Discourse production in aphasia: a current review of theoretical and methodological challenges

Anastasia Linnik, Roelien Bastiaanse & Barbara Höhle

To cite this article: Anastasia Linnik, Roelien Bastiaanse & Barbara Höhle (2015): Discourse production in aphasia: a current review of theoretical and methodological challenges, Aphasiology, DOI: 10.1080/02687038.2015.1113489

To link to this article: http://dx.doi.org/10.1080/02687038.2015.1113489

Published online: 02 Dec 2015.

Submit your article to this journal

Article views: 94

View related articles

View Crossmark data
Discourse production in aphasia: a current review of theoretical and methodological challenges

Anastasia Linnika, Roelien Bastiaanseb and Barbara Höhléc

aInternational Doctorate for Experimental Approaches to Language and Brain (IDEALAB), Universities of Groningen (NL), Newcastle (UK), Potsdam (GE), Trento (IT) & Macquarie University, Sydney (AU); bCenter for Language and Cognition Groningen (CLCG), University of Groningen, Groningen, Netherlands; cDepartment of Linguistics, University of Potsdam, Potsdam, Germany

(Received 12 June 2015; accepted 25 October 2015)

Background: Discourse abilities play an important role in the assessment, classification, and therapy outcome evaluation of people with aphasia. Discourse production in aphasia has been studied quite extensively in the last 15 years. Nevertheless, many questions still do not have definitive answers.

Aims: The aim of this review is to present the current situation in the research on a number of crucial aspects of discourse production in aphasia, focusing on methodological progress and related challenges. This review continues the discussion of the core themes in the field, aiming to render it as up-to-date as possible.

Main Contribution: The review focuses on a number of unexplored theoretical issues, specifically, the interface between micro- and macrolinguistic abilities, and the relationship between linguistic competence and communicative success in aphasia. The emphasis on theoretical challenges, along with the thorough discussion of methodological problems in the field, makes this review a starting point and a comprehensive information source for researchers planning to address language production in people with aphasia.

Conclusion: Although the picture is not yet complete, recent advancements lead to a better understanding of the processes involved in aphasic discourse production. Different approaches provide insights into the complex multifaceted nature of discourse-level phenomena; however, methodological issues, including low comparability, substantially slow down the progress in the field.

Keywords: Discourse production; aphasia

Motivation for the study of discourse in aphasia

Discourse is indispensable for human interactions, as well as for the expression of one’s feelings, thoughts, or ideas. Telling personal stories, engaging in long conversations, giving talks, and creating other forms of spoken and written discourse is essential for communication and establishing relations within a society. Due to their language impairments, people with aphasia (PWA) often have problems with social and professional integration, and, as a consequence, lower quality of life. The recovery of their language skills is supported by speech-language therapy, which traditionally focuses on smaller language components, such as single words and sentences. Discourse, however, is internally more complex than a group of words or sentences put together. The mechanisms...
underlying the organisation of speech into a coherent flow have not been fully understood yet, despite the fact that discourse production in healthy population has been extensively studied, compared with that in language-impaired. Investigation of the patterns of preservations and impairments related to different aspects of discourse, however, may provide insights both for clinical practice and for cognitive science, as it grants a unique opportunity to access the underlying linguistic and cognitive processes that are relevant for discourse production, and to devise a more targeted and effective approach to treatment.

In her comprehensive review of aphasic discourse studies, Armstrong (2000) expressed concern about the lack of a unified theoretical base to study discourse in aphasia. She emphasised that a large variety of existing methodologies and differences in the definitions of crucial concepts have yielded disparate findings. Armstrong (2000) addressed syntactic abilities during discourse production, discourse organisation issues along with coherence and cohesion, and a number of methodological problems. Fifteen years have passed since her review was published, and more effort has been devoted to this topic, bringing new results and addressing some of the questions raised, but also creating room for more discrepancies and contradictory conclusions.

The current review addresses a number of recurring theoretical topics in the field of discourse production in aphasia. Building on the work of Armstrong (2000), it offers an update on the findings and the discussion of practical issues in research design, their impact on data interpretation, and potential ways to overcome some of the related methodological problems. When thinking about discourse analysis, one needs to have an understanding of which components of language production this term comprises. We will consider several crucial concepts, namely informativeness, information structure, discourse structure, cohesion and coherence, existing multilevel approaches addressing the connection between them, and overall communicative effectiveness and efficiency. The first part of this review incorporates the literature published since the work of Armstrong (2000) into the discussion of relevant themes and issues, and highlights those theoretical and methodological aspects of discourse analysis which may warrant increased attention in the future, to optimally advance the field of clinical discourse studies. The second part of the review addresses several global methodological issues in discourse studies, including sample size, genre, modality, and crosslinguistic research. The discussion of these particular methodological matters was included in this review, because the quality, interpretability, and future comparability of the outcomes of every study on discourse in aphasia depend to a certain extent on the careful consideration of each of these factors.

Very few other works have attempted to bring together the findings on discourse in aphasia. Prins and Bastiaanse (2004) reviewed the body of work on spontaneous speech of adults with aphasia, including several existing pragmatic analysis tools, the application of Conversation Analysis (CA; Sacks, 1972; Sacks, Schegloff, & Jefferson, 1974) to aphasic data, as well as a group of standardised quantitative and qualitative linguistic measures. They stressed the lack of attention to the topic, and suggested using a combination of functional and statistical perspectives in future research. Recently a special issue of Aphasiology (Wright, 2011a) was dedicated to various topics in aphasic discourse research, namely, discourse level treatment techniques (Boyle, 2011; Kempler & Goral, 2011), new methodological developments (Fergadiotis & Wright, 2011; Marini, Andreotta, Del Tin, & Carломagno, 2011; Olness & Ulatowska, 2011), comparison of monologue and conversational discourse (Armstrong, Ciccone, Godecke, & Kok, 2011), coherence in personal narratives (Olness & Ulatowska, 2011), and aphasic speakers’ evaluation of their own verbal language abilities (Fromm et al., 2011).
Whereas the special issue of Aphasiology highlighted several directions of research on discourse in aphasia, the current paper provides a comprehensive overview of central theoretical constructs and methodological challenges in the field, with purposeful incorporation of the most recent publications, to guide the field towards potential refinements of study design and methodology.

Several concepts defining discourse production in aphasia

The existing body of work on discourse production in aphasia rendered a complex picture, according to which the overall ability to produce discourse consists of several aspects. Both early and current data demonstrated that some of these aspects remained relatively intact, while others were more or less impaired. Already in their early fundamental studies, Ulatowska, Allard, and Chapman (1990), Ulatowska, Freedman-Stern, Doyel, Macaluso-Haynes, and North (1983) and Ulatowska, North, and Macaluso-Haynes (1981) argued that although aphasic narratives are shorter and grammatically simpler, they contain all the essential elements of story structure and the chronological order of events. A number of other researchers agreed that overall text macrostructure, global coherence, and other pragmatic skills of aphasic speakers are preserved (e.g., Armstrong & Ulatowska, 2007; Glosser & Deser, 1991; Gordon, 2006; Ulatowska et al., 2003). On the other hand, an increasing number of studies have provided evidence that people with aphasia do experience various difficulties communicating at the discourse level (e.g., Armstrong et al., 2011; Fergatioidis & Wright, 2011; Wright, 2011b), such as production of excessive irrelevant proposition content, reduced efficiency, and low lexical informativeness (Andreetta, 2014; Andreetta & Marini, 2015; Christiansen, 1995; Nicholas & Brookshire, 1993).

The multifactorial nature of discourse production motivated a line of studies, which focused on the connectivity between aspects of local and global levels, both impaired and preserved. Holland (1982) used the notion of “functional communication”, the ability to communicate despite the language production difficulties, and claimed it to be available to PWA despite the microstructural linguistic difficulties. However, several researchers have noted that PWA experience difficulties with microstructuring, such as construction of cohesive ties (Armstrong, 2000; Bloom, 1994; Olness & Ulatowska, 2011), which impact the global coherence of aphasic discourse, making it vague and potentially ambiguous (Christiansen, 1995; Huber, 1990).

Some of the previous studies addressed the same concepts and produced seemingly contradictory outcomes. In the first part of this review, we will focus on determining the source of these contradictory findings by going step by step through several concepts defining language production. It will soon become clear to the reader that the discussion reveals not just the contradictory findings, but rather the inconsistencies with respect to the definitions and/or methodologies used to investigate the same constructs. This part of the review also focuses on the underexplored interrelationship between micro- and macrolinguistic levels, and highlights some less studied phenomena at the interface between different discourse analysis traditions (i.e., functional, structural, and cognitivist).

Informativeness

Discourse is a flow of information put into words, organised in order to meet specific communication goals, and shaped by situational factors (common ground, social context,
etc.). Hence, the first problem to be addressed is the information content, or the informativeness, of discourse in aphasia. A reduced amount of essential content, information gaps, tangential propositions, and topic shifts were found to contribute to the vagueness and incomprenhensibility of aphasic discourse (e.g., Andreetta, Cantagallo, & Marini, 2012; Capilouto, Wright, & Wagovich, 2006; Foka-Kavalieraki et al., 2008; Stark, 2010; Ulatowska & Chapman, 1994; Ulatowska et al., 1983).

