

# Coherence and structure in aphasic and non-aphasic spoken discourse

A thesis presented

by

**Anastasia Linnik**

to the Faculty of Human Sciences

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

University of Potsdam

July 2016



university of  
 groningen



MACQUARIE  
University  
SYDNEY · AUSTRALIA



UNIVERSITY OF TRENTO - Italy



Newcastle  
University



Education and Culture

# Erasmus Mundus

This work was supported by the Erasmus Mundus Joint Doctoral Programme (EMJD) of the European Union ‘International Doctorate for Experimental Approaches to Language and Brain (IDEALAB)’ of the Universities of Groningen (NL), Newcastle (UK), Potsdam (DE), Trento (IT), and Macquarie University, Sydney (AU).



Supervisors:

Prof. Dr. Manfred Stede

Prof. Dr. Roelien Bastiaanse

Prof. Dr. Barbara Höhle

Oral exam date: March 31, 2017

Published online at the

Institutional Repository of the University of Potsdam:

<https://nbn-resolving.org/urn:nbn:de:kobv:517-opus4-423202>

<https://doi.org/10.25932/publishup-42320>

# Acknowledgements

Although I suspect I may never be able to thank them enough, I would like to express my deepest gratitude to my advisors, Professor Manfred Stede, Professor Roelien Bastiaanse, and Professor Barbara Höhle, who inspired, encouraged, and challenged me to develop and grow as a researcher. This work would not have been possible without their endless patience, countless hours of their time, and their generous well-placed guidance.

Next I want to thank the wonderful Hong Kong team, Professor Sam-Po Law, Dr. Anthony Kong, and Waisa Shum, who accepted me as one of their own and worked hard to help me make the best of my short tenure at HKU. Many thanks go to the IDEALAB professors who crafted the program and gave me the opportunity to join, to Professor Brendan Weeks for his thoughtfulness, and to Dr. Olga Dragoy for her infectious enthusiasm about neuroscience. Special thanks go to Professor Lyndsey Nickels for having provided much-needed perspective at regular intervals. I am also very grateful to Professor Gloria Olness who took an impressive amount of time to provide her invaluable expertise on language pathology and showed interest in my work.

My journey with IDEALAB has been one of the most exciting and incredible ones of my entire life. This is largely due to my fantastic friends, Leigh Fernandez and Katharina Nimz, who have been there for me from day one. Their enthusiasm (not only) about research, never-ceasing encouragement, and ability to have the most fun at all times are priceless to me. I had the pleasure of working and spending time with several excellent groups of people in a number of countries. Thanks to my brilliant colleagues and friends, Seçkin Arslan, Adrià Rofes, Vânia de Aguiar, Rui Qin, Jinxing Yue, and Tina Marusch, for creating such a fun environment to work in. Extra special thanks go to Djaina Satoer, Silvia Martínez-Ferreiro, and Laura Bos, who made it all seem easy and became very dear friends. Thanks also to Anja Papke, Lena Jäger, Alice Pomstra, Felix Golcher, Maja Stegenwallner-Schütz, and Yulia Akinina for their help and support.

I am most grateful to the people who, knowing little of the work I was doing, provided support in many different ways: Sebastian Golly for many a good thought and deed, Sasha Khakalo for being the caring Jewish mother I never had, Anna Endrihovskaia for the most lovely birthdays and much more, Max Müller for all the fun we had, Jenny (Miss Jay) Schellenberg for her inextinguishable passion for music, Christian Laakman for his endless hospitality, Henry Southern who is a blast to be around, and Mariya Khudyakova who is my friend first and only then a collaborator. Thank you to other friends who have cheered me on along the way: Yael Farhy, Kristina Lepetyukha, Anna Kareva, Maria Psurtseva, Maria Khachatryan, and Anna Khorosheva.

