rating scales in clinical practice which act as a conduit to interpreting conversational language behaviours. Primarily, the time and clinical knowledge necessary to complete each step of the linguistic discourse analysis process contributes to the impracticality and inefficiency of the tool within the clinical environment at several stages of application (Armstrong, 2000; Togher, 2001). Clinicians must collect language samples, transcribe them and, depending on the analysis to be used, code linguistic behaviours. The analysis must be applied and the results of the discourse assessment interpreted. The time necessary to complete the process in full has been estimated to range from 6 to 12 minutes to for every minute of sampling collected (Elia, Liles, Duffy, Coelho, & Belanger, 1994; Boles, 1998). When addressing their rationale behind an investigation of transcription-less analysis, Armstrong and colleagues (2007) suggested that the time estimate may even increase up to an hour for every minute of language to complete transcription alone. While time is a leading barrier to clinical translation of linguistic discourse analysis research, it has been suggested that further barriers exist at each step of the process. The initial step of language sampling prior to linguistic discourse analysis requires collection of language that is representative of that used by the person with aphasia. Both the length and the type of sample that best represent language used for communication remains a debated issue. Collection of multiple samples covering different genres including narrative, procedure and conversation, has been suggested as a possible means of sampling that may be necessary to ensure discourse represents language in use (Armstrong, 2000; Olness, 2006). Following sample collection, appropriate measures and methods for analysis must be selected based upon the language presentation of the individual being assessed, and the parts of language that require assessment. Next, clinicians require training in order to identify target linguistic behaviours within the sample and perform an accurate and appropriate analysis. Finally, results of analysis must be interpreted with knowledge of linguistic processing, language structure and the patterns of impairment observed in aphasia required (Marini et al., 2011; Jaecks, Hielscher-Fastabend, & Stenneken, 2012). With such a specific skill set necessary to complete linguistic discourse analysis, clinicians may lack or perceive themselves to lack the expertise required to perform such an assessment.

These identified barriers may have a significant impact on the clinical use of linguistic discourse analysis. However, the literature has not been closely examined in order to observe and collate evidence that may provide insight into the use of linguistic discourse analysis. Ongoing applications of this method in research contexts despite the reported barriers may suggest ways to facilitate clinical applications. In order to describe research use, a review of the literature is necessary. The purpose of this review is to examine the methodological applications of linguistic discourse analysis in aphasiology in order to determine whether it is ongoing. A similar review that focused on the methodology of studies rather than on findings/outcomes has been applied to the aphasiology literature by Simmons-Mackie and Lynch (2013) in relation to the use of qualitative research methods, and the characteristics and patterns observed in its application. They asserted that knowledge of ongoing use of qualitative methods would guide future research practice and direction.

The aim of this review was to examine the research applications of linguistic discourse analysis. The specific research questions were:

- To what extent has the frequency of application of linguistic discourse analysis in research changed over time?
- What have been the main purposes to which linguistic discourse analysis has been applied?
- What methods have been used in the application of linguistic discourse analysis in relation to discourse elicitation, transcription and linguistic analysis?

Method

A search of the literature was conducted in October 2014 in order to identify studies utilising linguistic discourse analysis to examine language in adults with aphasia following left-hemisphere stroke. This was repeated in September 2015 in order to collect studies published in the period following the original literature search. Search terms were chosen to focus on adults who had received a diagnosis of aphasia (aphasia, aphasic, dysphasia or dysphasic) and studies using methods of linguistic discourse analysis, or linguistic analysis of spontaneously or semi-spontaneously elicited connected language (discourse, communication/communicative, narrative or story). The resulting search string was as follows: (aphas* OR dysphas*) AND (discourse OR communicat* OR narrative OR story) AND adult. The search terms were used for an electronic database search of MEDLINE, CINAHL Complete, Linguistics Language Behaviour Abstracts (LLBA), PsycINFO, Web of Science (Core collection) and Nursing and Allied Heath Source (via Proquest) and of the publisher database of Taylor and Francis. No restrictions were placed on the literature search in relation to dates of publication in order to collect all relevant studies. A total of 7248 studies were identified.

For inclusion, studies met the following criteria: (1) used descriptive analysis of language samples above the sentence level, or focused on the role that structures such as words and phrases served in meaningful conversational contributions; (2) participants were adults (over 18 years of age); (3) a diagnosis of aphasia had been made following left-hemisphere stroke; (4) the analysed language was spoken or written English and (5)

reported original research into language in aphasia, i.e. not a review or discussion of previously published research. Studies were excluded from the analysis if they did not meet the inclusion criteria, and if they were: (1) duplicate copies of included studies; (2) not published in a peer-reviewed journal and (3) not written in English. Linguistic discourse analysis studies were defined as those analysing structural linguistic elements of language including lexical, grammatical and semantic structures within samples, by providing a descriptive report or numeric representation of linguistic behaviours. These were identified by author statement of the type of linguistic discourse analysis used, or description of linguistic analysis of language samples elicited by spontaneous or controlled means. Conversational analysis approaches that assessed interactive and extralinguistic elements of language were not included in this review.

A diagram illustrating the assessment process for study inclusion is presented in Figure 1. First, studies with multiple copies retrieved during the search were removed, leaving only a single copy (629 results excluded). The titles and abstracts of each published study were then reviewed by the first author to assess whether they met the inclusion and exclusion criteria. Those items that did not represent research (book reviews, conference proceedings and editorials (i.e. the grey literature) (917 results) and publications reporting reviews of existing literature (557 results) were excluded as these sources did not provide detailed information regarding the linguistic discourse analysis methods used in research, as were the focus of analysis in this review. Those studies that explicitly stated a research population age or diagnosis that was not within the inclusion criteria (3851 results), a population language other than English (197 results), a focus on physical or psychological outcomes other than language (402 results) and use of assessments that did not use analysis of linguistic features in productions longer than a single sentence (667 results)
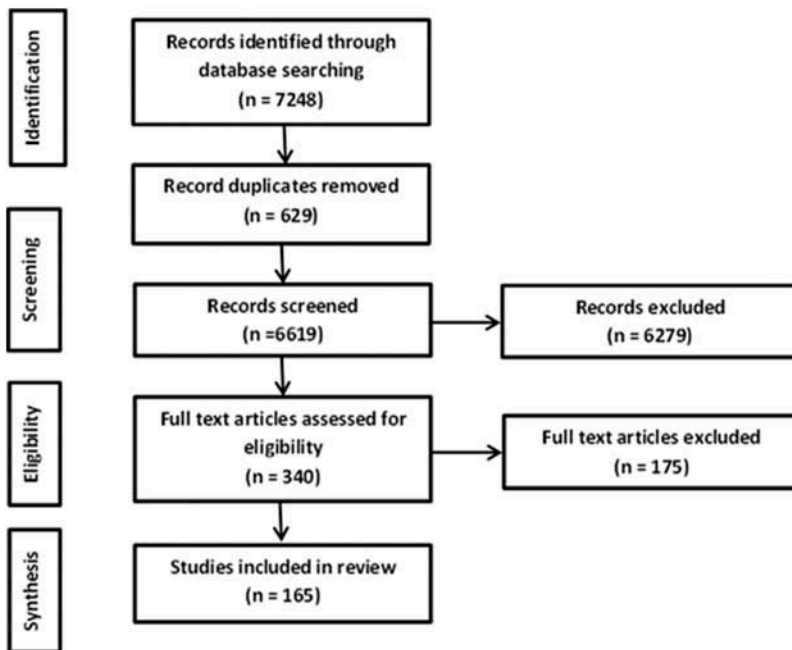

Figure 1. Flow diagram of the decision process of sorting retrieved literature.

were also excluded from the review. Following this exclusion process, 340 studies remained. Full-text copies of these studies were retrieved for more comprehensive review. Where it was unclear if the criteria for inclusion had been met after full-text review, a decision was made through discussion and consensus between all researchers. Of the full text-studies, a further 175 were excluded based on the previously mentioned criteria. One additional study was removed from the analysis as it failed to identify if participants had received a diagnosis of aphasia following their stroke.

The remaining 165 studies (see Appendix A) were assessed and relevant information was extracted for analysis. Target information analysed from each study included the date of publication retrieved from the reference; the purpose of linguistic discourse analysis application identified in the aims of the study and in the report of the method; and the methods of elicitation, types of language samples collected for analysis and the explicitly stated linguistic discourse analyses applied to collected samples as reported in the *Method* and *Results* sections of the studies. Only explicitly reported information was collected from the included studies. Any incomplete data were reported as such in the results of this review.

Once relevant data were extracted, it was analysed by grouping data into categories to code across studies (see Appendix B). For elicitation and analysis procedures, these categories were determined through the application of inductive reasoning to a list of results obtained from the included studies. The number of studies within each category was calculated. A random sample of 10% of included studies was selected for analysis by the second author to establish coding agreement. Disagreement on the coding of linguistic analysis measures was reviewed further through a collaborative training and discussion process between all three authors, and re-analysis of a further random sample of 15% of included studies, bringing the total to 25% of studies reviewed to assess agreement. Inter-rater agreement was as follows: year of publication, 100%; purpose of application, 86.7%; discourse elicitation method, 86.4% and linguistic discourse analysis methods, 89.6%. The reported results represent the consensus agreement achieved following discussion between all authors of this review.

Results

All results reported in this section of the article are based on analysis of the 165 included studies. As such, numbers represent a proportion of these 165 studies unless otherwise stated.

Timeline

With no time restrictions placed on the collection of linguistic discourse analysis studies, those identified as meeting the criteria for analysis spanned a period of 40 years (see Figure 2). The reporting of research utilising linguistic discourse analysis methods increased in the latter half of the 1990s by double the rate observed cumulatively over the previous 20 years. When breaking down this trend to assess the differences between descriptive analysis and treatment studies, the progression of descriptive research using linguistic discourse analysis plateaued over time. However, treatment studies continued to increase in number, with marked growth apparent in the late 2000s.
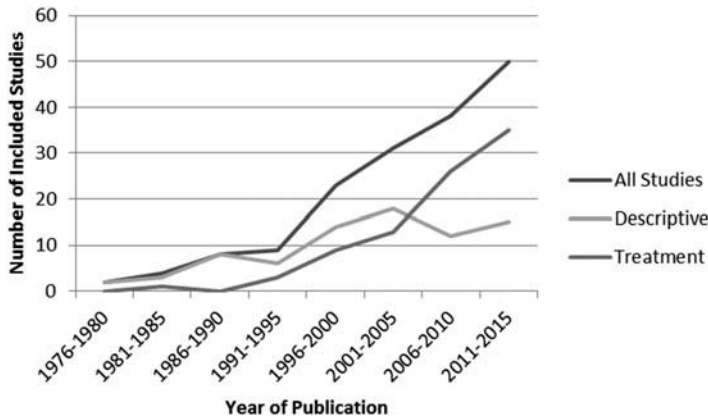
Figure 2. Number of included review studies by date of publication.

Purpose of application

The included studies collected during this review reflected two different research types: 87 treatment studies that used discourse to examine changes in language over time with intervention and 78 descriptive studies that described language impairment or analysis procedures. The type and scale of studies in each category are presented in Table 1.