Research findings regarding the informativeness of discourse produced by people with aphasia must be interpreted in light of the various methods used to assess informativeness. Several measures have been developed to assess informativeness in aphasia (Table 1). For a long time the only existing method was the analysis of content units (CU) developed by Yorkston and Beukelman (1980). A CU was defined as “a grouping of information always expressed as a unit by normal speakers” (p. 30), and consisted of a single word, a noun phrase, verb phrase, or a propositional phrase. A total count of CUs was identified for a task, such as picture-elicited storytelling, which included all CUs mentioned by at least one of the participants of the study. Nicholas and Brookshire (1993) devised a rule-based measure, which was not content specific, based on scoring correct information units (CIU), single words which are accurate, informative, and relevant to the story being told. Both studies reported lower informativeness, measured in CUs and CIUs, in aphasia as compared to healthy speakers’ discourse. Later several techniques were proposed to compete with those two widely used variables. For example, McNeil, Doyle, Fossett, Park, and Goda (2001) presented a simplified and arguably a more efficient scoring procedure, namely percent of information units (%IU), which they reported to be highly reliable. IUs, similar to CIUs, are intelligible and informative words or phrases that convey accurate and relevant information about the story. Ulatowska et al. (2003) suggested rating “emplotment”, or “the ability to express information in the narrative structural form” (p. 515), complemented by a quantitative measure, the number of propositions. Interestingly, they found employment, but not the number of propositions, to be correlated with the Western Aphasia Battery Aphasia Quotient scores. Wright and colleagues (Capilouto et al., 2006; Wright, Capilouto, Wagovich, Cranfill, & Davis, 2005) developed a main event analysis, which demonstrated that adults with aphasia conveyed a lower proportion of main events in picture-elicited narratives. Similarly, Marini and colleagues measured informativeness in thematic units, which they defined as “a main idea or detail in the story” (Andreetta et al., 2012; Marini et al., 2011: p. 1383). First, a large group of non-impaired speakers was asked to identify the thematic “backbone” of a story, after that the rate of thematic units was calculated. In addition, the count of lexical information units (LIU), grammatically and pragmatically appropriate content and function words, was used to account for lexical-semantic appropriateness. The results demonstrated that thematic informativeness was within normal limits in aphasic speech, whereas the number of lexical information units was reduced, indicating a certain number of tangential and uninformative words (Andreetta, 2014; Andreetta & Marini, 2015; Andreetta et al., 2012). The information on the design and results of the above mentioned studies, along with several other works addressing informativeness in aphasic discourse, is summarised in Table 1.

There are a few challenges with the informativeness measurements that have been used in previous studies. First, our understanding of informativeness in discourse may be constrained by the limitations of the current methods. Namely, for most of the mentioned lexical informativeness measures a certain number of raters have to be trained, and even then, for example, CIU analysis has demonstrated a low reliability when applied to
Table 1. Studies addressing informativeness in aphasic discourse: Methods and findings.

<table>
<thead>
<tr>
<th>Investigator(s)</th>
<th>Case/Multiple case/Group</th>
<th>Elicitation task</th>
<th>Units</th>
<th>Impaired?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Yorkston and Beukelman (1980)</td>
<td>50 PWA, 78 NBD</td>
<td>Single picture description</td>
<td>Content units</td>
<td>No, in mild and high-moderate aphasia; Yes, in more severe aphasia</td>
</tr>
<tr>
<td>Ulatowska et al. (1990)</td>
<td>30 PWA (10 mild, 15 moderate, 5 severe)</td>
<td>Personal narrative, fable retelling, picture series description, procedural discourse</td>
<td>Number of propositions (events), number of steps in the procedures</td>
<td>No, in mild and moderate aphasia; Yes, in severe aphasia</td>
</tr>
<tr>
<td>Nicholas and Brookshire (1993)</td>
<td>20 PWA, 20 NBD</td>
<td>Single pictures, picture series, personal narratives, procedural discourse</td>
<td>Correct Information units (CIU)</td>
<td>Yes</td>
</tr>
<tr>
<td>Doyle, Goda, and Spencer (1995), Doyle et al. (2000)</td>
<td>20 PWA, 15 PWA</td>
<td>Conversation; Story retelling procedure (SPR)</td>
<td>CIU</td>
<td>No</td>
</tr>
<tr>
<td>McNeil et al. (2001)</td>
<td>15 PWA, 31 NBD</td>
<td>SPR (Doyle et al. 2000)</td>
<td>Percentage of Information units</td>
<td>Yes</td>
</tr>
<tr>
<td>Wright et al. (2005), Capilouto et al. (2006)</td>
<td>40 healthy adults; 8 PWA, 8 NBD</td>
<td>Single and sequential picture description</td>
<td>Main events</td>
<td>Yes</td>
</tr>
<tr>
<td>Marini, Carfomagno, Caltagirone, and Nocentini (2005), Marini et al. (2011), Andreetta et al. (2012), Andreetta (2014), Andreetta and Marini (2015)</td>
<td>11 participants with right brain damage (RBD) without aphasia, 11 RBD with aphasia, 11 NBD; 2 cases; 10 PWA, 10 NBD; 20 PWA, 20 NBD</td>
<td>Single and sequential picture description</td>
<td>Thematic units Lexical information units (%LIU)</td>
<td>No for thematic informativeness Yes for lexical information content</td>
</tr>
</tbody>
</table>

Note: PWA, people with aphasia; NBD, non-brain-damaged people; RBD, people with right hemisphere brain damage.
naturally occurring conversation (Oelschlaeger & Throne, 1999). For the thematic informativeness analysis, a group of raters is required to identify main thematic elements for the stimuli (e.g., the main events of a fable or a story illustrated with a series of pictures), which makes it inapplicable to some types of discourse, such as spontaneous conversation or personal recounts. For example, Doyle et al. (1995) showed that lexical information is higher for conversational discourse, but thematic information content cannot be estimated based on the principles identified for semispontaneous narratives.

Second, when addressing informativeness, one should distinguish between thematic and lexical content, and choose a method accordingly. Although thematic and/or lexical informativeness have been systematically included in recent studies on discourse in aphasia, very few of them address the dissociation between the two (e.g., Andreetta et al., 2012). The existing measures of thematic informativeness appear to be directly related to some of the coherence assessment methods discussed further in this review, and refer to the level of discourse organisation at the macrolevel, whereas lexical informativeness is a microlinguistic variable. However, lexical informativeness, but not thematic, has been found to be correlated with coherence (Andreetta et al., 2012). Moreover, it was reported to be a statistically significant predictor of discourse coherence (Wright & Capilouto, 2012). Thematic informativeness was not directly addressed in the latter study, but the coherence measure used in it relied on the amount of relevant information included in a discourse sample. Further investigations of the interaction between these variables should address the relationship between thematic and lexical informativeness, and attempt to clarify the role of information content in establishing coherence, and more generally, in language processing.

Information structure

The notion of “information structure” (IS), first introduced by Halliday (1967), usually refers to the way information is “packaged” in a sentence (Chafe, 1976). Not much is known about IS in aphasic discourse (Table 2), although IS has received considerable attention in healthy discourse analysis (Chafe, 1976; Chomsky, 1971; Gundel, Hegarty, & Borthen, 2003; Krifka, 2006; Lambrecht, 1994; Roberts, 1996; Dipper, Götze, Stede, & Wegst, 2004, i.a.).

The central concepts of information structure are topic, comment, focus, and givenness (see Krifka, 2008; Von Heusinger, 1999; for an overview and discussion). We will adopt the definitions of these concepts provided by Krifka (2008). “Topic” stands for the object which a speaker is talking about, and “comment” refers to what the speaker is stating about it. One distinguishes between sentence topic and discourse topic, referring to what a sentence or a piece of discourse “is about” respectively (e.g., Gundel & Freltheim, 2004; Lambrecht, 1994; Van Dijk, 1977). According to Krifka, “focus” signals the presence of alternatives relevant for the interpretation of a linguistic expression, whereas “givenness” indicates whether an expression is in the immediate common ground of interlocutors, and if so, to what extent. These constructs are assumed to be linked to the cognitive states of interlocutors, and help build mental representations or modify existing representations. One of the crucial observations of the information structure theory is that topic usually precedes focus/comment, or given information tends to appear earlier in a sentence than new information. In many languages, focus is also typically prosodically marked as more salient (Gundel & Freltheim, 2004).

In the 1980s, it was claimed that the topic-focus function and the given-new distinction remained intact in aphasia (Bates, Hamby, & Zurif, 1983; Wulfeck et al.,
Later insensitivity of adults with aphasia to the given-new organisation of simple narratives, as well as their failure to mark given or new information appropriately was noted (Cannito, Jarecki, & Pierce, 1986; Early & VanDemark, 1985; Ulatowska & Chapman, 1994). The results from Bastiaanse, Koekkoek, and Van Zonneveld (2003) were in line with this claim; however, some evidence suggested that individuals with Broca’s aphasia may be aware of the pragmatic rule allowing the omission of given information, although they may use it when syntactic rules do not allow such omissions. Olness et al. (2010) investigated pragmatic use of narrative evaluative devices, such as expressive lexicon, pitch peaks, or direct speech, in aphasia and claimed that their aphasic subjects were able to transmit the “point” of their personal narratives and assign prominence to information in a way similar to their nonlanguage-impaired counterparts. The preserved ability to use evaluative devices to communicate the distinction between more and less salient information in discourse is indicative of the PWA’s awareness of the concepts of psychological and semantic focus. Table 2 presents crucial information on the studies focusing on information structure in aphasic discourse production.