Special thanks must go to the people who spent a great deal of time around me as I wrote my thesis. Anna Tschapke, Matt Moroz, Philipp and Lisa Block, Jenny Schellenberg, and Alex Anania were the best housemates one can wish for. My colleagues at the Software Development Center at Interfax and at Retresco deserve many thanks for their support too (and also to the latter for the most amazing PhD hat).

Most of all, I would like to thank my brother Stanislav (Stasik) and my parents, Sergey and Tatiana, who are always on my team and who continually remind me of the rewards of dedication and perseverance. Finally, I want to thank Anselm Schultes for just being who he is.



# Contents

<b>Acknowledgements</b>	<b>iii</b>
<b>1 General Introduction</b>	<b>1</b>
1.1 Coherence . . . . .	1
1.1.1 Coherence and discourse structure . . . . .	2
1.1.2 Coherence as a perceived quality of discourse . . . . .	3
1.2 Aphasia . . . . .	3
1.2.1 Aphasia syndromes . . . . .	4
1.2.2 Variability in aphasia manifestations . . . . .	4
1.3 Cross-linguistic studies . . . . .	5
1.4 Aims and scope . . . . .	5
1.5 Outline of the dissertation . . . . .	8
<b>2 Discourse production in aphasia</b>	<b>11</b>
2.1 Motivation . . . . .	12
2.2 Defining concepts . . . . .	13
2.2.1 Informativeness . . . . .	14
2.2.2 Information structure . . . . .	16
2.2.3 Discourse structure . . . . .	17
2.2.4 Cohesion . . . . .	19
2.2.5 Coherence . . . . .	21
2.2.6 Relationship between coherence and cohesion . . . . .	25
2.2.7 Multi-level approaches . . . . .	25
2.2.8 Effectiveness and efficiency . . . . .	27
2.3 Methodological issues . . . . .	30
2.3.1 Sample size . . . . .	30
2.3.2 Genre . . . . .	31
2.3.3 Modality . . . . .	32
2.3.4 Cross-linguistic studies . . . . .	33
2.4 Conclusion . . . . .	33
<b>3 Evaluating coherence in anomic aphasia</b>	<b>35</b>
3.1 Introduction . . . . .	36
3.1.1 Coherence . . . . .	36
3.1.2 Coherence in aphasia: methods . . . . .	36
3.1.3 Coherence in aphasia: previous findings . . . . .	37
3.1.4 Different aspects of coherence . . . . .	37
3.1.5 Genre-related differences . . . . .	38
3.1.6 The present study . . . . .	38
3.2 Rhetorical Structure Theory . . . . .	38

3.3	Aims and research questions . . . . .	40
3.4	Method . . . . .	41
3.4.1	Data . . . . .	41
3.4.2	Data analysis . . . . .	42
3.4.3	Reliability of EDU segmentation and discourse structure (RST) annotation . . . . .	44
3.4.4	Analyzing discourse organization using RST . . . . .	44
3.4.5	Coherence ratings . . . . .	45
3.4.6	Analysis of information content . . . . .	45
3.4.7	Statistical analysis . . . . .	46
3.5	Results . . . . .	46
3.5.1	Analysis of coherence ratings . . . . .	46
3.5.2	Analysis of micro- and macrolinguistic (RST) variables . . . . .	47
3.5.3	Correlations between the linguistic variables and coherence ratings . . . . .	49
3.5.4	Information content . . . . .	51
3.6	Discussion . . . . .	51
<b>4</b>	<b>Linguistic mechanisms of coherence</b>	<b>55</b>
4.1	Introduction . . . . .	56
4.1.1	Coherence . . . . .	56
4.1.2	Coherence in aphasia . . . . .	56
4.1.3	Discourse structure . . . . .	58
4.1.4	The present study . . . . .	60
4.2	Methods . . . . .	61
4.2.1	Participants . . . . .	61
4.2.2	Materials . . . . .	61
4.2.3	Procedure . . . . .	62
4.3	Results . . . . .	67
4.3.1	Ratings . . . . .	67
4.3.2	Classification analysis with random forests . . . . .	68
4.4	Discussion . . . . .	71
<b>5</b>	<b>General discussion</b>	<b>75</b>
5.1	Major Findings . . . . .	75
5.1.1	Coherence deficit in aphasic discourse . . . . .	75
5.1.2	Discourse structure in aphasia . . . . .	76
5.1.3	On the multifaceted nature of coherence . . . . .	78
5.1.4	How is coherence established linguistically? . . . . .	79
5.2	General conclusions . . . . .	80
5.3	Further directions . . . . .	80
	<b>References</b>	<b>85</b>
	<b>Appendices</b>	<b>101</b>