Treatment studies

Within the 87 identified treatment studies, two types of studies were evident amongst those included. Seventy-five studies utilised within- and between-subject repeated analysis designs, comparing measures at baseline to post-treatment, illustrating changes in

Table 1. Average number of research participants by type of included study.

Study design	Number of studies	Participants – mean (SD; range)	Participants with aphasia* – mean (SD; range)
Descriptive studies			
Case studies/series	16	5.31 (5.86; 1–20)	3.31 (2.73; 1–10)
Single group	14	30.00 (30.86; 11–101)	29.93 (30.90; 11–101)
Two groups	36	43.64 (43.29; 8–200)	22.03 (21.15; 3–100)
Three groups	7	44.86 (27.82; 15–103)	23.57 (35.15; 5–103)
Four+ groups	5	74.60 (44.25; 20–124)	23.60 (21.82; 10–62)
Subtotal	78	35.42 (38.92; 1–200)	19.85 (23.81; 1–103)
Treatment studies			
AB design – case	10	2.90 (2.13; 1–6)	2.70 (2.00; 1–5)
AB design – group means	1	13.00 (-; 13)	13.00 (-; 13)
Pre/post – case	66	4.45 (3.88; 1–17)	3.80 (3.16; 1–17)
Pre/post – single-group outcomes	3	31.00 (36.43; 8–73)	28.00 (31.24; 8–64)
Pre/post – across groups	7	21.86 (16.43; 8–56)	21.86 (16.43; 8–56)
Subtotal	87	6.69 (10.34; 1–73)	6.07 (9.61; 1–64)
Total	165	20.27 (31.21; 1–200)	12.58 (19.03; 1–103)

*English-speaking individuals with aphasia as a result of left-hemisphere cerebrovascular accident (CVA).

connected language samples as a result of intervention. The remaining 12 studies used an AB design to evaluate language ability in the presence and absence of a therapy condition, such as a language scaffold in reciprocal scaffolding treatment, or the use of voice recognition software to target written language. Across both study types, 76 of the total 87 studies reported either single case or case series research, with measures of change calculated within individuals and then compared between cases. Linguistic discourse analysis was used within all 87 treatment studies in three conditions: as a primary outcome measure of intervention effects in 36 studies, as a secondary outcome measure in 19 studies and as a measure that aimed to determine if treatment effects, measured by other means, generalised to language in use in 37 of the included studies.

Descriptive studies

The 78 descriptive studies reported the use of discourse measures to describe the nature of language impairment in aphasia. Included studies classified as 'descriptive' utilised both group comparisons, single group, case series and case studies. Comparative studies analysed the differences between the language skills of those with and without impairments, of different severities of aphasia and of different ages. Forty-eight studies analysed between two and five different groups of individuals to understand the patterns of language impairment observed in aphasia. The remaining 30 studies analysed the language of single individuals or a single group of individuals.

Descriptive studies examined two primary aims in the use of linguistic discourse analysis with 10 studies focusing on multiple aims. Thirty-nine studies reported a case or comparisons between two or more participant populations in order to describe characteristics of language impairments, while 49 studies investigated the validity of components of the linguistic discourse analysis process for language assessment. In doing so, the sensitivity of analysis measures to the existence and severity of aphasic language impairments was investigated. A subset of these studies ($n = 14/42$) explored language sampling methods, and the effects that these had on the analysis of language structures. These studies investigated the effects that the genre, task instruction and context used to elicit language had on measured outcomes.

Review of included descriptive studies also identified 39 studies (of 78 descriptive studies) that stated or suggested in discussions and conclusions that linguistic discourse analysis had benefits that applied to its clinical application as an assessment of aphasia. For example, Armstrong and colleagues (2011) identified significant differences in the outcomes of linguistic analysis applied to monologic and dialogic discourse samples, indicating that the factors that affected the ability to produce dialogic discourse were multiple and complex. They concluded:

> 'The further exploration of such factors in clinical assessments of individuals with aphasia is hence essential if the clinician is to get a true picture of the speaker's overall linguistic skills' (pp. 1367).

Suggestions of clinical applicability of linguistic discourse analysis were evident in examples provided by researchers to illustrate how their methods could be used in the assessment and delivery of intervention to people with aphasia, for example, Capilouto, Wright and Wagovich (2006) wrote:

'From a clinical perspective the findings suggest, first, that the measurement of narrative discourse is an important aspect of the assessment of individuals with aphasia, because of the inherent importance of relaying the relationships between characters and events in daily life' (pp. 214–215).

Another 17 included descriptive studies used the assertion that use of linguistic discourse analysis is a vital component of clinical aphasia assessment as a rationale for the investigation illustrated in their research paper. Such statements were used to explain why the research was necessary for the discipline of speech pathology. For example, when discussing the rationale for investigating the performance of persons with aphasia on a particular discourse genre, Purdy (2002) stated that:

'Clinicians often infer how patients [with aphasia] will actually perform tasks or activities of daily living based on performance on procedural discourse tasks' (p. 174).

Discourse elicitation

All reviewed studies involved the collection of discourse to provide a language sample for analysis. In 158 studies, discourse was collected only in spoken format, while three studies collected only written discourse samples. Four studies collected both spoken and written language samples. Multiple methods were used across studies to elicit language samples for analysis (see Figure 3). Language collected ranged from structured, predictable responses to specific questions and stimuli, such as telling of the Cinderella story, to unstructured free conversation between familiar and unfamiliar individuals. In some cases, the same picture stimuli was used to elicit a structured sample; however the instruction provided with the presentation of the stimuli altered the elicited discourse genre. For example, single pictures could be used to elicit an expository description of content, or a narrative based on depicted events. Most samples used only structured language samples, whereas unstructured discourse was collected as the sole sample for analysis in 11 studies. Thirteen studies used both structured and unstructured language samples. The most commonly reported sampling
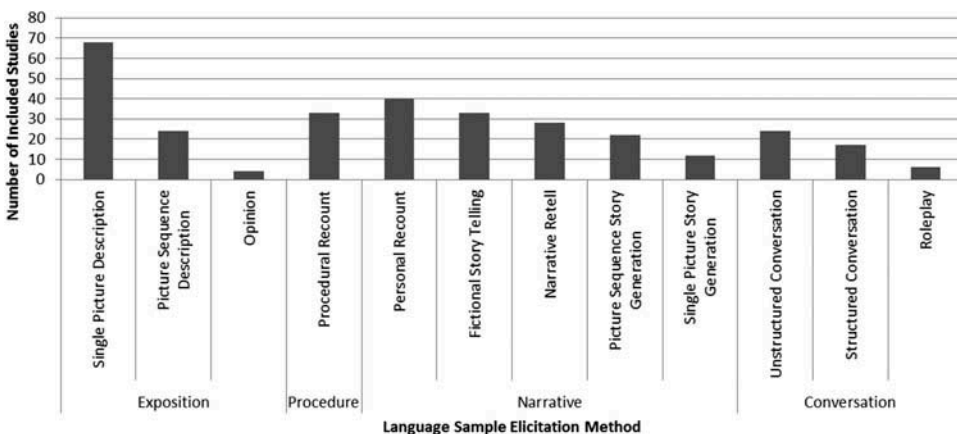
Figure 3. Methods used to elicit language samples in studies for review.

stimuli was the expository description of the Cookie Theft Picture from the Boston Diagnostic Aphasia Examination (BDAE) (Goodglass, Kaplan, & Barresi, 2001), which was utilised in 37 studies, closely followed by the telling of the Cinderella fairy tale after viewing a wordless picture book, used in 29 studies. Narratives were the most used discourse genre, utilised in 101 studies.

To elicit language samples, 27 studies used a single structured stimuli (e.g. The Cookie Theft picture) that elicited samples that could be easily compared between individuals (excluding personal narratives and structured conversation, which have highly variable content between individuals) and did not collapse this with other samples. Across these studies, six different elicitation methods were used, precluding the comparison of language samples and collation of data for meta-analytic purposes in this review. This resulted in an absence of language samples that could be compared across studies precluded collation of data for meta-analytic purposes in this review. Another 25 studies used a single elicited sample, though these samples had less structure and varied in content. The remaining 113 studies elicited language using multiple genres and topics including procedures, personal narratives, picture descriptions and narratives structured using picture sequences, and 77 of these collapsed the samples to form one larger corpus for their analysis. The protocol developed by Nicholas and Brookshire (1993), used in 19 studies, was the most used multiple-sample protocol across all included studies, with the Story Retell Procedure (Doyle et al., 1998) (10 studies) and AphasiaBank Protocol (MacWhinney, Fromm, Forbes, & Holland, 2011) (three studies) also used for this purpose.

Sample preparation

Transcription of samples was explicitly reported in 121 of the 162 studies using spoken language samples. In 50 of these studies, transcription was orthographic, with direct transcription of words as they were spoken. The remaining 71 studies did not specify transcription methods. This process was reportedly completed manually in each of these cases by a trained transcriptionist, typist, trained student, researcher or research assistant. Reliability of transcription was established through checking by a second individual prior to analysis in 49 studies. The remaining 113 studies did not report their process for establishing the reliability of transcription. Audio and/or visual recordings of spoken language samples were reported in 109 studies (of 162 spoken discourse studies), and 92.7% of those using recording also reported transcription. In 24 studies, parts of the transcribed language sample were removed prior to analysis to provide comparable sample lengths between individuals. Further, in 43 studies, other elements were removed from the language sample to prevent linguistic errors (such as paraphasias and unintelligible utterances) from affecting the outcome of analysis.

Discourse analysis methods

Once collected and transcribed, a large variety of methods were applied in order to analyse language samples. Across the 165 studies included in this review, 536 different measures (e.g. counts, proportions and ratios) of various linguistic elements were applied to describe various language behaviours observable through the analyses used. Due to this number and the variety of base units used in the analysis of samples (T-units, C-units, utterances,
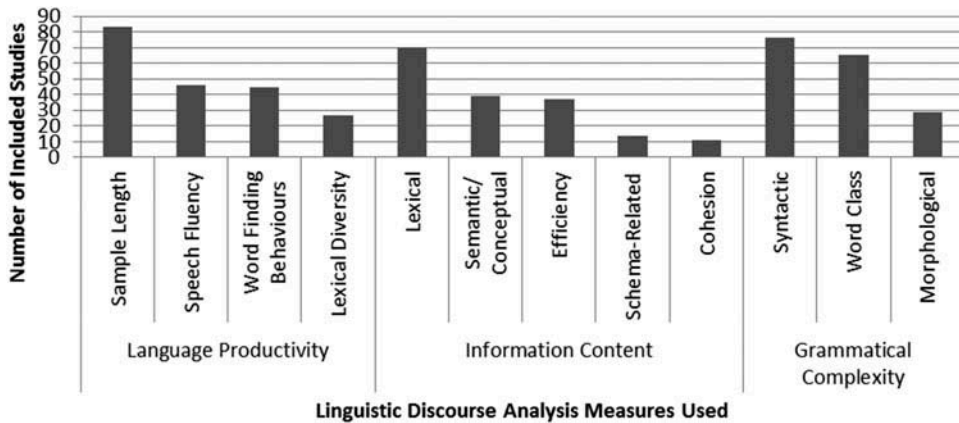
Figure 4. Number of studies utilising measures of linguistic structure in each category of analysis.

sentences, words, narrative words, etc.), measures were grouped into 12 categories of analysis which were related to three broad clusters of analytic foci: measures related to language productivity (sample length, lexical diversity, speech fluency, word finding behaviours); measures related to information content (efficiency, cohesion, lexical, semantic/conceptual, schema-related) and measures related to grammatical complexity (morphological, word class, syntactic), shown in Figure 4. Calculation of sample length in number of words was the most frequently applied measure ($n = 55$), followed by quantification of information content at the lexical level in the proportion of words that were correct information units (CIUs) (Nicholas & Brookshire, 1993) ($n = 43$). The calculation of measure outcomes was conducted manually by the authors in the majority of studies ($n = 147$). It was the sole method used in 140 studies (84.8%), indicating that it remained the dominant method of linguistic discourse analysis application. Computerised analysis tools were used to generate measure outcomes in 15.2% of studies ($n = 25$). The most frequently used computer software applications were Computerized Language Analysis (CLAN) (MacWhinney, 2000) (used in 12 studies) and Systematic Analysis of Language Transcripts (SALT) (Miller & Iglesias, 2012) (used in nine studies). Both of these tools calculate a suite of measures from language samples. Across all studies, 61.2% (101 of 165) reported that coding and/or analysis was checked by a second individual to determine the reliability of the reported results.