The inconsistent findings, do not allow us to draw any conclusions on whether or not problems PWA experience are rooted in the information structure. Investigating IS may...
shed light on the cognitive mechanisms related to speech production, such as attention and the ability to establish common ground. An impairment of these processes can be detrimental to discourse clarity. It has been suggested that PWAs’ ability to use information structure devices is limited due to syntactic deficits (Ulatowska, Allard, & Chapman, 1990). This hypothesis is yet to be systematically verified. Confirming or rejecting it would contribute to the understanding of a more general question of whether discourse-level impairments are a result of microlinguistic difficulties or not.

**Discourse structure**

Discourse is often rather simplistically defined as a unit above sentence. Apart from being a linguistic construct, it is in fact an action involving a number of cognitive processes, shaped by interactive and social factors (Fox, 1987; Van Dijk, 1997; cf. Cameron, 2001; Schiffrin, 1994; for the discussion on the definition of discourse). Thoughts, ideas, and information transferred through discourse are not chaotic, but organised, which lead to the idea of discourse being internally structured. Thus, the term “discourse structure” in this paper refers to the internal organisation of discourse into a coherent whole. Discourse is commonly considered to have two dimensions—local and global, also referred to as micro- and macrostructure, respectively (cf. Van Dijk, 1980). Although an extensive body of research exists on this subject in healthy population (e.g., Grosz & Sidner, 1986; Mann & Thompson, 1988; Moser & Moore, 1996; Redeker, 1990, 2000; Taboada, 2004; Wolf & Gibson, 2005), there is not one commonly accepted approach to studying discourse organisation at the macrolevel, and it follows that there is a corresponding possibility that no one common aspect of macro-organisation is being assessed by each given approach.

Van Dijk (1976, 1980) introduced the term “macrostructure of discourse” and defined it as a semantic object representing global meaning, also called “topic”, “theme”, or “gist”. Schematic organisation of global meaning through the use of narrative elements (e.g., setting, evaluation, and coda) he then called “superstructure”. These definitions of macro- and superstructure were used in several studies investigating discourse in aphasia (Ulatowska & Chapman, 1994; Ulatowska & Sadowska, 1992; Ulatowska et al., 1981, 1983, 1990). Their results showed that reduction in information content and poor distribution of information disrupts macrostructure even in simple narratives. Nonetheless, they argue that superstructure remains relatively well-preserved. Similarly, others claimed that people with aphasia displayed a remarkable ability for “maintaining conceptual and pragmatic organization at the supraglobal level” (Glosser & Deser, 1991: p. 68). Olness defined superstructure in the terms of Labov (1972) as setting, complicating action, and resolution, and reported it to be intact independently of aphasia severity (Olness, 2007; Olness & Ulatowska, 2011; Olness et al., 2010). However, a few studies addressing discourse organisation in terms of propositional content provided evidence against the preservation of this aspect of discourse organisation in aphasia (e.g., Christiansen, 1995; Huber, 1990). Another discourse production macrophenomenon is the ability of speakers to reduce a full-length discourse to a pithy encapsulation of its explicit and implicit content, expressed in a short phrase or two. Studies, in which aphasic participants were asked to produce a moral and a gist of a fable, demonstrated that PWA tend to experience difficulties at different levels of manipulation of discourse structure, in particular, with abstraction and generalisation (Ulatowska & Chapman, 1994;
A brief overview of the methods and results of the studies investigating different aspects of aphasic discourse organisation at the macrolevel discussed in this section are presented in Table 3.

Different theoretical approaches to the investigation of discourse structure used in the existing literature provided insights on various aspects of discourse macrostructure. Bringing these different perspectives together can potentially lead to a more in-depth, multifaceted understanding of language processing in aphasia at the macrolevel. Furthermore, different methodological approaches, which in turn assess different aspects of discourse structure, presumably make very different demands on the linguistic system of the speaker (Ulatowska & Chapman, 1994), potentially also providing access to various cognitive mechanisms, such as inferencing and cognitive planning, involved in production of discourse of different complexity. In her review, Armstrong (2000) noted the lack of information about “how text macrostructure is realized through words and sentences” (Armstrong, 2000: p. 876). It is unclear how the macrostructure of discourse is exactly built, and how it is represented linguistically. Further research should focus on establishing at which point the problems experienced by aphasic speakers at the lower levels of language organisation start affecting global connectedness and meaning of discourse, and which aspects of discourse structuring are involved in the process. So far very few studies have made contributions to this line of research (e.g., Ulatowska & Chapman, 1994; cf. the “Multi-level approaches” section further in this review).

**Coherence and cohesion**

**Cohesion**

Cohesion refers to the semantic connectedness between elements of discourse, reached via lexical and grammatical means, such as coreference, substitution, or conjunction (Halliday & Hasan, 1976). The term stands for a set of surface means used to achieve connectedness. Halliday and Hasan (1976) claimed that it “occurs where the interpretation of some elements of discourse depends on that of another” (p. 4). Several studies on cohesion in aphasia analysed semantic relations between elements (content words or their replacement) in a text, or “cohesive ties”, including pronominal reference (cf. Table 4). Referential function, which is a key component in establishing cohesion, is known to be particularly vulnerable in aphasia (Olness & Ulatowska, 2011; Ulatowska et al., 1990); PWA tend to omit antecedents of pronouns and create anaphoric ambiguity (cf., Andreetta et al., 2012; Armstrong, 2000; Bloom, 1994; Boyle, 2011; Glosser & Deser, 1991; Liles & Coelho, 1998; Marini et al., 2011; Ulatowska et al., 1981, 1983). Piehler and Holland (1984) investigated cohesion recovery in two individuals with aphasia, and noted that despite different recovery patterns, the two participants restored their ability to use lexical cohesion (e.g., synonyms). Armstrong et al. (2011) presented two case studies, where cohesion in monologues and dialogues between aphasic and NBD participants was explored. Their findings suggested that cohesion is facilitated for PWA in dialogues. Time reference requiring discourse linking, such as reference to the past, was also shown to be challenging for speakers with agrammatic aphasia (Bastiaanse, 2013; Bastiaanse et al., 2011; Bos, Dragoy, Avrutin, Iskra, & Bastiaanse, 2014). The approaches to study cohesion in aphasia are outlined in Table 4.
<table>
<thead>
<tr>
<th>Investigator(s)</th>
<th>Group/multiple case/case</th>
<th>Elicitation task</th>
<th>Method</th>
<th>Impaired?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ulatowska et al. (1981), 1983, 1990; Ulatowska and Sadowska (1992)</td>
<td>Groups (5–15 PWA), several cases</td>
<td>Personal narratives, picture description, story retelling</td>
<td>Analysis of relevant/irrelevant content, distribution of information within superstructure components, distribution of important information and detail</td>
<td>No</td>
</tr>
<tr>
<td>Huber (1990)</td>
<td>72 PWA (global, Wernicke, Broca, amnesic), examples for 6</td>
<td>Picture description (cartoon)</td>
<td>Evaluation of essential/optional propositional content</td>
<td>Yes</td>
</tr>
<tr>
<td>Glosser and Deser (1991)</td>
<td>9 PWA (fluent)</td>
<td>Personal narratives</td>
<td>Associated thematic organisation with global coherence and rated every verbalisation’s appropriateness with respect to the general topic of conversation</td>
<td>No</td>
</tr>
<tr>
<td>Ulatowska and Chapman (1994); Ulatowska et al. (1999)</td>
<td>4 case studies; 15 PWA, 15 NBD</td>
<td>3 Aesop’s fables (written), Spontaneous speech, expository discourse 2 picture stories, 6 verbal fables</td>
<td>Analysis of the ability to manipulate macrostructure, namely provide the main idea/gist and the moral of the story; analysis of linguistic means of signalling macrostructure (e.g., use of connectors, reference)</td>
<td>Restricted in some aspects of macrostructure</td>
</tr>
<tr>
<td>Christiansen (1995)</td>
<td>15 PWA (fluent)</td>
<td>Picture series description (4 cartoon stories)</td>
<td>Propositional content analysis (events, states, elaborations, comments)</td>
<td>Yes</td>
</tr>
<tr>
<td>Ulatowska and Olness (2007), Olness and Ulatowska (2011)</td>
<td>Single case examples</td>
<td>Personal narratives</td>
<td>Analysis of the event line, temporal organisation, evaluative devices</td>
<td>No</td>
</tr>
</tbody>
</table>

Note: PWA, people with aphasia; NBD, non-brain-damaged people; RBD, people with right hemisphere brain damage.
Table 4. Studies addressing cohesion in aphasic discourse: Methods, findings, and the relationship between cohesion and coherence.