# List of Figures

3.1	Example of the RST analysis in Cantonese . . . . .	39
3.2	Picture sequence “Refused umbrella” from the AphasiaBank protocol used as a stimulus in the story-telling task. . . . .	42
3.3	Distribution of RST relations in the (a) story-telling and (b) procedural description tasks. . . . .	49
4.1	Example of an RST tree with a multi-nuclear relation (Contrast), mononuclear relations (Concession, Consequence), and an embedded EDU (a restrictive relative clause) . . . . .	59
4.2	Example of discourse structures demonstrating different attachment points (A-B) and different constituents (A-C). . . . .	63
4.3	Random forest built-in feature selection plot. . . . .	69
4.4	Distribution of ratings in the NBD and PWA groups . . . . .	70
4.5	Distribution of Content ratings based on top 3 selected features . . . . .	70
4.6	The continuum of coherence ratings . . . . .	73



# List of Tables

1.1	Classical aphasia types in the Boston group system and their rough equivalents in Luria’s system (Akhutina, 2016; Bastiaanse, 2010) . . . . .	4
2.1	Studies addressing informativeness in aphasic discourse: methods and findings. . . . .	15
2.2	Studies addressing information structure in aphasic discourse: methods and findings. . . . .	17
2.3	Studies addressing discourse structure in aphasic discourse: methods and findings. . . . .	19
2.4	Studies addressing cohesion in aphasic discourse: methods, findings, and the relationship between cohesion and coherence. . . . .	20
2.5	Studies addressing coherence in aphasic discourse: definitions, methods, and findings. . . . .	23
2.6	Multi-level approaches to studying discourse coherence in aphasia. . . . .	26
2.7	Communicative effectiveness and efficiency in aphasic discourse. . . . .	28
3.1	Descriptive statistics of coherence ratings by naive listeners. . . . .	47
3.2	Comparisons of speaker groups in terms of perceptual judgment by naive listeners. . . . .	47
3.3	Comparisons of naive listeners’ judgment in the two genres by speaker group. . . . .	47
3.4	Descriptive statistics of macro- and microlinguistic measures . . . . .	48
3.5	Statistical comparisons between performances of aphasic and non-brain-damaged groups. . . . .	50
3.6	Correlation between macro- and micro-linguistic variables and subjective coherence ratings. . . . .	50
4.1	Results of the quantitative analysis of the discourse samples produced by aphasic (PWA) and non-brain-damaged (NBD) participants of the current study. Means (ranges). . . . .	61
4.2	Linguistic variables used for the analysis at the micro- and macrolinguistic levels. . . . .	66
4.3	Correlations between different coherence rating scales. . . . .	67
4.4	Results of the group and coherence ratings classification using random forests and different sets of predictors (variables). . . . .	68
4.5	Variable importance for each set of predictors. . . . .	69
4.6	Examples of cases problematic for classification into low/high coherence groups. . . . .	71