The large number of measures was reflected in the use of multiple analyses in many studies. While 28 studies used only a single category of analysis, only 10 applied a single measure within that category to the language samples collected. These 10 studies were distributed across treatment and descriptive studies. Where multiple analyses were used, some studies attempted to establish a system that provided a comprehensive linguistic assessment of discourse. Quantitative Production Analysis (QPA) (Saffran, Berndt, & Schwartz, 1989), the most frequently applied system used in 17 studies, combined a series of measures investigating sample length, speech fluency, lexical informativeness and grammatical complexity at all levels of analysis. Other established multi-measure systems, including the Shewan Spontaneous Language Analysis (SSLA) (Shewan, 1988), Profile of Word Errors and Retrieval in Speech (POWERS) (Herbert, Hickin, Howard, Osborne, & Best, 2008) and Language Assessment,

Remediation and Screening Procedure (LARSP) (Crystal, Fletcher, & Garman, 1976), were applied in no more than two studies each.

Summary of results

This review identified 165 studies that had applied linguistic analysis of discourse to assess the language of people with aphasia. Over the past 40 years, the number of studies using this method has increased from two identified studies in the late 1970s to 50 between 2011 and 2015. Applications of linguistic analysis of discourse fulfilled two key purposes. In 78 studies, linguistic analysis was used to evaluate the outcomes of language interventions, and in 87 studies, the method was used to describe the nature of language function and impairment in aphasia. The samples of discourse analysed across studies varied in the degree of structure. The most structured samples, expository language elicited using the description of a single picture, were used most frequently. However, narrative language samples were analysed in 101 studies, making it the most analysed genre of discourse. A total of 536 different linguistic measures were identified in the included studies, analysing language productivity, information content and grammatical complexity. The volume of language calculated by counting the number of words produced was the most frequently used measure, reported in 53 studies.

Discussion

Timeline of research

Research into the use of linguistic discourse analysis to assess the language of people with aphasia has been reported for the past 40 years. The amount of research conducted in this area over time has continued to increase providing growing evidence and insight to illustrate the process of linguistic discourse analysis application. Particular growth was evident in the number of treatment-based studies using linguistic discourse analysis over this period.

The growth in studies utilising linguistic discourse analyses reflects the growth evident in other areas of aphasiology research. For example, an analysis of studies utilising qualitative methods in aphasia research by Simmons-Mackie and Lynch (2013) demonstrated a fourfold increase in studies in the second half of the 20-year period of the review, which analysed studies from 1993 to 2013. The number of linguistic discourse analysis studies increases by more than five times in the second half of the review period for this study, though the review period was much longer. When making a direct comparison to the publication of qualitative aphasia studies, the rate of growth in linguistic discourse analysis research was less than half of that observed by Simmons-Mackie and Lynch (2013). The increase in treatment-based studies alone, however, closely mirrored the research growth observed in this comparable study.

The observed increase in the frequency of linguistic discourse analysis in the literature did not appear to be greater than increases in other areas of speech pathology and aphasiology research. Particularly in the domain of treatment studies, consistent use of linguistic discourse analysis as a language assessment was observable despite the variability and reported clinical limitations evident in the studies included in this review. This

highlighted several interesting findings that required further discussion, principally in the purposes for which linguistic discourse analysis was used, and the methods of analysis applied to measure language outcomes.

Purpose of application

It was found that nearly half of the research studies included in this review investigated the outcomes of various forms of language therapy, with linguistic discourse analysis applied to measure changes in language associated with intervention. The application of discourse in this context was necessary to indicate the social and ecological validity of interventions for aphasia (Ballard & Thompson, 1999). In both clinical and research environments, language treatments targeting language structures and impairments at the word and sentence levels are common. While therapy may improve targeted language structures, the use of linguistic discourse analysis can improve assessment with a view to social participation and activity and indicate if therapy has brought about change in the language used by persons with aphasia in their day-to-day communication. Studies applying linguistic discourse analysis to measure generalisation illustrate a means by which comparison of samples before and after treatment can be applied to demonstrate the effects of treatment on connected language. Where such generalisation measures are not applied, it remains unknown if the intervention has any effect on the language that is of central importance – that which is used to interact and participate in social life.

A particularly interesting finding was the frequency with which studies made reference to the clinical use of linguistic discourse analysis. These references were evident in descriptive studies where the focus of research was on linguistic discourse measures and language outcomes, rather than in treatment-based studies where the focus was placed on the particular intervention method. Often reported and widely recognised limitations to the clinical use of linguistic discourse analysis, particularly the time necessary to complete an analysis, were thought to restrict this form of language assessment to the domain of research. While assessment of discourse and connected language samples illustrated the ways in which language was impaired in the context of communication to achieve function in daily life, application of such practices to the clinical environment appeared to be impractical. However, with more than half of included descriptive studies either directly stating or suggesting that linguistic discourse analysis is a valid and used clinical tool, the reported limitations may not have the impact on clinical application that has been previously alleged. Future research will need to investigate the clinical use of linguistic discourse analysis to further understand this finding.

Discourse elicitation

An issue of particular interest in the elicitation of discourse for analysis was the frequency of structured sample use. The use of a language sample collected within the context of conversation, whether structured or unstructured, in nearly a third of studies provided an opportunity to collect a sample of the researcher's desired length without changes in discourse genre, by progression through conversational topics. Armstrong (2000) has suggested that as language used in conversation is the form of discourse used for human interaction, it is the most representative genre of language in use. However,

such samples varied considerably from individual to individual on the basis of topic and therefore semantic content, vocabulary and grammatical structure. This complicated comparison of conversational samples within and between individuals, and may suggest a reason for the apparent preference of researchers to use structured language samples.

The use of highly structured stimuli allowed for the collection of language that was predictable, and therefore comparable between individuals. Semantic analysis especially benefited from the similarity between sample content, with omission of important information apparent in reference to normal or expected responses (Yorkston & Beukelman, 1980; Nicholas & Brookshire, 1995). Single-picture descriptions in particular were the most used elicitation stimuli and provided the most structured language samples by restricting the content and structure of language that could be used in responses. Despite being the most applied method, outcomes have suggested that the amount of language generated by such stimuli was insufficient for the examination of discourse ability in aphasia, typically resulting in a descriptive listing of observed items, rather than a narrative sample that illustrated the grammatical and cohesive structures that were central to discourse construction (Olness, 2006; Armstrong, Ferguson & Simmons-Mackie, 2013).

While the use of structured samples simplified language sample collection and linguistic discourse analysis, the absence of some discourse structures and the length of sample produced highlighted the issue of representation. A single expositional picture description may not demonstrate actual linguistic ability as would be evident in conversational language. Brookshire and Nicholas (1994) suggested that the combination of multiple structured language samples including expositional picture descriptions, procedures, picture-supported narratives and personal narrative recounts generated a language sample that better represented actual language use. Multiple included studies collapsed language samples collected using different elicitation methods to generate a single language sample for analysis. This combination of samples also generated the 300- to 400-word samples that Nicholas and Brookshire reported as necessary for reliable analysis (Brookshire & Nicholas, 1994). However, while this generated a greater volume of language, it created an additional source of uncontrolled variability. Different discourse genres varied in the lexical and grammatical complexity of responses (Li, Volpe, Ritterman, & Williams, 1996), and within genres the variability occurred as a function of familiarity and difficulty of the stimuli presented (Williams, Li, Volpe, & Ritterman, 1994). As such, collapsing multiple discourse genres into a single language sample may have led to difficulty in one particular genre being disguised by a lesser impairment in the linguistic structures required to produce a less complicated sample.

Most studies included in this review reported the use of recording and transcription of language samples prior to analysis. The application of these steps during the linguistic discourse analysis process permits greater accuracy in the results obtained from the assessment of language samples by permitting review of samples and ease of access for measurement of linguistic structures (Ball, Howard, Müller, & Granese, 2013; Rutter & Cunningham, 2013). As recording and orthographic transcription (i.e. transcription of words) is integral to determining the accuracy and reliability of analysis, it is noteworthy that so few studies reported that transcripts had been examined to determine accuracy. Similarly, only just over half of included studies reported that the consistency and reliability of their results had been determined. Without a means of establishing the

accuracy of the language samples analysed and the linguistic outcomes reported, the precision and replicability of reported research cannot be assured. An agreed method to perform transcription and determine transcription reliability may be required to improve the linguistic discourse analysis process. This is of particular importance when considering that analysis can only be as accurate as the language sample transcript. The time and resources available when completing the research may be a barrier to this process as a second individual is needed to repeat the time-consuming transcription and analysis process. Despite this, perhaps the reporting of inter- and intra-rater agreement should be a requirement in research studies utilising linguistic discourse analysis, particularly where it is used as a primary outcome measure to ensure the rigor of the analysis.

Discourse analysis methods

The large number of linguistic discourse analysis measures identified through this review highlighted a key issue impacting on clinical translation of linguistic discourse analysis research in aphasia. Selecting appropriate analysis measures for the assessment of language is made more complex by the wealth of options identified in the research literature. Such a substantial variety of linguistic measures were necessary across the studies as each permitted the measurement of a unique component of linguistic structure across the domains of language productivity, information content and grammatical structure. The greater number of measures than studies occurred due to the use of multiple analyses in the overwhelming majority of studies. In most cases, these measures defined language structures across multiple aspects of discourse structure and provided a comprehensive assessment of language. For example, language volume defined in the number of words, fluency in words per minute, information content in the number and proportion of CIUs and efficiency in CIUs per minute were often used together and provided greater detail than a single measure (Nicholas & Brookshire, 1993). This provided a systematic and detailed report of language production ability, specific to the individual as no single measure could capture the complexity of language as a whole (Saffran et al., 1989). Where fewer measures were used, the result targeted a more specific domain of language and may not have provided a comprehensive view of a person with aphasia's language function and impairment. This more specific analysis was generally selected by researchers to meet the specific research aims and goals of their study. For example, analysis of the number of verb form errors and the number of semantically accurate verbs used occurred only in a study examining the effects of a verb retrieval therapy (Maul, Conner, Kempler, Radvanski, & Goral, 2014).