<table>
<thead>
<tr>
<th>Investigator(s)</th>
<th>Group/multiple case/case</th>
<th>Method</th>
<th>Impaired?</th>
<th>Is cohesion necessary for coherence?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Piehler and Holland (1984)</td>
<td>2 case studies</td>
<td>Halliday and Hasan (1976) system to study cohesive ties. Five major classes of cohesion devices: reference, substitution, ellipsis, conjunction, lexical reiteration and collocation</td>
<td>Recovery over time</td>
<td>No</td>
</tr>
<tr>
<td>Ulatowska et al. (1981), 1983</td>
<td>10 PWA, 10 NBD; 15 PWA, 15 NBD</td>
<td>4-point rating scale of clarity as a rough measure of cohesion; Marking of verb tenses, definiteness of nominals, appropriate connectors for sequencing</td>
<td>No</td>
<td>–</td>
</tr>
<tr>
<td>Armstrong (1987)</td>
<td>3 case studies</td>
<td>Halliday and Hasan (1976) system; Cohesive Harmony Index (CHI)—the percentage of tokens in the chains to the total number of tokens</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Glosser and Deser (1991)</td>
<td>9 PWA (fluent), 9 patients with Alzheimer's disease; 9 patients with CHI</td>
<td>Halliday and Hasan (1976) system</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Bloom, Borod, Santschi-Haywood, Pick, and Obler (1996)</td>
<td>12 PWA, 9 RBD, 12 NBD</td>
<td>Analysis of specific lexical devices used to establish cohesion (Gleason et al., 1980, Nicholas, Obler, Albert, &amp; Helm-Estabrooks, 1985, Ulatowska et al. 1981); analysis of referential system (anaphora, deictic terms, indefinite terms, definite articles) and connective devices</td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Andreetta et al. (2012), Andreetta (2014)</td>
<td>10 PWA (anomic), 10 NBD; 20 PWA, 20 NBD</td>
<td>Index of cohesiveness (misuse of cohesive ties: errors in anaphoric pronouns, of function words or semantically related content words, number and gender agreement between pronouns and nouns, abrupt interruption)</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>

Note: PWA, people with aphasia; NBD, non-brain-damaged people; RBD, people with right hemisphere brain damage.
Coherence

Micro- and macrostructural phenomena together allow us to perceive a collection of words as sentences, or utterances, and a group of sentences or utterances as text or connected speech. This quality of discourse—its unity, connectedness—is called “coherence”. Coherence can be divided into local, established at the sentence level, and global coherence that binds larger constituents together, although this distinction is often disregarded. Glosser and Deser (1991) refer to the overall theme, goal or plan of discourse as “global coherence”, while “local coherence” determines the conceptual ties between individual propositions.

Findings on coherence in aphasia have been largely controversial. Behind this controversy are variations in definitions and methodologies used to study this phenomenon, as well as a lack of congruity in the aphasia types of participants in different studies (Table 4). The concept of coherence is rather complex. Depending on the research group, it is explained, fully or partially, through thematic relatedness, topic maintenance (Glosser & Deser, 1991; Halliday & Hasan, 1976), or a general notion of semantic unity, in which every part of discourse “hangs together” (Flotz, 2007; Olness, 2006; Ulatowska, Olness, & Williams, 2004). The problem of defining the nature of coherence in discourse has been addressed by a number of researchers outside of the field of language pathology (Foltz, 2007; Gernsbacher & Givón, 1995; Kehler, 2002, 2004; Kintsch & Van Dijk, 1978; Sanders & Spooren, 2001; Van Dijk, 1977).

One of the methods commonly used in aphasia research is rating on the five-point scale developed by Glosser and Deser (1991; e.g., Laine et al. 1998; Rogalski et al. 2010). The procedure includes segmenting samples into verbalisation, a verbalisation being an independent clause with all its dependent clausal and nonclausal elements, and rating the contextual appropriateness of every verbalisation by two independent trained raters. To assess global coherence, the relevance of a verbalisation to the topic of conversation is rated, whereas for local coherence it is the appropriateness with respect to the immediately preceding utterance. Glosser and Deser (1991) found no difference between coherence ratings for their NBD and fluent aphasic groups. Several alternative shorter—three- or four-point—scales have been suggested to measure coherence in a similar way (Koutsoftas, Wright, & Capilouto, 2009; Ulatowska et al., 1983; Van Leer & Turkstra, 1999; Wright, Koutsoftas, Fergadiotis, & Capilouto, 2010, 2013). Although Ulatowska and colleagues claimed the discourse of their aphasic participants to be well-structured, other studies reported higher scores for nonlanguage-impaired speakers. Fergadiotis and Wright (2011) compared the classic analysis developed by Glosser and Deser (1991) with a similar discourse coherence rating scale (Koutsoftas et al., 2009) and a computationally calculated objective coherence measure based on a cognitive model of knowledge acquisition—Latent Semantic Analysis (Landauer & Dumais, 1997). Their findings indicated the existence of a strong correlation between the three measures, and a direct link between global coherence and aphasia severity.

Christiansen (1995) was the first to demonstrate that coherence in the discourse of aphasic speakers was different than that of NBD speakers. She analysed the discourse of three groups of people with fluent aphasia of different types, namely anomic, conduction, and Wernicke’s, from a perspective of propositional coherence, that is, the texts were divided into propositions, and the propositional content of the discourse samples was studied in terms of four functional categories: events, states, elaborations, and comments. Propositions were then rated in terms of coherence violations, such as information gaps, progression and relevance violations. Christiansen’s analysis revealed different patterns of coherence impairments in the three
aphasic groups, as well as individual variability within the groups. Christiansen emphasised the potential impact of aphasia type on narrative production strategies and suggested that aphasia types needs to be considered when interpreting results of discourse coherence studies.

The approach of Marini and colleagues (e.g., Marini, Carlomagno, et al., 2005; Marini et al., 2008, 2011) is based on the analysis of cohesion and coherence errors. This method was used to study discourse of a group of individuals with anomic aphasia and a larger group of participants with other types of fluent aphasia (Andreetta, 2014; Andreetta & Marini, 2015; Andreetta et al., 2012). Their clinical groups’ performance differed from that of NBDs on almost all of the measures. Specifically, speakers with aphasia had more local and global coherence errors than their healthy counterparts. Interestingly, the author’s qualitative analysis of global coherence in anomia showed that it was disrupted by propositional repetitions and filler utterances, and not by irrelevant and tangential propositions. Table 5 contains definitions, methods, and results of the above mentioned studies to help navigate through the growing body of research on coherence in aphasia.

While ratings-based assessment presents an opportunity to capture coherence as an overall property of discourse, the main disadvantage of the available rating scales consists in their addressing slightly different constructs, which raises construct and convergent validity issues, and potentially leads to incomparable outcomes. On the other hand, methods based on error counts are generally more reliable, provided the technique is well-tested, but they risk only partially grasping the complex combination of processes behind coherence. The discussion in this section is centred around “textual” coherence, striped of extralinguistic context, such as common ground, world knowledge, or shared visual space provide, and without direct consideration of the multimodal nature of natural communication. These factors, however, have been considered to influence discourse comprehension, rendering linguistically limited and/or incoherent input coherent and adequate to situation (e.g., Goodwin, 2000; Hobbs, 1979; cf. subsection “Modality” further in this review). Of all the methods used for the assessment of discourse coherence, only the perceptual rating scales, which require human raters to evaluate discourse as a whole, potentially adjust for some of the extralinguistic content, including gesturing, in case raters are presented with a video recording. Combining subjective ratings and text-based measures, and taking into account other factors, such as thematic informativeness (e.g., Ulatowska et al., 1990), can shed more light on how coherence is achieved, and what causes its disruption.

**Relationship between coherence and cohesion**

It has been suggested that a large number of incomplete cohesive ties and a limited range of connective forms are responsible for discourse in aphasia often being perceived as vague and ambiguous (cf. Bloom, 1994). The question about the contribution of cohesion to coherence has not been answered yet (e.g., Armstrong, 2000; Reinhart, 1980; Ulatowska et al., 1981), although the concept of cohesion has been extensively explored in discourse studies (cf. Xi, 2010 for a review). A direct dependency has been hypothesised to exist between cohesion and the overall coherence of discourse (e.g., Coelho et al., 1994; Halliday & Hasan, 1976; Hasan, 1985; Piehler & Holland, 1984). Armstrong (1987) introduced the notion of “cohesive harmony”, which refers to the interaction of cohesive elements within a text, and demonstrated that its amount correlated with listeners’ perception of coherence. Glosser and Deser (1991) also claimed that coherence is expressed through cohesive devices, such as coreference.
Table 5. Studies addressing coherence in aphasic discourse: definitions, methods, and findings.