# Abbreviations

EDU – Elementary Discourse Unit

NBD – Non-Brain-Damaged People

PWA – People With Aphasia

RBD – People with Right Hemisphere Brain Damage

RST – Rhetorical Structure Theory



# Chapter 1

## General Introduction

The ability to produce connected understandable spoken discourse is vitally important in daily life. Discourse of different genres is a crucial part of a large variety of everyday communicative situations, from explaining how to make a cake to a friend or telling a family member about the events of the day, to sharing exciting stories of the past with a new acquaintance. People with aphasia, a language disorder resulting from focal brain damage, experience a multitude of language deficits that often deprives them of or limits their ability to participate in these and many more different activities. This dissertation focuses on the very property distinguishing discourse from a random or unstructured collection of sentences/utterances, called “coherence”, and the ability of people with aphasia to establish coherence in spoken discourse of different genres. The relationship between different linguistic impairments and coherence is explored through the analysis of formal structure in spoken discourse of aphasic and non-brain-damaged speakers of Russian and Cantonese, and the coherence perception by listeners.

### 1.1 Coherence

One must begin the discussion about coherence by addressing the existence of various definitions for this concept, which has been a prominent topic in a number of previous studies (e.g., Foltz, 2007; Kehler, 2002, 2004; Sanders & Spooren, 1999, ; cf. Section 2.2.5). Definitions and operationalizations have varied depending on the chosen approach – structural, functional, or cognitive (Armstrong, 2000), as well as the purpose of investigation. For example, in a functionalist investigation focusing on communicative success, coherence could be defined as the linguistic component of effectiveness and evaluated based on the amount of information successfully transferred in an interactive setting (e.g., Glosser & Deser, 1991; Halliday & Hasan, 1976; Ulatowska, Olness, Samson, Keebler, & Goins, 2004). From a structuralist point of view, which considers discourse a unit above sentence (e.g., Grimes, 1975; Hobbs, 1985; Mann & Thompson, 1985), however, coherence could refer to the well-formedness or integrity of this unit and be evaluated based on the number of missing elements or links between parts of a discourse or gaps in discourse structure.

Further in this and the following chapters, I will expand on the ways in which the concept of coherence is defined and operationalized. The difficulty with finding a single definition for this concept results largely from the fact that different, albeit related, phenomena have been referred to as coherence. In this dissertation, it is argued that coherence is not a unitary notion, but a multifaceted construct, comprised of a combination of these phenomena, and co-created in an interaction between a speaker and a listener. As such, coherence is perhaps the most inherent property of discourse, indispensable for

communication. Consequently, it has been extensively studied in theoretical and computational linguistics (e.g., Gernsbacher & Givón, 1995; Hovy, 1993; Kintsch & Van Dijk, 1978; Marcu, 1997; Van Dijk, 1977). However, the means by which it is achieved are still not entirely understood.

### 1.1.1 Coherence and discourse structure

Discourse is considered to have two dimensions: a local, or micro-structural dimension, associated with local connectivity between pairs of clauses or sentences and cohesion, and a global, or macrostructural dimension, where overall semantic unity is established (see Van Dijk, 1980, for a discussion). These two dimensions comprise the phenomena that transform a series of utterances or sentences into connected speech or text. It has been suggested that this semantic wholeness and connectedness of discourse – its coherence – is established through its internal organization, specifically, through the construction of discourse structure. A number of theories have stemmed from this idea, suggesting, among others, tree (Mann & Thompson, 1985, 1988), graph (Wolf & Gibson, 2005), stack (Grosz & Sidner, 1986), or underspecified representations (Reitter & Stede, 2003; Stede, 2004) of discourse structure. These formal approaches are based on the assumption that discourse consists of basic meaningful elements that are connected by semantic relations (also called “rhetorical”, “coherence”, or “discourse” relations). Although communicative intentions do not define relations per se, it is often true that there is an intention behind a relation (Taboada & Mann, 2006b). The main goal of relations is to connect spans of discourse to each other, building up the discourse as a semantic entity, of which every part makes sense and has its place and reason to be.