To encourage the use of a more comprehensive range of linguistic measures, researchers such as Saffran and colleagues (1989) proposed multi-measure systems such as QPA. These systems were designed to provide greater detail to the linguistic assessment of language samples, though they appeared to focus more specifically on grammatical structures and outcomes (McEntee & Kennedy, 1995; Gordon, 2006). The absence of language productivity and information content measures within these systems may have limited their more widespread use as they were unable to provide a complete picture of language. Researchers appeared to select measures tailored to their needs rather than use a broad system that provided results that were not specific to the investigated research outcomes. This was particularly evident in treatment-based studies where elements of a

multi-measure method were often used, rather than the complete system. Descriptive studies appeared to use more complete systems when investigating a concept such as agrammatism, rather than the outcomes of a therapy that targeted a single grammatical structure.

The most frequently applied analysis methods defined linguistic behaviours at the word and sentence levels, measuring the volume of language produced, the fluency of production and the grammatical characteristics. These measures provided examination of the sentence-level structures that are typically assessed using standardised assessments (Chapman & Ulatowska, 1992). These particular measures may have been applied more frequently within the specific population of people with aphasia due to the oft studied and reported nature of aphasic impairment, whereby sentential structures are affected while higher level semantic and schematic structures may be preserved (Ulatowska, Doyle, Freedman-Stern, & Macaluso-Haynes, 1983; Ulatowska, Allard, Reyes, Ford, & Chapman, 1992). However, while evaluating the ability to use these structures in the context of language in use is important, it was interesting to note the limited use of measures that quantified aspects of cohesion, and information content at the schema-related level. Measures of cohesion in particular were observed in the fewest studies despite breakdowns in cohesion previously being recognised as an area of aphasic impairment (Lock & Armstrong, 1997). As cohesive and schema-related structures can only be assessed at the discourse level (Sherratt, 2007), measures that examined them specifically may have been expected more frequently in studies of linguistic discourse analysis in aphasia.

While the large variation observed in discourse measures served a purpose in quantifying linguistic structures of language productivity, information content and grammatical context, it also raised questions about the nature of linguistic discourse analysis in research. The analysis of grammatical complexity in particular highlighted issues in the variation between measures. Where even the base unit of analysis was not consistent across linguistic discourse analysis applications – including phrases, clauses, T-units, C-units, utterances and sentences – measures themselves could not be compared between studies. The absence of an agreed analysis procedure precluded a meta-analysis of discourse outcomes across research studies, and, clinically, the large number of measures applied without any indication of those which may be most beneficial to the assessment of aphasia potentially confused the implementation of evidence to practice. While further investigation may determine if an agreed core set of linguistic analysis measures could be used to establish a recognised assessment of discourse in aphasia, this would perhaps limit the efficiency of such a functional approach to language assessment. The included studies demonstrated that linguistic measures were typically selected to meet a specific goal, or measure a specific therapy target. An agreed form of analysis, while simplifying the interpretation of results, could increase the inefficiency of discourse analysis as it would introduce measures that may not be suited to all aphasia types and severities, and may not be appropriate to measure the outcomes of activity and social participation goals at which therapy is aimed.

In the current computer age, it was interesting to note that most studies used only manual analysis of language transcripts. Despite the continued increase in the availability of automated tools since they first became available 20 years ago, their use remained limited in included studies. These computer programs, particularly SALT (Miller &

Iglesias, 2012) and CLAN (MacWhinney, 2000), provided researchers with a means to analyse multiple linguistic structures at the lexical and morpho-syntactic levels simultaneously. The reliability of such tools was reported across studies, suggesting that they could be highly effective to facilitate a reliable and efficient linguistic discourse analysis. While the cost of software may have posed a barrier to the use of some programs, others were freely available downloads that were available for access by researchers (MacWhinney, 2000). The training necessary to appropriately use these tools may also have posed a barrier to more widespread use in the research literature, with specific transcription conventions needed to ensure the program can accurately identify target behaviours (MacWhinney, 2000; Miller & Iglesias, 2012). However, as many studies reported the use of appropriate transcription conventions even in the absence of computerised analysis, it is unlikely that this particular barrier was a significant limiter of use.

The use of computer-assisted measures may also improve the clinical and research feasibility of linguistic discourse analysis by reducing the amount of time necessary to complete an adequate assessment (Frederiksen, Bracewell, Breuleux, & Renaud, 1990). For computerised tools, training is needed for only the transcription process as computer-assisted analysis removes the need for expertise related to each specific analysis completed, making clinical education to assess discourse more viable. The completion of a multi-level assessment with multiple measures calculated simultaneously from a single transcript generates the necessary time efficiency that does not exist during manual discourse analysis (Togher, 2001). Should computer-assisted transcription tools using voice-to-text software also prove effective in future trials, the efficiency of computer-assisted analysis would be further supported. This could address reported concerns and limitations to facilitate translation of linguistic discourse analysis to clinical aphasia practice.

Limitations

It should be noted that this review examined only the methodological applications of linguistic discourse analysis to assess language in aphasia. As only the methodology was investigated, no critical evaluation of included studies was completed. Studies were not evaluated to examine potential bias that may have influenced discourse outcomes as only the assessment protocol was examined. All observed applications of linguistic discourse analysis were included in the review without examination of the quality of the research reported. While the results of this review included evaluation of analyses that may not have been implemented or reported in a rigorous manner, the results that were obtained revealed a comprehensive picture of the research applications of discourse. Future research may look to critically examine discourse applications and outcomes to provide further insight into the use and reporting of such measures.

The focus on methodology resulted in large number of studies identified through the literature search. The limited resources available to examine the wealth the data collected influenced the decision to have only one researcher examining studies for inclusion in the research and extracting relevant information regarding the use of linguistic discourse analysis. The use of two examiners may have improved the rigor of the review process. The decision was made to have a second reviewer examine a portion of included studies to determine agreement in the classification of the purposes of discourse use, language elicitation methods and analysis measures used, in line with a previously reported

methodological review (Simmons-Mackie & Lynch, 2013). The reliability of classification of analysis measures in particular was initially low, so further discussion and training was implemented to establish consensus agreement on the classification of measures. Some disagreement remained due to fundamental differences in the interpretation of measures, and the linguistic focus used by researchers in their applications of linguistic discourse analysis. For example, the analysis of the number of different nouns in a sample could be interpreted as either a measure of grammar in word class use, or as a measure of linguistic diversity. As the interpretation of linguistic measures and their classification was somewhat subjective, this type on ambiguity and therefore disagreement in the classification of measures was to be expected.

Conclusions

This review revealed a large volume of research studies that have applied linguistic discourse analysis to assess language in use in aphasia. These studies suggested that over time, despite many reported barriers, the use of linguistic discourse analysis has continued, and studies in fact continue to propose that this pattern of use may reflect similar ongoing application in clinical contexts. The use of discourse to examine generalisation of therapy outcomes was a particular positive trend, encouraging the development of interventions with functional outcomes for people with aphasia. However, this review also revealed shortcomings in the application and reporting of linguistic discourse analysis in the speech pathology literature. The lack of agreed processes for conducting and reporting transcription and analysis resulted in large variety of observed procedures, particularly for language analysis. As such, questions were raised regarding the need for agreed procedures that could facilitate the implementation of research in clinical practice. Further, the incidental finding that computer-assisted analysis, while used minimally to date, may support greater efficiency in the analysis process requires greater examination. Investigating this in depth may provide detailed information to better understand and facilitate the translation of linguistic discourse analysis research to the clinical assessment of aphasia within the speech pathology profession. Such developments will further align assessment procedures with the ICF to improve the social participation and functional activity of people with aphasia.

Acknowledgments

Declaration of interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of this paper.

References

Armstrong, E. (2000). Aphasic discourse analysis: the story so far. *Aphasiology*, *14*(9), 875–892.

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

Armstrong, E., Ciccone, N., Godecke, E., & Kok, B. (2011). Monologues and dialogues in aphasia: some initial comparisons. *Aphasiology*, *25*(11), 1347–1371.

Armstrong, E., Ferguson, A., & Simmons-Mackie, N. (2013). Discourse and functional approaches to aphasia. In I. Papathanasiou, P. Coppens & C. Potagas (Eds.), *Aphasia and related neurogenic communication disorders* (pp. 217–232). Boston, MA: Jones and Bartlett Publishers.

Ball, M. J., Howard, S., Müller, N., & Granese, A. (2013). Data processing: transcriptional and impressionistic methods. In N. Müller & M. J. Ball (Eds.), *Research methods in clinical linguistics and phonetics: a practical guide* (pp. 177–194). West Sussex, UK: Wiley-Blackwell.

Ballard, K. J., & Thompson, C. K. (1999). Treatment and generalization of complex sentence production in agrammatism. *Journal of Speech Language & Hearing Research*, *42*(3), 690–707.

Beeke, S., Maxim, J., & Wilkinson, R. (2008). Rethinking agrammatism: factors affecting the form of language elicited via clinical test procedures. *Clinical Linguistics & Phonetics*, *22*(4–5), 317–323.

Boles, L. (1998). Conversational discourse analysis as a method for evaluating progress in aphasia: a case report. *Journal of Communication Disorders*, *31*, 261–274.

Brady, M., Kelly, H., Godwin, J., & Enderby, P. (2012). Speech and language therapy for aphasia following stroke. *Cochrane Database of Systematic Reviews*, *16*(5). doi:10.1002/14651858. CD000425.pub3.

Brookshire, R. H., & Nicholas, L. E. (1994). Speech sample-size and test-retest stability of connected speech measures for adults with aphasia. *Journal of Speech and Hearing Research*, *37*(2), 399–407.

Capilouto, G. J., Wright, H. H., & Wagovich, S. A. (2006). Reliability of main event measurement in the discourse of individuals with aphasia. *Aphasiology*, *20*(2–4), 205–216.

Chapman, S. B., & Ulatowska, H. K. (1992). Methodology for discourse management in the treatment of aphasia. *Clinical Communication Disorders*, *2*(1), 64–81.

Cherney, L. R. (1998). Pragmatics and discourse: an introduction. In L. R. Cherney, B. B. Shadden & C. A. Coelho (Eds.), *Analyzing discourse in communicatively impaired adults* (pp. 1–8). Gaithersburg, Maryland: Aspen Publishers.

Coelho, C. A., Grela, B., Corso, M., Gamble, A., & Feinn, R. (2005). Microlinguistic deficits in the narrative discourse of adults with traumatic brain injury. *Brain Injury*, *19*(13), 1139–1145.

Crystal, D., Fletcher, P., & Garman, M. (1976). *The grammatical analysis of language disability*. London: Edward Arnold.