<table>
<thead>
<tr>
<th>Investigator(s)</th>
<th>Group/multiple case/case</th>
<th>Definition of coherence</th>
<th>Elicitation task</th>
<th>Method</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ulatowska et al. (1981), 1983)</td>
<td>10 PWA, 10 NBD; 15 PWA, 15 NBD</td>
<td>Well-formedness of a text in terms of plausibility, conventionality, and conclusiveness of text</td>
<td>Personal narrative, procedural discourse, story retelling</td>
<td>Analysis of superstructure (settings, resolutions, etc.) and procedural steps Analysis of evaluation and adverb. modification Subjective ratings: 2- or 3-point scales for ambiguity and content, 3-point clarity scale</td>
<td>Narratives in aphasia were found to be well formed, and included all the superstructure elements. Discourse error analysis showed the existence of a continuum from disrupted to normal linguistic performance</td>
</tr>
<tr>
<td>Glosser and Deser (G&amp;D, 1991)</td>
<td>9 PWA (fluent), 9 patients with Alzheimer’s disease, 9 patients with closed head injury (CHI)</td>
<td>A term used to characterise conceptual organisational aspects of discourse at suprasentential level; it depends at least partially on the speaker’s ability to maintain thematic unity (p. 69)</td>
<td>Personal narrative</td>
<td>5-point rating scale</td>
<td>No difference between aphasic and control group in ratings of either global or local coherence</td>
</tr>
<tr>
<td>Coelho et al. (1994);</td>
<td>Longitudinal case study</td>
<td>Story grammar (number of complete episodes consisting of an initiating event, an action, and a direct consequence)</td>
<td>Picture series (19 frames) and single picture description</td>
<td>Number of episodes</td>
<td>While cohesion improved, story grammar remained moderately depressed over 12 months</td>
</tr>
<tr>
<td>Christiansen (1995)</td>
<td>15 PWA (5 anomic, 5 conduction, 5 Wernicke). 20 NBD</td>
<td>C. implies that consecutive actions are really segments of a single global intention and that there is an overall purpose or goal to be achieved (p. 292)</td>
<td>Picture series description (4 cartoon stories)</td>
<td>Analysis of propositional content (events, states, elaborations, and comments), coherence violations (information gaps, relevance and progression)</td>
<td>Discourse skills of PWA and NBD were different</td>
</tr>
</tbody>
</table>

(continued)
Table 5. (Continued).

<table>
<thead>
<tr>
<th>Investigator(s)</th>
<th>Group/multiple case/case</th>
<th>Definition of coherence</th>
<th>Elicitation task</th>
<th>Method</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coelho and Flewellyn (2003)</td>
<td>Longitudinal case study</td>
<td>Global c. refers to how discourse is organised with respect to the global plan, goal, or topic, local—to the maintenance of meaningful conceptual links between individual sentences within a text (p. 174)</td>
<td>Picture series (19 frames) and single picture description</td>
<td>3-point modification of G&amp;D (1991) scale (Van Leer &amp; Turkstra, 1999 version converted into z-scores)</td>
<td>As severity of aphasia decreased and microlinguistic skills recovered, the coherence remained impaired and showed minimal change over 12 months</td>
</tr>
<tr>
<td>Olness (2006); Olness and Ulatowska (2011)</td>
<td>12 PWA</td>
<td>“Hanging together” of the text as a unit (p. 176)/The overall semantic unity of themes and topics in a discourse (Glosser 1993)</td>
<td>Picture description; examples from personal narratives</td>
<td>Analysis of narrative superstructure (sequential order, setting, initiating event, complicating action, resolution, coda (Berman 1977, Labov, 1972)). A single rating of semantic unity is used</td>
<td>Abbreviation of narrative event line in aphasia, aphasia severity level correlates with narrative production skills</td>
</tr>
<tr>
<td>Wright et al. (2010, 2013)</td>
<td>14 PWA, 14 NBD; 50 healthy adults</td>
<td>Global c. reflects how discourse relates to the overall topic (G&amp;D 1991)</td>
<td>Picture series description (two wordless books)</td>
<td>4-point rating scale (Koutsoftas et al., 2009)</td>
<td>Control group received significantly higher scores than the aphasic group</td>
</tr>
<tr>
<td>Marini et al. (2011); Andreetta et al. (2012); Andreetta (2014); Andreetta and Marini (2015)</td>
<td>2 case studies; 10 PWA (anomic), 10 NBD; 20 PWA, 20 NBD</td>
<td>The ability to semantically relate remote utterances in the framework of a given discourse</td>
<td>Picture series and single picture description</td>
<td>Local and global coherence errors, cohesion errors</td>
<td>Different performance of clinical and control groups</td>
</tr>
</tbody>
</table>

Note: PWA, people with aphasia; NBD, non-brain-damaged people; RBD, people with right hemisphere brain damage.
However, several authors argued that coherence may be impaired while cohesion is not (Bloom et al., 1996; Coelho & Flewellyn, 2003; Coelho et al., 1994), and conversely, referential cohesion is not a prerequisite for establishing coherence (Glosser & Deser, 1991; Keenan, Baillet, & Brown, 1984; Ulatowska et al., 1981, 1983). Many researchers agree that micro- and macrolinguistic abilities are independently organised (e.g., Giora, 1985; Glosser & Deser, 1991; Lenk, 1998; Tanskanen, 2006; Ulatowska et al., 1981, 1983). Although it may seem logical that cohesion, which is related to the local coherence of discourse, is necessary for its overall coherence, other factors, or even a combination of factors, may have a stronger contribution in establishing coherence. Cohesion belongs in between micro- and macrolinguistic levels, which makes it harder to disentangle lexical and syntactic deficits in aphasia from cohesion impairment, and, in turn, the effect of all of them on coherence. Once again, one is faced with the issue of the understudied interplay between different levels of language production.

**Multilevel approaches**

It has been noted before that the existing measures have separately failed to capture all the aspects of the complex multidimensional process of discourse production (Elvevåg, Foltz, Weinberger, & Goldberg, 2007; Halliday & Hasan, 1976; Lorch & O’Brien, 1995; Sanders, Spooren, & Noordman, 1992). Thus, several combinations of different scales and methods have been suggested (cf. Table 6). For example, such features as relevance, discourse grammar analysis, clarity disruptors, and cohesion were included in the analysis of the interaction between structural and functional aspects of narrative and procedural discourse by Sherratt (2007). Although it has only been applied to healthy speakers, and adults with right hemisphere damage (Sherratt & Bryan, 2012), Sherratt concluded that multilevel analysis of discourse production, including pragmatic and linguistic measures, is important for theory and for therapy, as it provides the understanding of the underlying mechanisms of the process and their interrelations.

A number of multilevel approaches have been devised combining word, sentence, and discourse level measures to study discourse in aphasia (Marini et al., 2011; Prins & Bastiaanse, 2004; Wright & Capilouto, 2012). Glosser and Deser (1991) used 11 different measures, including syntactic and lexical errors, cohesion, and thematic coherence analyses. Their method was implemented in a multilevel approach developed by Wright and Capilouto (2012) that combined micro- and macrolinguistic measures, including syntactic complexity, information content, lexical diversity, and global coherence. Marini and colleagues (Andreetta & Marini, 2015; Marini et al., 2011) developed a multilevel procedure specifically designed for the assessment of macro- and microlinguistic skills of PWA. Some of the variables at the microstructural level, for example, semantic paraphasias, omissions of morphosyntactic information, sentence completeness, were complemented by the count of errors of cohesion and both local and global coherence (e.g., Andreetta et al., 2012; Marini, Boewe, Caltagirone, & Carloomagnos, 2005; Marini, Carloomagnos, et al., 2005). Both groups of researchers (Andreetta et al., 2012; Marini et al., 2011; Wright & Capilouto, 2012) found evidence of macro-, as well as microlinguistic impairments in aphasic discourse.

The main advantage of multilevel analyses is the possibility to account for the interrelatedness among linguistic processes at different levels, which results in a more comprehensive understanding of the factors influencing language production. Two major problems, however, pertain. First, the comprehensive picture drawn by the multilevel approaches is blurred by the lack of consideration of the interactions between different
Table 6. Multilevel approaches to studying discourse in aphasia.