Doyle, P. J., McNeil, M. R., Spencer, K. A., Goda, A. J., Cottrell, K., & Lustig, A. P. (1998). The effects of concurrent picture presentations on retelling of orally presented stories by adults with aphasia. *Aphasiology*, *12*(7–8), 561–574.

Elia, D., Liles, B. Z., Duffy, R. J., Coelho, C. A., & Belanger, S. A. (1994). *An investigation of sample size in conversational analysis*. Paper presented at the ASHA Convention, New Orleans, Louisiana.

Frederiksen, C. H., Bracewell, R. J., Breuleux, A., & Renaud, A. (1990). The cognitive representation and processing of discourse: function and dysfunction. In Y. Joanette & H. H. Brownwell (Eds.), *Discourse ability and brain damage: theoretical and empirical perspectives* (pp. 69–112). New York, NY: Springer-Verlag.

Goodglass, H., Kaplan, E., & Barresi, B. (2001). *Boston Diagnostic Aphasia Examination* (3rd ed.). Philadelphia: Lippencott, Williams & Wilkins.

Gordon, J. K. (2006). A quantitative production analysis of picture description. *Aphasiology*, *20*(2–4), 188–204.

Herbert, R., Hickin, J., Howard, D., Osborne, F., & Best, W. (2008). Do picture-naming tests provide a valid assessment of lexical retrieval in conversation in aphasia? *Aphasiology*, *22*(2), 184–203.

Huber, W. (1990). Text comprehension and production in aphasia: analysis in terms of micro- and macroprocessing. In Y. Joanette & H. H. Brownwell (Eds.), *Discourse ability in brain damage: theoretical and empirical perspectives* (pp. 154–179). New York, NY: Springer-Verlag.

Jaecks, P., Hielscher-Fastabend, M., & Stenneken, P. (2012). Diagnosing residual aphasia using spontaneous speech analysis. *Aphasiology*, 26(7), 953–970.

Kemper, S., McDowd, J., Pohl, P., Herman, R., & Jackson, S. (2006). Revealing language deficits following stroke: the cost of doing two things at once. *Aging, Neuropsychology, and Cognition*, 13 (1), 115–139.

Li, E. C., Volpe, A. D., Ritterman, S., & Williams, S. E. (1996). Variations in grammatic complexity across three types of discourse. *Journal of Speech-Language Pathology and Audiology*, 30, 180–186.

Lock, S., & Armstrong, L. (1997). Cohesion analysis of the expository discourse of normal, fluent aphasic and demented adults: a role in differential diagnosis? *Clinical Linguistics & Phonetics*, 11 (4), 299–317.

MacWhinney, B. (2000). *The CHILDES Project: tools for analysing talk* (3rd ed.). Muhwah, NJ: Lawrence.

MacWhinney, B., Fromm, D., Forbes, M., & Holland, A. (2011). AphasiaBank: methods for studying discourse. *Aphasiology*, 25(11), 1286–1307.

Marini, A., Andreetta, S., del Tin, S., & Carlomagno, S. (2011). A multi-level approach to the analysis of narrative language in aphasia. *Aphasiology*, 25(11), 1372–1392.

Maul, K. K., Conner, P. S., Kempler, D., Radvanski, C., & Goral, M. (2014). Using informative verbal exchanges to promote verb retrieval in nonfluent aphasia. *American Journal of Speech-Language Pathology*, 23(4), 407–420.

Mayer, J., & Murray, L. (2003). Functional measures of naming in aphasia: word retrieval in confrontation naming versus connected speech. *Aphasiology*, 17(5), 481–497.

McEntee, L. J., & Kennedy, M. (1995). Profiling agrammatic spoken language: towards a government and binding framework. *European Journal of Disorders of Communication*, 30 (3), 317–332.

Miller, J., & Iglesias, A. (2012). Systematic Analysis of Language Transcripts (SALT) (Version Research version 2012). Middleton, WI: SALT Software, LLC.

Murray, L. L., Holland, A. L., & Beeson, P. M. (1998). Spoken language of individuals with mild fluent aphasia under focused and divided-attention conditions. *Journal of Speech Language & Hearing Research*, 41(1), 213–227.

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(2), 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.

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

Prins, R., & Bastiaanse, R. (2004). Analysing the spontaneous speech of aphasic speakers. *Aphasiology*, 18(12), 1075–1091.

Purdy, M. H. (2002). Script knowledge following stroke. *Journal of Medical Speech-Language Pathology*, 10(3), 173–181.

Rutter, B., & Cunningham, S. (2013). The recording of audio and video data. In N. Müller & M. J. Ball (Eds.), *Research methods in clinical linguistics and phonetics: a practical guide* (pp. 160–176). West Sussex, UK: Wiley-Blackwell.

Sacks, H., Schegloff, E. A., & Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, 50, 696–735.

Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (1989). The quantitative analysis of agrammatic production: procedure and data. *Brain & Language*, 37(3), 440–479.

Sherratt, S. (2007). Multi-level discourse analysis: a feasible approach. *Aphasiology*, 21(3–4), 375–393.

Shewan, C. M. (1988). The Shewan Spontaneous Language Analysis (SSLA) system for aphasic adults: description, reliability and validity. *Journal of Communication Disorders*, 21, 103–138.

Simmons-Mackie, N., & Lynch, K. (2013). Qualitative research in aphasia: a review of the literature. *Aphasiology*, 27(11), 1281–1301.

Togher, L. (2001). Discourse sampling in the 21st century. *Journal of Communication Disorders*, *34* (1–2), 131–150.

Ulatowska, H. K., Allard, L., & Chapman, S. B. (1990). Narrative and procedural discourse in aphasia. In Y. Joanette & H. H. Brownwell (Eds.), *Discourse ability and brain damage: theoretical and empirical perspectives* (pp. 180–198). New York, NY: Springer-Verlag.

Ulatowska, H. K., Allard, L., Reyes, B. A., Ford, J., & Chapman, S. (1992). Conversational discourse in aphasia. *Aphasiology*, *6*(3), 325–330.

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

Ulatowska, H. K., Olness, G. S., Wertz, R. T., Samson, A. M., Keebler, M. W., & Goins, K. E. (2003). Relationship between discourse and Western Aphasia Battery performance in African Americans with aphasia. *Aphasiology*, *17*(5), 511–521.

Williams, S. E., Li, E. C., Volpe, A. D., & Ritterman, S. (1994). The influence of topic and listener familiarity on aphasic discourse. *Journal of Communication Disorders*, *27*, 207–222.

World Health Organization. (2011). *ICF: International Classification of Functioning, Disability and Health*. Geneva, Switzerland: World Health Organization.

Yorkston, K. M., & Beukelman, D. R. (1980). An analysis of connected speech samples of aphasic and normal speakers. *Journal of Speech & Hearing Disorders*, *45*(1), 27–36.

Appendix A. Studies included in the review of linguistic discourse analysis in aphasia

Abuom, T. O., & Bastiaanse, R. (2012). Characteristics of swahili-english bilingual agrammatic spontaneous speech and the consequences for understanding agrammatic aphasia. *Journal of Neurolinguistics, 25*(4), 276–293.

Albright, E., & Purves, B. (2008). Exploring SENTENCESHAPER: treatment and augmentative possibilities. *Aphasiology, 22*(7–8), 741–752.

Altmann, L. J. P., Hazamy, A. A., Carvajal, P. J., Benjamin, M., Rosenbek, J. C., & Crosson, B. (2014). Delayed stimulus-specific improvements in discourse following anomia treatment using an intentional gesture. *Journal of Speech Language and Hearing Research, 57*(2), 439–454.

Antonucci, S. M. (2009). Use of semantic feature analysis in group aphasia treatment. *Aphasiology, 23*(7–8), 854–866.

Armstrong, E. (2001). Connecting lexical patterns of verb usage with discourse meanings in aphasia. *Aphasiology, 15*(10–11), 1029–1045.

Armstrong, E., Ciccone, N., Godecke, E., & Kok, B. (2011). Monologues and dialogues in aphasia: some initial comparisons. *Aphasiology, 25*(11), 1347–1371.

Avent, J., & Austermann, S. (2003). Reciprocal scaffolding: a context for communication treatment in aphasia. *Aphasiology, 17*(4), 397–404.

Avent, J. R. (1997). Group treatment in aphasia using cooperative learning methods. *Journal of Medical Speech-Language Pathology, 5*(1), 9–26.

Ballard, K. J., & Thompson, C. K. (1999). Treatment and generalization of complex sentence production in agrammatism. *Journal of Speech Language & Hearing Research, 42*(3), 690–707.

Bates, E., Reilly, J., Wulfeck, B., Dronkers, N., Opie, M., Fenson, J.,..., Herbst, K. (2001). Differential effects of unilateral lesions on language production in children and adults. *Brain & Language, 79* (2), 223–265.

Bloom, R. L., Borod, J. C., Santschi-Haywood, C., Pick, L. H., & Obler, L. K. (1996). Left and right hemispheric contributions to discourse coherence and cohesion. *International Journal of Neuroscience, 88*(1–2), 125–140.

Boles, L. (2000). Aphasia therapy in a bilingual speaker: treatment in language one, with spousal support in language two. *Asia Pacific Journal of Speech, Language & Hearing, 5*(2), 137–142.

Boles, L. (2015). Establishing alignment in aphasia couples therapy in a woman with wernicke's aphasia. *Communication Disorders Quarterly, 36*(4), 219–230.

Boles, L., & Bombard, T. (1998). Conversational discourse analysis: appropriate and useful sample sizes. *Aphasiology, 12*(7–8), 547–560.

Boo, M., & Rose, M. L. (2011). The efficacy of repetition, semantic, and gesture treatments for verb retrieval and use in Broca's aphasia. *Aphasiology, 25*(2), 154–175.

Boyle, M. (2014). Test–retest stability of word retrieval in aphasic discourse. *Journal of Speech, Language & Hearing Research, 57*(3), 966–978.

Breedin, S. D., Saffran, E. M., & Schwartz, M. F. (1998). Semantic factors in verb retrieval: an effect of complexity. *Brain & Language, 63*(1), 1–31.

Brodsky, M. B., McNeil, M. R., Doyle, P. J., Fossett, T. R. D., Timm, N. H., & Park, G. H. (2003). Auditory serial position effects in story retelling for non-brain-injured participants and persons with aphasia. *Journal of Speech Language and Hearing Research, 46*(5), 1124–1137.

Brookshire, R. H., & Nicholas, L. E. (1994). Speech sample-size and test-retest stability of connected speech measures for adults with aphasia. *Journal of Speech and Hearing Research, 37*(2), 399–407.

Bruce, C., Edmundson, A., & Coleman, M. (2003). Writing with voice: an investigation of the use of a voice recognition system as a writing aid for a man with aphasia. *International Journal of Language & Communication Disorders, 38*(2), 131–148.

Bryant, L., Spencer, E., Ferguson, A., Craig, H., Colyvas, K., & Worrall, L. (2013). Propositional idea density in aphasic discourse. *Aphasiology, 27*(8), 992–1009.

Busch, C. R., Brookshire, R. H., & Nicholas, L. E. (1988). Referential communication by aphasic and nonaphasic adults. *Journal of Speech & Hearing Disorders, 53*(4), 475–482.

Cahana-Amitay, D., Oveis, A. C., Sayers, J. T., Pineles, S. L., Spiro, A., & Albert, M. L. (2015). Biomarkers of "Linguistic Anxiety" in aphasia: a proof-of-concept case study. *Clinical Linguistics & Phonetics, 29*(5), 401–413.

Cameron, R. M., Wambaugh, J. L., & Mauszycki, S. C. (2010). Individual variability on discourse measures over repeated sampling times in persons with aphasia. *Aphasiology, 24*(6–8), 671–684.

Cameron, R. M., Wambaugh, J. L., Wright, S. M., & Nessler, C. L. (2006). Effects of a combined semantic/phonologic cueing treatment on word retrieval in discourse. *Aphasiology, 20*(2–4), 269–285.

Capilouto, G., & Wright, H. H. (2009). Scripting information components to improve narrative discourse performance. *Journal of Medical Speech-Language Pathology, 17*(2), 99–110.

Capilouto, G. J., Wright, H. H., & Wagovich, S. A. (2006). Reliability of main event measurement in the discourse of individuals with aphasia. *Aphasiology, 20*(2–4), 205–216.

Carlomagno, S., Losanno, N., Emanuelli, S., & Casadio, P. (1991). Expressive language recovery or improved communicative skills - effects of PACE therapy on aphasics referential communication and story retelling. *Aphasiology, 5*(4–5), 419–424.

Carragher, M., Sage, K., & Conroy, P. (2013). The effects of verb retrieval therapy for people with non-fluent aphasia: evidence from assessment tasks and conversation. *Neuropsychological Rehabilitation, 23*(6), 846–887.

Carragher, M., Sage, K., & Conroy, P. (2015a). Outcomes of treatment targeting syntax production in people with Broca's-type aphasia: evidence from psycholinguistic assessment tasks and every-day conversation. *International Journal of Language & Communication Disorders, 50*(3), 322–336.

Carragher, M., Sage, K., & Conroy, P. (2015b). Preliminary analysis from a novel treatment targeting the exchange of new information within storytelling for people with nonfluent aphasia and their partners. *Aphasiology, 29*(22), 1381–1408.

Cherepski, M. A., & Drummond, S. S. (1987). Linguistic description in nonfluent dysphasia: utilization of pictograms. *Brain and Language, 30*(2), 285–304.

Cherney, L. R. (2010a). Oral Reading for Language in Aphasia (ORLA): evaluating the efficacy of computer-delivered therapy in chronic nonfluent aphasia. *Topics in Stroke Rehabilitation, 17*(6), 423–431.

Cherney, L. R. (2010b). Oral Reading for Language in Aphasia: impact of aphasia severity on cross-modal outcomes in chronic nonfluent aphasia. *Seminars in Speech & Language, 31*(1), 42–51.

Cherney, L. R., Halper, A. S., Holland, A. L., & Cole, R. (2008). Computerized script training for aphasia: preliminary results. *American Journal of Speech-Language Pathology, 17*(1), 19–34. doi:10.1044/1058-0360(2008/003).

Cherney, L. R., Harvey, R. L., Babbitt, E. M., Hurwitz, R., Kaye, R. C., Lee, J. B., & Small, S. L. (2012). Epidural cortical stimulation and aphasia therapy. *Aphasiology, 26*(9), 1192–1217.

Conroy, P., Sage, K., & Ralph, M. L. (2009). Improved vocabulary production after naming therapy in aphasia: can gains in picture naming generalize to connected speech? *International Journal of Language & Communication Disorders, 44*(6), 1036–1062.

Correia, L., Brookshire, R. H., & Nicholas, L. E. (1990). Aphasic and non-brain-damaged adults' descriptions of aphasia test pictures and gender-biased pictures. *Journal of Speech & Hearing Disorders, 55*(4), 713–720.

Craig, H. K., Hinckley, J. J., Winkelseth, M., Carry, L., Walley, J., Bardach, L.,…, Sheimo, D. (1993). Quantifying connected speech samples of adults with chronic aphasia. *Aphasiology, 7*(2), 155–163.

Damico, J., Tetnowski, J., Lynch, K., Hartwell, J., Weill, C., Heels, J., & Simmons-Mackie, N. (2014). Facilitating authentic conversation: an intervention employing principles of constructivism and conversation analysis. *Aphasiology*, 1–22. doi:10.1080/02687038.2014.945388.

Davis, C. H., & Harrington, G. (2006). Intensive semantic intervention in fluent aphasia: a pilot study with fMRI. *Aphasiology, 20*(1), 59–83.

del Toro, C. M., Altmann, L. J. P., Raymer, A. M., Leon, S., Blonder, L. X., & Gonzalez Rothi, L. J. (2008). Changes in aphasic discourse after contrasting treatments for anomia. *Aphasiology, 22*(7–8), 881–892.

Doyle, P. J., McNeil, M. R., Park, G., Goda, A., Rubenstein, E., Spencer, K.,…, Szwarc, L. (2000). Linguistic validation of four parallel forms of a story retelling procedure. *Aphasiology, 14*(5–6), 537–549.

Doyle, P. J., McNeil, M. R., Spencer, K. A., Goda, A. J., Cottrell, K., & Lustig, A. P. (1998). The effects of concurrent picture presentations on retelling of orally presented stories by adults with aphasia. *Aphasiology, 12*(7–8), 561–574.

Doyle, P. J., Tsironas, D., Goda, A. J., & Kalinyak, M. (1996). The relationship between objective measures and listeners' judgments of the communicative informativeness of the connected discourse of adults with aphasia. *American Journal of Speech-Language Pathology, 5*(3), 53–60.

Edmonds, L. A., & Babb, M. (2011). Effect of verb network strengthening treatment in moderate-to-severe aphasia. *American Journal of Speech-Language Pathology, 20*(2), 131–145.

Edmonds, L. A., Obermeyer, J., & Kernan, B. (2014). Investigation of pretreatment sentence production impairments in individuals with aphasia: towards understanding the linguistic variables that impact generalisation in Verb Network Strengthening Treatment. *Aphasiology*, 1–33. doi:10.1080/02687038.2014.975180.

Edwards, S. (1995). Profiling fluent aphasic spontaneous speech: a comparison of two methodologies. *European Journal of Disorders of Communication, 30*(3), 333–345.

Edwards, S., & Bastiaanse, R. (1998). Diversity in the lexical and syntactic abilities of fluent aphasic speakers. *Aphasiology, 12*(2), 99–117.

Edwards, S., & Tucker, K. (2006). Verb retrieval in fluent aphasia: a clinical study. *Aphasiology, 20*(7), 644–675.

Ernestbaron, C. R., Brookshire, R. H., & Nicholas, L. E. (1987). Story structure and retelling of narratives by aphasic and non-brain-damaged adults. *Journal of Speech and Hearing Research, 30*(1), 44–49.

Falconer, C., & Antonucci, S. M. (2012). Use of semantic feature analysis in group discourse treatment for aphasia: extension and expansion. *Aphasiology, 26*(1), 64–82.

Faroqi-Shah, Y., & Virion, C. R. (2009). Constraint-induced language therapy for agrammatism: role of grammaticality constraints. *Aphasiology, 23*(7–8), 977–988.

Faroqi-Shah, Y., & Waked, A. N. (2010). Grammatical category dissociation in multilingual aphasia. *Cognitive Neuropsychology, 27*(2), 181–203.

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

Fergadiotis, G., & Wright, H. H. (2015). Modelling confrontation naming and discourse performance in aphasia. *Aphasiology*, 1–17. doi:10.1080/02687038.2015.1067288.

Fergadiotis, G., Wright, H. H., & West, T. M. (2013). Measuring lexical diversity in narrative discourse of people with aphasia. *American Journal of Speech-Language Pathology, 22*(2), S397–S408.

Fillingham, J., & Lum, C. (1998). A modality-specific mapping impairment: spoken versus written production. *International Journal of Language & Communication Disorders, 33*(S1), 196–201.

Fink, R. B., Bartlett, M. R., Lowery, J. S., Linebarger, M. C., & Schwartz, M. F. (2008). Aphasic speech with and without SENTENCESHAPER: two methods for assessing informativeness. *Aphasiology, 22*(7–8), 679–690.

Fridriksson, J., Hubbard, H. I., Hudspeth, S. G., Holland, A. L., Bonilha, L., Fromm, D., & Rorden, C. (2012). Speech entrainment enables patients with Broca's aphasia to produce fluent speech. *Brain, 135*(12), 3815–3829.

Furnas, D. W., & Edmonds, L. A. (2014). The effect of computerised Verb Network Strengthening Treatment on lexical retrieval in aphasia. *Aphasiology, 28*(4), 401–420.

Georgeadis, A. C., Brennan, D. M., Barker, L. M., & Baron, C. R. (2004). Telerehabilitation and its effect on story retelling by adults with neurogenic communication disorders. *Aphasiology, 18*(5–7), 639–652.

Glosser, G., Wiener, M., & Kaplan, E. (1988). Variations in aphasic language behaviors. *Journal of Speech & Hearing Disorders, 53*(2), 115–124.

Goldberg, S., Haley, K. L., & Jacks, A. (2012). Script training and generalization for people with aphasia. *American Journal of Speech-Language Pathology, 21*(3), 222–238.

Gordon, J. K. (2006). A quantitative production analysis of picture description. *Aphasiology, 20*(2–4), 188–204.

Gordon, J. K. (2007). A contextual approach to facilitating word retrieval in agrammatic aphasia. *Aphasiology, 21*(6–8), 643–657.

Gordon, J. K. (2008). Measuring the lexical semantics of picture description in aphasia. *Aphasiology, 22*(7–8), 839–852.

Greenwood, A., Grassly, J., Hickin, J., & Best, W. (2010). Phonological and orthographic cueing therapy: a case of generalised improvement. *Aphasiology, 24*(9), 991–1016.

Helmick, J. W., & Berg, C. B. (1976). Perseveration in brain-injured adults. *Journal of Communication Disorders, 9*(2), 143–156.

Herbert, R., Hickin, J., Howard, D., Osborne, F., & Best, W. (2008). Do picture-naming tests provide a valid assessment of lexical retrieval in conversation in aphasia? *Aphasiology, 22*(2), 184–203.

Herbert, R., Webster, D., & Dyson, L. (2012). Effects of syntactic cueing therapy on picture naming and connected speech in acquired aphasia. *Neuropsychological Rehabilitation, 22*(4), 609–633.

Hickin, J., Mehta, B., & Dipper, L. (2015). To the sentence and beyond: a single case therapy report for mild aphasia. *Aphasiology, 29*(9), 1038–1061.

Hoover, E. L., Caplan, D., Waters, G., & Budson, A. (2015). Effects of impairment-based individual and socially oriented group therapies on verb production in aphasia. *Aphasiology, 29*(7), 781–798.

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*(5), 523–528.