<table>
<thead>
<tr>
<th>Investigator(s)</th>
<th>Aim</th>
<th>Included measures</th>
<th>Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vermeulen et al. (1989); Bastiaanse et al. (1996);</td>
<td>Analysis of spontaneous speech in aphasia</td>
<td>Speech rate (words/min), MLU (Mean Length of Utterances), analysis of conjunctions, auxiliary verbs, prepositions, empty words, incomprehensible speech Type-token ratio (TTR), number of semantic and phonetic paraphasias, neologisms; Distribution of subordinate/main clauses, distribution of lexical words/copulas;</td>
<td>Impaired ability to elaborate the main clause in the aphasic group, but normal ability to construe matrix clauses (reduced ability to use complex grammatical structures); reduced frequency and diversity of verbs</td>
</tr>
<tr>
<td>Glosser and Deser (1991)</td>
<td>Investigation of discourse organisation, in particular the dissociation between micro- and macrolinguistic cognitive processes, in speakers with different types of brain damage (Alzheimer’s disease, head injury, and post-stroke fluent aphasia)</td>
<td>Local and global coherence, cohesion, lexical measures (literal and verbal paraphasias, indefinite terms), syntactic complexity and occurrence of syntactic errors</td>
<td>Aphasic group was found to be impaired on measures of syntactic completeness and complexity, macrolinguistic aspects of discourse production, lexical-semantic abilities, and demonstrated inappropriate use of cohesive ties; thematic coherence within normal limits</td>
</tr>
<tr>
<td>Sherratt (2007), Sherratt and Bryan (2012)</td>
<td>Assessment of the applicability and utility of using a multi-level discourse processing model to examine the interaction between different linguistic levels (RBD and healthy speakers)</td>
<td>Fluency, lexical, and morphologic processing: Cohesion (Hallida &amp; Hasan 1976), clarity disruptors (nonspecific elements and word substitutions); Syntactic analysis (including the analysis of clausal structures); Discourse grammar, content and fluency disruptors, relevance, productivity</td>
<td>Using a multilevel procedure gives a more realistic perspective of discourse. Certain discourse features that can be assessed more objectively, and correlations can provide explanatory information regarding more subjective concepts</td>
</tr>
<tr>
<td>Investigator(s)</td>
<td>Aim</td>
<td>Included measures</td>
<td>Results</td>
</tr>
<tr>
<td>----------------</td>
<td>-----</td>
<td>-------------------</td>
<td>---------</td>
</tr>
<tr>
<td>Marini et al. (2011), Andreetta et al. (2012), Andreetta and Marini (2015)</td>
<td>Investigation of discourse processing in anomic aphasia, and the connection between lexical impairment and discourse organisation</td>
<td>Productivity: MLU, total count of units—phonologically well-formed words except for fillers, non-words, false-starts, sounds and syllable repetition Lexical processing: semantic paraphasias, phonological errors Syntactic encoding: percent of complete sentences Discourse organisation: local and global coherence errors, index of cohesiveness Information content: number of thematic units, percent of lexical information units (grammatically, pragmatically, and phonologically well-formed and used content and function words)</td>
<td>Reduced abilities of lexical retrieval, increased amount of global coherence errors, reduced lexical informativeness Slower than normal speech rate The data suggests that PWAs don’t have conceptual organisation problems, and that their macro linguistic deficit is an epiphenomenon of the micro linguistic problems</td>
</tr>
<tr>
<td>Wright and Capilouto (2012)</td>
<td>Exploring maintenance of global coherence in aphasia</td>
<td>Syntactic complexity: complexity index (CI)—total clauses/total independent clauses; Informativeness: percent of information units (words that were intelligible, neologisms, fillers, partial words, and commentary on the task); Lexical diversity (using voc-D tool from CLAN); 4-point scale for global coherence (Wright et al., 2010, Fergadiotis &amp; Wright, 2011)</td>
<td>Lower global coherence scores for the aphasic group. In terms of influence of micro linguistic processes on maintenance of global coherence, the best predictors were the percent of information units and lexical diversity</td>
</tr>
</tbody>
</table>
factors. For example, Glosser and Deser (1991) conducted a principal component analysis (PCA) and found that the variables they included are naturally clustered into three groups: those related to suprasentential organisation, those reflecting lexical processing, and syntactic measures. The results of the PCA, along with separate ANOVAs for lexical, syntactic, cohesion, and coherence measures, provided rather compelling evidence to the dissociation between micro- and macrolinguistic levels. However, correlation analysis presented by Andreetta et al. (2012), as well as the regression analysis in the study of Wright and Capilouto (2012) spoke for the existence of a connection between coherence and lexical informativeness. At the same time, in the study of Andreetta et al. (2012), both syntactic completeness and coherence were impaired in the PWA group, but the relationship between these two variables was not addressed. Some clarity was added by Wright & Capilouto (2012), who specifically aimed at exploring connections between lexical and syntactic variables, and coherence. Their regression analysis suggested that, despite the correlations between syntactic and coherence measurements in both groups, syntactic processes did not contribute significantly to the establishment of global coherence. It is noteworthy, that coherence rating scales used by Glosser and Deser (1991), and Wright and Capilouto (2012) have the same underlying concept of coherence, but a comparison of the two scales suggested that the latter one was possibly more reliable (Wright et al., 2010).

The issue of comparability between the results of different studies, which has been one of the key points of this review, is even more pronounced here, as different measures and combinations of measures are included in the existing multilevel approaches. Some of the measures refer to the same concepts, but use different mechanisms to assess them, while others target the same concepts, but operationalise them differently. Future investigations of the relations between different levels of language production should focus on ascertaining the construct validity of the existing metrics and the degree to which they converge. After the methodological foundation is stable, the interactions between variables at different levels should be explored, keeping in mind that some of the correlations are potentially arbitrary.

Effectiveness and efficiency

Despite the fact that many linguistic abilities and structural components of discourse may be impaired in aphasia, people with aphasia are often able to maintain functional communication (Holland, 1982; Meuse & Marquardt, 1985). Substantially more work has been done, however, on the exploration of separate components of the linguistic apparatus than on the overall communication success in aphasia. In addition to “functional communication” and “communicative success”, the term “effectiveness” has been used to describe the ability of people with aphasia to produce meaningful and understandable discourse, that is, to reach their communicative goal. Whereas “efficiency” reflects how effortlessly and timely they manage to do so, the distinction between these two notions has been often disregarded.

Manochioping et al. (1992) reviewed 15 communicative effectiveness measures and subdivided them into five types: observational profiles, communicative efficiency measures, standardised testing in real and/or simulated situations, significant others questionnaires, and composite assessment. They noted that several “communicative efficiency measures [...] reflect the combined effectiveness of a number of pragmatic behaviors in achieving functional communication” (1992, p. 521).
It has been argued that communicative success should be the ultimate goal of aphasia treatment, and that the relevant assessment tools should be able to capture the improvement of functional communication skills over time, which makes them good therapy outcome measures. Supporting this theory, the Communicative Effectiveness Index (CETI) was found to correlate with Western Aphasia Battery scores (Bakheit, Carrington, Griffiths, & Searle, 2005; Lomas et al., 1989). Recently, however, commonly used techniques, such as CETI and Communication Activities in Daily Living-Second Edition (CADL-2; Holland, Frattali, & Fromm, 1999), have been re-evaluated and several new methods have been proposed to complement them. Thus, the Scenario Test (Van Der Meulen, Van De Sandt-Koenderman, Duivenvoorden, & Ribbers, 2010) extends the Amsterdam–Nijmegen Everyday Language Test (ANELT, 1995; Blomert, Kean, Koster, & Schokker, 1994), an instrument for the assessment of verbal communication in aphasia, for multimodal communication. It provides an outcome measure for alternative and augmentative communication (AAC) therapy, through which people with severe and moderate aphasia learn to rely not only on verbal but also nonverbal strategies, such as gesturing, to transmit information in a conversation. The importance of communication effectiveness assessment stimulated the development and adaptation of the existing measures for English (e.g., Long, Hesketh, Paszek, Booth, & Bowen, 2008) and other languages, for example, Danish (Pedersen, Vinter, & Olsen, 2001), Italian (Muò et al, 2015), and Cantonese (Kong & Law 2004, 2009), although the latter was focused on the linguistic component of functional communication.

A different group of methods developed for the assessment of purely linguistic aspects of communication success includes one of the most popular measures of informative efficiency calculated based on informativeness, namely correct information units per minute (CIU/min, Nicholas & Brookshire, 1993). Although this measure is related to a very specific aspect of functional communication, it was found to correlate with naive listeners’ perception of communicative abilities (Jacobs, 2001) and to contribute to the classification of aphasia into mild, severe, and moderate (Im, Kwon, & Sim, 2001). Poor inter- and intra-rater reliability, the drawback of the correct information unit analysis (Nicholas & Brookshire, 1993), stimulated the development of a reliable, ecologically valid measure of interaction success in conversations (Ramsberger & Rende, 2002). The latter is based on the number of main ideas transferred between a person with aphasia and a non-aphasic partner in a conversation. Another quantitative method of this group, content unit analysis (Yorkston & Beukelman, 1980, see above), was adapted for the ANELT scenarios, and was suggested to be more sensitive than the original ANELT Comprehensibility scale in detecting changes in PWA’s verbal effectiveness over time (Ruiter, Kolk, Rietveld, Dijkstra, & Lotgering, 2011). The new techniques for the assessment of communicative effectiveness and efficiency in aphasia reviewed in this section are gathered in Table 7.