Jacobs, B. J. (2001). Social validity of changes in informativeness and efficiency of aphasic discourse following linguistic specific treatment (LST). *Brain & Language, 78*(1), 115–127.

Jacobs, B. J., & Thompson, C. K. (2000). Cross-modal generalization effects of training noncanonical sentence comprehension and production in agrammatic aphasia. *Journal of Speech, Language & Hearing Research, 43*(1), 5–20.

Kambanaros, M. (2010). Action and object naming versus verb and noun retrieval in connected speech: comparisons in late bilingual greek–english anomic speakers. *Aphasiology, 24*(2), 210–230.

Kempler, D., & Goral, M. (2011). A comparison of drill- and communication-based treatment for aphasia. *Aphasiology, 25*(11), 1327–1346.

Kendall, D. L., Rosenbek, J. C., Heilman, K. M., Conway, T., Klenberg, K., Gonzalez Rothi, L. J., & Nadeau, S. E. (2008). Phoneme-based rehabilitation of anomia in aphasia. *Brain & Language, 105*(1), 1–17.

Kendall, D. P., Raymer, A. P., Rose, M. P., Gilbert, J. M. S., & Rothi, L. J. G. P. (2014). Anomia treatment platform as behavioral engine for use in research on physiological adjuvants to neurorehabilitation. *Journal of Rehabilitation Research and Development, 51*(3), 391–400.

Kennedy, M., & Murdoch, B. E. (1991). Patterns of speech and language recovery following left striato-capsular haemorrhage. *Aphasiology, 5*(6), 489–510.

Kim, M., Adingono, M. F., & Revoir, J. S. (2007). Argument structure enhanced verb naming treatment: two case studies. *Contemporary Issues in Communication Science & Disorders, 34*, 24–36.

King, J. M., & Hux, K. (1995). Intervention using talking word processing software: an aphasia case study. *AAC: Augmentative & Alternative Communication, 11*(3), 187–192.

Kintz, S., Wright, H. H., & Fergadiotis, G. (2015). Semantic knowledge use in discourse produced by individuals with anomic aphasia. *Aphasiology*, 1–14. doi:10.1080/02687038.2015.1081140.

Knoph, M. I. N., Lind, M., & Simonsen, H. G. (2015). Semantic feature analysis targeting verbs in a quadrilingual speaker with aphasia. *Aphasiology*, 1–24. doi:10.1080/02687038.2015.1049583.

LaFrance, C., Garcia, L. J., & Labreche, J. (2007). The effect of a therapy dog on the communication skills of an adult with aphasia. *Journal of Communication Disorders, 40*(3), 215–224.

Lee, J. B., Kaye, R. C., & Cherney, L. R. (2009). Conversational script performance in adults with non-fluent aphasia: treatment intensity and aphasia severity. *Aphasiology, 23*(7–8), 885–897.

Linebarger, M., McCall, D., Virata, T., & Berndt, R. S. (2007). Widening the temporal window: processing support in the treatment of aphasic language production. *Brain & Language, 100*(1), 53–68.

Llinàs-Grau, M., & Martínez-Ferreiro, S. (2013). On the presence and absence of that in aphasia. *Aphasiology, 28*(1), 62–81.

Lock, S., & Armstrong, L. (1997). Cohesion analysis of the expository discourse of normal, fluent aphasic and demented adults: a role in differential diagnosis? *Clinical Linguistics & Phonetics, 11*(4), 299–317.

Lustig, A. P., & Tompkins, C. A. (2002). A written communication strategy for a speaker with aphasia and apraxia of speech: treatment outcomes and social validity. *Aphasiology, 16*(4–6), 507–521.

MacWhinney, B., Fromm, D., Forbes, M., & Holland, A. (2011). AphasiaBank: methods for studying discourse. *Aphasiology, 25*(11), 1286–1307.

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

Maher, L. M., Kendall, D., Swearengin, J. A., Rodriguez, A., Leon, S. A., Pingel, K.,. . ., Rothi, L. J. G. (2006). A pilot study of use-dependent learning in the context of Constraint Induced Language Therapy. *Journal of the International Neuropsychological Society, 12*(6), 843–852.

Marshall, R. S., Laures-Gore, J., DuBay, M., Williams, T., & Bryant, D. (2015). Unilateral forced nostril breathing and aphasia: exploring unilateral forced nostril breathing as an adjunct to aphasia treatment: a case series. *Journal of Alternative and Complementary Medicine, 21*(2), 91–99.

Mason, C., Nickels, L., McDonald, B., Moses, M., Makin, K., & Taylor, C. (2011). Treatment of word retrieval impairments in aphasia: evaluation of a self-administered home programme using personally chosen words. *Aphasiology, 25*(2), 245–268.

Maul, K. K., Conner, P. S., Kempler, D., Radvanski, C., & Goral, M. (2014). Using informative verbal exchanges to promote verb retrieval in nonfluent aphasia. *American Journal of Speech-Language Pathology, 23*(4), 407–420.

Mayer, J., & Murray, L. (2003). Functional measures of naming in aphasia: word retrieval in confrontation naming versus connected speech. *Aphasiology, 17*(5), 481–497.

McCall, D., Virata, T., Linebarger, M. C., & Berndt, R. S. (2009). Integrating technology and targeted treatment to improve narrative production in aphasia: a case study. *Aphasiology, 23*(4), 438–461.

McCarney, C. T., & Johnson, A. F. (2001). Examining conversational characteristics of persons with chronic aphasia during three group conversations. *Aphasiology, 15*(10–11), 1017–1028.

McEntee, L. J., & Kennedy, M. (1995). Profiling agrammatic spoken language: towards a government and binding framework. *European Journal of Disorders of Communication, 30*(3), 317–332.

McNeil, M. R., Doyle, P. J., Fossett, T. R. D., Park, G. H., & Goda, A. J. (2001). Reliability and concurrent validity of the information unit scoring metric for the story retelling procedure. *Aphasiology, 15*(10–11), 991–1006.

McNeil, M. R., Doyle, P. J., Park, G. H., Fossett, T. R. D., & Brodsky, M. B. (2002). Increasing the sensitivity of the Story Retell Procedure for the discrimination of normal elderly subjects from persons with aphasia. *Aphasiology, 16*(8), 815–822.

McNeil, M. R., Doyle, P. J., Spencer, K. A., Goda, A. J., Flores, D., & Small, S. L. (1997). A double-blind, placebo-controlled study of pharmacological and behavioural treatment of lexical-semantic deficits in aphasia. *Aphasiology, 11*(4–5), 385–400.

McNeil, M. R., Sung, J. E., Yang, D., Pratt, S. R., Fossett, T. R. D., & Doyle, P. J. (2007). Comparing connected language elicitation procedures in persons with aphasia: concurrent validation of the story retell procedure. *Aphasiology, 21*(6–8), 775–790.

Meuse, S., & Marquardt, T. P. (1985). Communicative effectiveness in Broca's aphasia. *Journal of Communication Disorders, 18*(1), 21–34.

Milman, L., Clendenen, D., & Vega-Mendoza, M. (2014). Production and integrated training of adjectives in three individuals with nonfluent aphasia. *Aphasiology, 28*(10), 1198–1222.

Milman, L., Vega-Mendoza, M., & Clendenen, D. (2014). Integrated training for aphasia: an application of part-whole learning to treat lexical retrieval, sentence production, and discourse-level communications in three cases of nonfluent aphasia. *American Journal of Speech-Language Pathology, 23*(2), 105–119.

Milman, L. H., Dickey, M. W., & Thompson, C. K. (2008). A psychometric analysis of functional category production in English agrammatic narratives. *Brain & Language, 105*(1), 18–31.

Mortensen, L. (2005). Written discourse and acquired brain impairment: evaluation of structural and semantic features of personal letters from a systemic functional linguistic perspective. *Clinical Linguistics & Phonetics, 19*(3), 227–247.

Muñoz, M. L., & Marquardt, T. P. (2004). The influence of language context on lexical retrieval in the discourse of bilingual speakers with aphasia. *Journal of Multilingual Communication Disorders, 2*(1), 1–17.

Murray, L., Ballard, K., & Karcher, L. (2004). Linguistic specific treatment: just for Broca's aphasia? *Aphasiology, 18*(9), 785–809.

Murray, L., Timberlake, A., & Eberle, R. (2007). Treatment of underlying forms in a discourse context. *Aphasiology, 21*(2), 139–163.

Murray, L. L., & Holland, A. L. (1995). The language recovery of acutely aphasic patients receiving different therapy regimens. *Aphasiology, 9*(4), 397–405.

Murray, L. L., Holland, A. L., & Beeson, P. M. (1998). Spoken language of individuals with mild fluent aphasia under focused and divided-attention conditions. *Journal of Speech Language & Hearing Research, 41*(1), 213–227.

Murray, L. L., & Karcher, L. (2000). A treatment for written verb retrieval and sentence construction skills. *Aphasiology, 14*(5–6), 585–602.

Nicholas, M., Sinotte, M., & Helm-Estabrooks, N. (2005). Using a computer to communicate: effect of executive function impairments in people with severe aphasia. *Aphasiology, 19*(10/11), 1052–1065.

Oelschlaeger, M. L., & Thorne, J. C. (1999). Application of the correct information unit analysis to the naturally occurring conversation of a person with aphasia. *Journal of Speech Language and Hearing Research, 42*(3), 636–648.

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

Olness, G. S., Ulatowska, H. K., Wertz, R. T., Thompson, J. L., & Auther, L. L. (2002). Discourse elicitation with pictorial stimuli in African Americans and Caucasians with and without aphasia. *Aphasiology, 16*(4–6), 623–633.

Pashek, G. V., & Tompkins, C. A. (2002). Context and word class influences on lexical retrieval in aphasia. *Aphasiology, 16*(3), 261–286.

Patterson, K., Graham, N. L., Ralph, M. A. L., & Hodges, J. R. (2006). Progressive non-fluent aphasia is not a progressive form of non-fluent (post-stroke) aphasia. *Aphasiology, 20*(9–11), 1018–1034.

Peach, R. K., & Reuter, K. A. (2010). A discourse-based approach to semantic feature analysis for the treatment of aphasic word retrieval failures. *Aphasiology, 24*(9), 971–990.

Penn, C. (1987). Compensation and language recovery in the chronic aphasic patient. *Aphasiology, 1*(3), 235–245.

Pickersgill, M. J., & Lincoln, N. B. (1983). Prognostic indicators and the pattern of recovery of communication in aphasic stroke patients. *Journal of Neurology, Neurosurgery & Psychiatry, 46* (2), 130–139.

Pritchard, M., Dipper, L., Morgan, G., & Cocks, N. (2015). Language and iconic gesture use in procedural discourse by speakers with aphasia. *Aphasiology, 29*(7), 826–844.

Purdy, M. H. (2002). Script knowledge following stroke. *Journal of Medical Speech-Language Pathology, 10*(3), 173–181.

Raymer, A. M., Ciampitti, M., Holliway, B., Singletary, F., Blonder, L. X., Ketterson, T.,..., Rothi, L. J. G. (2007). Semantic-phonologic treatment for noun and verb retrieval impairments in aphasia. *Neuropsychological Rehabilitation, 17*(2), 244–270.

Raymer, A. M., Rowland, L., Haley, M., & Crosson, B. (2002). Nonsymbolic movement training to improve sentence generation in transcortical motor aphasia: a case study. *Aphasiology, 16*(4–6), 493–506.

Rider, J. D., Wright, H. H., Marshall, R. C., & Page, J. L. (2008). Using semantic feature analysis to improve contextual discourse in adults with aphasia. *American Journal of Speech-Language Pathology, 17*(2), 161–172.

Rochon, E., Laird, L., Bose, A., & Scofield, J. (2005). Mapping therapy for sentence production impairments in nonfluent aphasia. *Neuropsychological Rehabilitation, 15*(1), 1–36.

Rochon, E., Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (2000). Quantitative analysis of aphasic sentence production: further development and new data. *Brain & Language, 72*(3), 193–218.

Rodriguez, A. D., Worrall, L., Brown, K., Grohn, B., McKinnon, E., Pearson, C.,..., Copland, D. A. (2013). Aphasia LIFT: exploratory investigation of an intensive comprehensive aphasia programme. *Aphasiology, 27*(11), 1339–1361.

Rogalski, Y., Edmonds, L. A., Daly, V. R., & Gardner, M. J. (2013). Attentive Reading and Constrained Summarisation (ARCS) discourse treatment for chronic Wernicke's aphasia. *Aphasiology, 27*(10), 1232–1251.

Rose, M. L., Attard, M. C., Mok, Z., Lanyon, L. E., & Foster, A. M. (2013). Multi-modality aphasia therapy is as efficacious as a constraint-induced aphasia therapy for chronic aphasia: a phase 1 study. *Aphasiology, 27*(8), 938–971.

Ross, K. B. (1999). Comparison of impairment and disability measures for assessing severity of, and improvement in, aphasia. *Aphasiology, 13*(2), 113–124.

Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (1989). The quantitative analysis of agrammatic production: procedure and data. *Brain & Language, 37*(3), 440–479.

Shewan, C. M. (1988). Expressive language recovery in aphasia using the Shewan Spontaneous Language Analysis (SSLA) system. *Journal of Communication Disorders, 21*(2), 155–169.

Silkes, J. P., Dierkes, K. E., & Kendall, D. L. (2013). Masked repetition priming effects on naming in aphasia: a phase I treatment study. *Aphasiology, 27*(4), 381–397.

Singh, S., & Bookless, T. (1997). Analysing spontaneous speech in dysphasic adults. *International Journal of Applied Linguistics, 7*(2), 165–182.

Thompson, C. K., Shapiro, L. P., Ballard, K. J., Jacobs, B. J., Schneider, S. S., & Tait, M. E. (1997). Training and generalized production of wh- and NP-movement structures in agrammatic aphasia. *Journal of Speech, Language, and Hearing Research, 40*(2), 228–244.

Thompson, C. K., Shapiro, L. P., Swathi, K., & Sobecks, J. (2003). The role of syntactic complexity in treatment of sentence deficits in agrammatic aphasia: the complexity account of treatment efficacy (CATE). *Journal of Speech, Language, and Hearing Research, 46*(3), 591–607.

Thompson, C. K., Shapiro, L. P., Tait, M. E., Jacobs, B. J., & Schneider, S. L. (1996). Training wh-question production in agrammatic aphasia: analysis of argument and adjunct movement. *Brain & Language, 52*(1), 175–228.

Ulatowska, H. K., Allard, L., Reyes, B. A., Ford, J., & Chapman, S. (1992). Conversational discourse in aphasia. *Aphasiology, 6*(3), 325–330.

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

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

Ulatowska, H. K., Olness, G. S., Wertz, R. T., Samson, A. M., Keebler, M. W., & Goins, K. E. (2003). Relationship between discourse and Western Aphasia Battery performance in African Americans with aphasia. *Aphasiology, 17*(5), 511–521.

Ulatowska, H. K., Olness, G. S., Wertz, R. T., Thompson, J. L., Keebler, M. W., Hill, C. L., & Auther, L. L. (2001). Comparison of language impairment, functional communication, and discourse measures in African-American aphasic and normal adults. *Aphasiology, 15*(10–11), 1007–1016.

Ulatowska, H. K., Reyes, B., Santos, T. O., Garst, D., Vernon, J., & McArthur, J. (2013). Personal narratives in aphasia: understanding narrative competence. *Topics in Stroke Rehabilitation, 20*(1), 36–43.

Wallace, S. E., & Kimelman, M. D. Z. (2013). Generalization of word retrieval following semantic feature treatment. *NeuroRehabilitation, 32*(4), 899–913.

Wambaugh, J. L., & Ferguson, M. (2007). Application of semantic feature analysis to retrieval of action names in aphasia. *Journal of Rehabilitation Research and Development, 44*(3), 381–394.

Wambaugh, J. L., & Martinez, A. L. (2000). Effects of modified response elaboration training with apraxic and aphasic speakers. *Aphasiology, 14*(5–6), 603–617.

Wambaugh, J. L., Martinez, A. L., & Alegre, M. N. (2001). Qualitative changes following application of modified response elaboration training with apraxic-aphasic speakers. *Aphasiology, 15*(10–11), 965–976.

Wambaugh, J. L., Mauszycki, S., & Wright, S. (2014). Semantic feature analysis: application to confrontation naming of actions in aphasia. *Aphasiology, 28*(1), 1–24.

Wambaugh, J. L., Nessler, C., & Wright, S. (2013). Modified response elaboration training: application to procedural discourse and personal recounts. *American Journal of Speech-Language Pathology, 22*(2), S409–425.

Wambaugh, J. L., Wright, S., & Nessler, C. (2012). Modified response elaboration training: a systematic extension with replications. *Aphasiology, 26*(12), 1407–1439.

Wambaugh, J. L., Wright, S., Nessler, C., & Mauszycki, S. C. (2014). Combined Aphasia and Apraxia of Speech Treatment (CAAST): effects of a novel therapy. *Journal of Speech, Language & Hearing Research, 57*(6), 2191–2207.

Webster, J., Franklin, S., & Howard, D. (2001). An investigation of the interaction between thematic and phrasal structure in nonfluent agrammatic subjects. *Brain & Language, 78*(2), 197–211.

Weinrich, M., McCall, D., Boser, K. I., & Virata, T. (2002). Narrative and procedural discourse production by severely aphasic patients. *Neurorehabilitation and Neural Repair, 16*(3), 249–274.

Winans-Mitrik, R. L., Hula, W. D., Dickey, M. W., Schumacher, J. G., Swoyer, B., & Doyle, P. J. (2014). Description of an intensive residential aphasia treatment program: rationale, clinical processes, and outcomes. *American Journal of Speech-Language Pathology, 23*(2), S330–342.

Wolf, R., Sidtis, D. V. L., & Sidtis, J. J. (2014). The ear craves the familiar: pragmatic repetition in left and right cerebral damage. *Aphasiology, 28*(5), 596–615.

Wright, H. H., & Capilouto, G. J. (2012). Considering a multi-level approach to understanding maintenance of global coherence in adults with aphasia. *Aphasiology, 26*(5), 656–672.

Wright, H. H., Silverman, S. W., & Newhoff, M. (2003). Measures of lexical diversity in aphasia. *Aphasiology, 17*(5), 443–452.

Yorkston, K. M., & Beukelman, D. R. (1980). An analysis of connected speech samples of aphasic and normal speakers. *Journal of Speech & Hearing Disorders, 45*(1), 27–36.

Youmans, G., Holland, A., Muñoz, M., & Bourgeois, M. (2005). Script training and automaticity in two individuals with aphasia. *Aphasiology, 19*(3–5), 435–450.

Appendix B. Coding categories for information extracted from included studies

Sample elicitation methods

Exposition: an opinion or discussion describing or explaining an issue or topic
> **Picture description** – a description of the content of a single picture
> **Picture sequence description** – a description of the content of multiple pictures forming a sequence. This may also be the description of a video sequence with no linguistic content
> **Opinion** – a personal opinion or perspective on an event, concept or idea (e.g. a current affair item, a popular book or movie)

Procedure: description of a process (e.g. how to complete an activity)

Narrative: an account of a fictional or factual story with a background or setting, series of events and a conclusion
> **Single-picture story** – formulation of story with some temporal-causal sequence based on the content of a single picture
> **Picture sequence story** – formulation of story with some temporal-causal sequence based on the content of multiple pictures forming a sequence
> **Narrative retell** – relaying a story following a verbal structure/model provided by an examiner. This may also be picture-supported
> **Fictional story telling** – relaying a familiar fairy tale (e.g. Cinderella) after revision of a wordless picture book, with pictures removed prior to narrative production
> **Personal narrative recount** – recounting an event that has happened to the individual in the past (e.g. story of stroke)

Conversation: an interactive communication between two or more people
> **Role play** – acting out a monologue or dialogue that may be used in another context (e.g. conversational scripts)
> **Structured conversation** – conversation guided by specific questions of examiner to control content (e.g. an interview)
> **Unstructured conversation** – general conversation unrestricted by topic

Linguistic discourse analysis measures

(NOTE: analyses are applied both above and below the sentence/utterance level)
   (For a comprehensive list of specific measures within each category, please contact the authors)

Verbal productivity:
> **Sample length** – quantification of the amount of language units within a sample (e.g. number of words, clauses, sentences, etc.)
> **Lexical diversity** – the variety of vocabulary used in a sample (e.g. type-token ratio (TTR), number of different words (NDW), vocabulary diversity (D), etc.)
> **Speech fluency** – the rate of speech production in units per minute, or measurement of the length or number of dysfluent moment (e.g. words per minute (WPM), proportion of dysfluencies per word, etc.)

Word finding behaviours – linguistic structures or behaviours that indicate difficulty retrieving words (e.g. paraphasias, delays, attempts at word production, etc.)

Information content:

Efficiency - the rate of information production per minute (e.g. Correct Information Units per minute (CIUs/min), etc.)

Lexical – analysis of single words that are integral to the expression of the information content of the sample as a whole (e.g. correct information units (CIUs), propositional density, number of content words, etc.)

Semantic/conceptual – the elements that communicate the gist, theme or main ideas of the language sample (e.g. information units (IUs), main events, story propositions, etc.)

Schema related – analysis of the over-arching schematic structure or framework, specific to the elicited language genre, onto which information content is mapped (e.g. number of temporal-causal sequences, utterances conveying background or setting, number of procedural steps communicated, etc.)

Cohesion – the linguistic structures that tie the smaller lexical and grammatical elements of a language sample together to form a whole (e.g. number of cohesive ties, number of referential errors, use of conjunctions, etc.)

Grammatical complexity:

Morphological - the grammatical structure of language measured in the use of bound and unbound morphemes (e.g. number of bound morphemes, use of tense or plurals, etc.)

Word classes – the grammatical categories of words used within a language sample (e.g. numbers of nouns, verbs, adjectives, closed-class words, etc.)

Syntactic – the structure of language at the sentential level (e.g. proportion of sentences that are grammatically complete, number of words in sentences, etc.)