The connection between linguistic competence and communicative effectiveness is not very straightforward (e.g., Fridriksson, Nettles, Davis, Morrow, & Montgomery, 2006). For example, measures of pragmatic performance and functional communication have been shown to address different, though possibly overlapping, aspects of performance in PWA (Irwin, Wertz, & Avent, 2002; McCollough et al., 2006). Armstrong and Ferguson (2010) addressed the role of language in “functional communication” and suggested that further investigation of different behaviours and skills contributing to functional communication, both expressive and receptive, and within context, is crucial for the improvement of aphasia assessment and treatment. Generally, the connection between the main approaches to studying discourse—structural, functional, and
cognitivist—has been largely ignored in previous studies (Armstrong, 2000), leaving the phenomena occurring at their interface unexplored. Combining approaches from different perspectives and assessing both linguistic parameters and the overall effectiveness of speech in aphasia within a multilevel procedure similar to the ones described in the

<table>
<thead>
<tr>
<th>Investigator(s)</th>
<th>Group/multiple case/case</th>
<th>Aim</th>
<th>Method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pedersen et al. (2001)</td>
<td>68 PWA</td>
<td>Adaptation and psychometric evaluation of CETI for Danish</td>
<td>CETI for Danish, test-retest reliability</td>
</tr>
<tr>
<td>Irwin et al. (2002)</td>
<td>20 PWA</td>
<td>Establish the relationship among language impairment, functional communication, and pragmatic performance in aphasia</td>
<td>Porch Index of Communicative Abilities (PICA, Porch, 1967), Rating of Functional Performance (RFP, Wertz et al., 1981), and Pragmatic protocol (Prutting &amp; Kirchner, 1987); Correlation analysis</td>
</tr>
<tr>
<td>Ramsberger and Rende (2002)</td>
<td>14 PWA, 56 NBD</td>
<td>Measuring transactional success in conversation</td>
<td>Cantonese Linguistic Communication Measure (CLCM), including the Index of communication efficiency (N of informative verbs per min)</td>
</tr>
<tr>
<td>Kong and Law (2004, 2009)</td>
<td>10 PWA, 30 NBD; 5 PWA (longitudinal)</td>
<td>Communication effectiveness assessment in Cantonese aphasics speakers</td>
<td>Cantonese Linguistic Communication Measure (CLCM), including the Index of communication efficiency (N of informative verbs per min)</td>
</tr>
<tr>
<td>Van Der Meulen et al. (2010)</td>
<td>122 PWA, 25 NBD</td>
<td>Assessment of verbal and non-verbal communication in severe aphasia</td>
<td>Scenario test for both verbal and non-verbal communication in daily life situations (4-point scale) and interactive settings</td>
</tr>
<tr>
<td>Ruiter et al. (2011)</td>
<td>10 PWA, 20 NBD</td>
<td>Development of a quantitative measure of verbal effectiveness and efficiency in the ANELT</td>
<td>Content unit analysis (Yorkston &amp; Beukelman, 1980) adapted for the ANELT scenarios</td>
</tr>
<tr>
<td>Muò et al. (2015)</td>
<td>60 PWA, 20 patients with traumatic brain injury, 100 NBD</td>
<td>Adaptation and psychometric evaluation of ASHA FACS for Italian</td>
<td>Italian version of ASHA-FACS for Italian (I-ASHA-FACS)</td>
</tr>
</tbody>
</table>

Note: PWA, people with aphasia; NBD, non-brain-damaged people; RBD, people with right hemisphere brain damage.
Methodological issues

A combination of qualitative and quantitative methods allows assessing both the overall quality and success of discourse, and the linguistic processes underlying discourse production. Bringing together these two perspectives can result in the development of more effective treatment programmes and methods. Although the research on discourse in aphasia advanced significantly over the past years, there is still a noticeable lack of congruity in the findings. We will now consider a few possible reasons as to why the growing number of studies and significant outcomes have not resulted in a better overall picture of aphasic discourse abilities yet. Armstrong (2000) suggested that the disparity in findings may be associated with a lack of congruity in the definitions of certain concepts or methodological differences. Indeed, we have already discussed that definitions, methods, and analyses vary largely from study to study. This part of the review touches upon the existing variability in study design, namely, in sample size, genre, elicitation task, and modality, and addresses crosslinguistic studies of discourse in aphasia.

Sample size

Depending on the goal of a particular study, the choice has to be made between a group and a multiple or a single case design. While group studies aim at demonstrating patterns and interactions of certain variables within a population, case studies can provide valuable counterevidence, and multiple cases can be used to demonstrate the existence of dissociations or opposite tendencies within a population. Case studies demonstrated that discourse abilities can be impaired in some individuals with mild aphasia (Coelho et al., 1994), whereas they may remain within the normal range in other aphasic speakers (Armstrong, 1992). Case studies have also been used to confirm the existence of participant- and task-related variations in the performance (Armstrong et al., 2011) and to demonstrate the applications of newly proposed methodologies (Boles, 1998; Marini et al., 2011; Olness & Gober, 2013). Stark (2010) reported on an aphasic speaker who improved over time in lexical and syntactic skills, and in narrative informativeness at story retelling, while the individual with aphasia in the study of Coelho and Flewellyn (2003) had no consistent improvement over a year of similar training. Though they are able to shed light on the internal language organisation and provide challenging counterexamples, case studies do not allow making generalisations about performance of a clinical population.

One of the reasons for the lack of large-scale group studies is that aphasic data collection and analysis is a long and complicated process. Most of the studies on aphasic discourse have been based on the analysis of small-to-medium-sized groups of participants (e.g., 3–10 PWA in Armstrong & Ulatowska, 2007; Glosser & Deser, 1991; Hough, 1990; Ruigendijk, Vasić, & Avrutin, 2006; Ulatowska et al., 1981; 11–15 in Christiansen, 1995; Goodglass, Christiansen, & Gallagher, 1993; Olness, 2006; Ulatowska et al., 1983), although for several studies larger number of participants were recruited (e.g., 20 agrammatic PWA in Miceli, Silveri, Romani, & Caramazza, 1989; 28 in Olness, 2007; 74 in Wagenaar, Snow, & Prins, 1975; 121 in Vermeulen, Bastiaanse, & Van Wageningen, 1989).

For the purpose of reducing the amount of time and labour that data collection and analysis take, and to make larger data samples available to researchers in various languages, a considerable effort has recently been put into creating corpora of aphasic
speech and test results. For example, AphasiaBank contains spoken language samples, action and object naming, repetition, as well as general assessment data of 311 PWA speaking eight languages by 1 May 2014 (Forbes, Fromm, & MacWhinney, 2012; Fromm, Forbes, Holland, & MacWhinney, 2014; MacWhinney, 2000; Macwhinney, Fromm, Forbes, & Holland, 2011, 2013). Corpus of Dutch Aphasic Speech (CoDAS; Westerhout & Monachesi, 2005) has been designed to have part of speech, syntactic and prosodic annotation of aphasic spoken language recorded in different communicational settings, but only a pilot study with six participants has been reported so far (Westerhout, 2006). Another project—the Moss Aphasia Psycholinguistics Project Database (MAPPD; Mirman et al., 2010)—contains a collection of behavioural test data from aphasic speakers; the core of the archive consists of the Philadelphia Naming Test data for more than 170 participants. One of the greatest advantages of aphasic speech corpora is the possibility to test new methodologies and implement multiple analyses on the same data sets. Hence, a much better comparability of results can be achieved. The relevance and the practical benefits of shared data sets are also recognised by funding agencies. However, when choosing a corpus collected by a third party for generalised research purposes, its compatibility with the goals of the research should be thoroughly assessed. For example, researcher interested in the factor of genre would have to ascertain that the genres of interest are available within the corpus, while for studies focusing on different modalities, a corpus has to be not just audio but also video based.

**Genre**

Many researchers noted that the choice of elicitation task influences quality and quantity of the discourse produced (Coelho, 2002; Olness, 2006, 2007; Olness, Ulatowska, Wertz, Thompson, & Author-Steefan, 2002; Van Leer & Turkstra, 1999). Picture description has perhaps been the most widely used technique, as it guarantees comparability of produced discourse samples (e.g., Brookshire & Nicholas, 1994; Nicholas & Brookshire, 1993; Olness et al., 2002; Wright et al., 2005). It has been demonstrated that single picture descriptions should not be chosen for discourse studies, as the descriptive discourse genre does not require establishing coherence (Olness, 2006, 2007). It was also suggested that there is an impact of aphasia severity on narrative production, while it is not the case in picture descriptions (Olness, 2006). Procedural discourse (Ulatowska et al., 1981, 1983; Weinrich, Mccall, Boser, & Virata, 2002) and personal narratives (Armstrong & Ulatowska, 2007; Behrns, 2009; Olness & Ulatowska, 2011; Ulatowska, Reyes, Santos, Garst, Vernon, et al., 2013; Ulatowska et al., 2004) received special attention because they are essential for the everyday life of people with aphasia, and because they provide a view on a wide range of linguistic and extralinguistic skills. However, these two genres were shown to impose different linguistic and cognitive demands on the speakers. For example, previous studies reported that procedural discourse of both people with and without aphasia had lower syntactic complexity compared to narratives, whereas narratives of PWA had lower information content, shorter storyline, more errors in the order of events, and syntactically less complex language than those of control participants (Ulatowska et al., 1990, 2004).

Gernsbacher and Givón (1995) emphasised that coherence is a property emerging during speech production as well as comprehension, allowing a listener/reader to reconstruct discourse as a reader/writer had it in mind, or the mental representation of it. Studies on conversation in aphasia focused on the ability of PWA to co-construct meaning in communication through the analysis of phenomena such as turn-taking, repair strategies, collaborative referencing, and the effects of aphasia severity, conversation partner, topic,
and other potential factors on discourse production. (e.g., Beeke, Maxim, & Wilkinson, 2007; Damico, Oelschlaeger, & Simmons-Mackie, 1999; Ferguson & Harper, 2010; Hengst, 2003; Linebaugh, Kryzer, Oden, & Myers, 2006; Perkins, 1995; Perkins & Goodwin, 2003). Different grammatical patterns were discovered in aphasic informal conversation compared to monologues or picture-induced discourse (Armstrong et al., 2011; Wilkinson, Beeke, & Maxim, 2010). Although aphasic speakers have been reported to successfully use conversational repair strategies, monologue speech has been considered to be more grammatical (Beeke, Wilkinson, & Maxim, 2003). The special issue of Aphasiology (Wilkinson, 2015) on CA application to aphasic data recently addressed topics such as repair (Barnes & Ferguson, 2015; Laasko, 2015; Penn, Frankel, & Wilkinson, 2015), adapted behaviours of conversation partners (Klippi, 2015), and interaction-focused therapy for aphasia (Beeke et al., 2015; Damico et al., 2015; Saldert, Johansson, & Wilkinson, 2015), emphasising the importance of studying language in interaction and taking into account the role of conversation partners in the recovery dynamics of people with aphasia.

**Modality**

Genre-related differences in discourse production bring up a related question of modality. For example, spontaneous speech has been studied much more than written discourse (Prins & Bastiaanse, 2004; Rossi & Bastiaanse, 2008; Vermeulen et al., 1989; Wagenaar et al., 1975). Multimodality in aphasic communication received substantially more attention in the years after Armstrong’s (2000) review. Behrns, Wengelin, Broberg, and Hartelius (2009) compared written and oral narrative production in PWA and non-impaired subjects and found written discourse to be generally better structured in both groups. De Riesthal (2011) noted that PWA performed better in speaking and pantomime compared to writing and drawing. He also argued that pictorial stimuli evoke better scores than printed and auditory ones.

Whereas written language has been studied less than oral discourse in aphasia due to the frequent inability of adults with aphasia to use this modality, gesturing has been understudied despite being potentially complementary or even compensatory to language (Goodwin, Goodwin, & Olsher, 2002; Klippi, 2015; Lanyon & Rose, 2009; Scharp, Tompkins, & Iverson, 2007). Previous findings suggest that the processes underlying gesture and language production are shared or closely related (e.g., Dipper, Cocks, Rowe, & Morgan, 2011; Goodwin, 2000; Mol, Krahmer, & Van De Sandt-Koenderman, 2012). Including gestures in further analyses of conversations with PWA can be insightful with respect to the general mechanisms of meaning co-construction in interaction (Pritchard, Dipper, Morgan, & Cocks, 2015).

**Crosslinguistic studies**

According to Beveridge and Bak (2011), 62% of all papers on aphasia between 2000 and 2009 were based on English material. Discourse-level representation, however, is arguably not language specific. Thus, crosslinguistic comparisons are extremely valuable in discourse studies. While a considerable amount of evidence comes from English, crosslinguistic studies in aphasiology started emerging as early as in the 1980s. Edwards (1981) examined spoken language samples of a Japanese, a Turkish, a Russian, and a Zulu speakers with Broca’s aphasia and found language output to be strikingly similar across languages. Case studies of narratives in aphasia in 14 languages have been collected in the
Agrammatic Aphasia: A Cross-language Narrative Sourcebook (Menn, Obler, & Miceli, 1990). A significant amount of work on the subject has also been done by Bates and colleagues, who demonstrated that there are language-specific differences within “the same” aphasic syndromes (e.g., Bates & Wulfeck, 1989; Bates, Wulfeck, & MacWhinney, 1991; Wulfeck et al., 1989). Bastiaanse, Edwards, and Kiss (1996) discussed certain grammatical features of fluent aphasia in three languages to demonstrate that PWA had linguistic deficits rather than an impairment of control of speech production. MacWhinney and Holland promoted crosslinguistic research in aphasia further and initiated the creation of AphasiaBank, an open corpus of aphasic spoken data (Macwhinney et al., 2011). Various aspects of language production, such as time reference (Bastiaanse et al., 2011), textual coherence (Korpijaakko-Huuhka & Lind, 2012), syntactic deficit in Broca’s aphasia (Friedmann, 2006), were investigated in crosslinguistic perspective to disentangle language-specific impairments and general deficits of language production mechanisms. Some of the characteristic manifestations of aphasia are language specific, while others are not. Hence, investigating the same phenomena and similar deficits in different languages offers the possibility of generalisation. Studies on different languages can stimulate the development of therapeutic techniques for the language-impaired speakers of those languages. At the same time, crosslinguistic comparisons are important for understanding the universal mechanisms of language production and its deterioration in aphasia.

Conclusion

Discourse production is the most important channel for communication. Comparing discourse produced by NBD and aphasic people provides a valid source of information on the mechanisms behind human language generation and the nature of aphasia. Due to the complexity of this phenomenon, previous research has focused on disassembling the process of language production and studying its components. Although we have an idea about the building blocks of discourse, their functional load and the way they are organised into a whole is not yet entirely clear. Understanding how language production functions is crucial for understanding the reasons of its deterioration in aphasia, and is, thus, informative for therapy. Whereas currently available techniques mainly aim at refining assessment, further research should concentrate on establishing the connection between different components of language and identifying what constitutes verbal communication deficits in aphasia.

Although there is an ongoing discussion on various aspects of this process, some of the findings remain rather controversial. In recent years, substantially more attention has been given to the macrolinguistic skills of people with aphasia, who undeniably have difficulties with language-based communication. It is yet unclear, however, if their ability to construct understandable connected discourse is impaired per se, or if it is the result of various microlinguistic deficits.

In the latest comprehensive review, Armstrong (2000) highlighted a number of theoretical and practical issues in aphasic discourse studies. In the current review, we aimed to show that some of the questions raised by Armstrong have been addressed in the literature to date. For example, it has been demonstrated that there are genre-related differences in discourse produced by PWA. Armstrong also noted the importance of exploring the connection between micro- and macrolinguistic levels. Several recent studies on aphasic discourse targeted exactly this issue (e.g., Andreetta et al., 2012; Marini et al., 2011; Wright & Capilouto, 2012). However, a number of issues
remain unresolved. Methodological variability continues as a critical source of disparate findings. Though new methods and tests allow for a more objective and in-depth examination of certain phenomena, such as coherence or informativeness, they often lead to incomparable results. Future research on discourse in aphasia can benefit from comparative evaluation of existing methods and replications of earlier studies with groups of aphasic speakers with different levels of severity and types of aphasia.

When the methodological foundation is solidified, additional effort should be put into in-depth investigations of several multifaceted phenomena, specifically information content and distribution (informativeness and information structure, respectively), discourse structure or discourse organisation at the macrolinguistic level, and discourse coherence. Investigation of these features’ complex nature, for example, through double dissociations with other correlated linguistic variables and through their interaction with the micro- and extralinguistic levels, is essential for understanding the mechanisms underlying communication. Furthermore, only through a combination of approaches from both structural and functional perspectives can a complete picture of the mechanisms of aphasic discourse production be formed. That is, bringing together studies on purely linguistic features and those focusing on the overall conversation success is essential for understanding the role language plays in establishing communication. This could be achieved, for example, by devising a comprehensive multilevel procedure, which includes measures at different linguistic levels, and a perceptual component to account for communicative effectiveness. Data analysis for such a procedure apart from main effects should include interactions between different variables. It is crucial to focus on studying language in context, and adapting existing methods or creating new ones for the analysis of naturally occurring conversations. Settings similar to natural conversations can also be manipulated to investigate the effect of the degree of common ground between interlocutors and communicative strategy (e.g., more and less cooperative) of aphasic participants’ conversation partners on communication success. Including speakers with a range of patterns of linguistic impairments would make the contribution of different linguistic variables more obvious.

Although there is still a need to continue studying the means through which meaningful and understandable discourse is created and, in particular, the relationship between different linguistic levels involved in this process (Armstrong, 2000; Ulatowska & Olness, 1997, 2000), much valuable insight on discourse production in aphasia has been gained since 2000, when Armstrong’s review was published. Introducing new approaches, including a range of powerful theoretical resources and frameworks developed in the normative discourse analysis, is very important not only for aphasiology but also because aphasic data can be immensely informative in testing linguistic theories. Nonetheless, reaching better comparability between methodologies and reproducing results of previous studies can certainly accelerate future research on the matter.

Acknowledgements
We would like to express our gratitude to Leigh Fernandez for her critical opinion and suggestions on the earlier version of this review. We want to acknowledge Erasmus Mundus International Doctorate for Experimental Approaches to Language and Brain (IDEALAB).

Disclosure statement
No potential conflict of interest was reported by the authors.
Funding
The first author’s research was supported by Erasmus Mundus International Doctorate for Experimental Approaches to Language and Brain (IDEALAB) under grant number 2012-1713, Framework Partnership Agreement number 2012-0025.

References


Lanyon, L., & Rose, M. L. (2009). Do the hands have it? The facilitation effects of arm and hand gesture on word retrieval in aphasia. Aphasiology, 23(7–8), 809–822. doi:10.1080/02687030802642044


Aphasiology