While there is a variety of theories and studies addressing the means through which coherence is established, its breakdown is not as well-studied. In few studies, elements of structural organization have been investigated through such constructs as story grammar (Coelho, Liles, Duffy, Clarkson, & Elia, 1994) and superstructure (Olness & Ulatowska, 2011; Ulatowska, Doyel, Stern, Haynes, & North, 1983; Ulatowska, North, & Macaluso-Haynes, 1981). While these approaches focused on the appropriateness of information content in a discourse within a specific genre and/or with respect to a selected topic, none of them considered relations between its different parts. Consider, for example, a situation in which a speaker strays from the expected story line to reminisce on an personal memory. The points of interest in this case would be whether the deviation is connected to the initial story, whether the speaker can make his/her way back to the initial main story line or not, and how abruptly or logically they could do it. All of these points are related not just to the amount of information conveyed, but to the way it is structured. The relations established between different parts of the discourse, in this case, will be what sets apart a coherent story and a disjointed collection of information. Consider the following example from the study of Wright and Capilouto (2012), in which a participant is asked to recount a previous weekend’s activities. It is suggested that answer (a) would be an example of good maintenance of coherence, whereas answer (b) would be an example of the opposite, as it is not related to the topic of the conversation. However, if a continuation (c) is added to answer (b), it can be seen that (b) is connected to (c), and together they represent a reasonable answer to (a).

(a) *“Last weekend we went to the shore.”*

(b) *“You don’t get much when you are retired.”*

(c) *“So we usually just stay at home and sometimes go out to watch a movie or have breakfast at a local café”*

It must be added that genre plays an important role when considering discourse structure. Building blocks of discourse are not necessarily the same in different genres, and the role of relations between them may not be as crucial for coherence and communicative success in some of them as it is in others (e.g., Olness, 2006). For example, semantic and pragmatic relations between events play a significantly more prominent role in a narrative than in a procedural discourse.

Coherence impairments have been observed in people with certain medical conditions, such as Alzheimer’s disease (Chapman, Ulatowska, King, Johnson, & McIntire, 1995; Glosser & Deser, 1991; Ripich & Terrell, 1988) or schizophrenia (Barch & Berenbaum, 1997; Ditman & Kuperberg, 2007, 2010). Perhaps the largest language-impaired group with relatively intact cognitive abilities is people with aphasia. However, findings on structural aspects of discourse production in aphasia are scarce. In this dissertation, a structural component is integrated into the investigation of the multifaceted phenomenon of coherence through the use of a relational approach to discourse structure.

### 1.1.2 Coherence as a perceived quality of discourse

Armstrong (2000) suggested that a structural, or formal, approach should be combined with elements of functional and cognitive approaches to study discourse. In this dissertation, discourse is regarded not just as a linguistic entity – the result of semantic organization of smaller meaningful units of language – but as a product of the speaker-listener interaction. Coherence is a property that emerges during both speech production **and** comprehension, allowing a listener/reader to reconstruct discourse as a speaker/writer had it in mind, or the mental representation of it (Gernsbacher & Givón, 1995). The aforementioned controversial reports on macrolinguistic processing in people with aphasia stem from the fact that aphasic speakers even with quite severe linguistic impairments often manage to transfer the message. This observation is a viable counter-example of what is usually presumed about the interactions between different linguistic levels. Namely, it is not obvious anymore whether intact microlinguistic abilities are a prerequisite for maintaining the macrostructure of discourse. As a consequence, it raises the question about the extent to which linguistic competence is indispensable for a successful communication (e.g., Armstrong & Ferguson, 2010b; Olness & Ulatowska, 2011). For example, nonverbal strategies, such as gesturing, can help to a certain extent to compensate language deficits in aphasia, just like beginner level second-language learners are often able to make themselves understood using the most basic vocabulary and almost no grammar. Although it is too rough to assume that regular daily communication is possible without language, the degree to which people with aphasia rely on the collaborative construction of meaning in their discourse production is an important issue. This dissertation addresses the connection between linguistic properties of discourse, and their relationship with the perception of coherence by listeners. It is thereby the initial step towards a better understanding of coherence as a phenomenon containing a pragmatic component, potentially more evident and significant in aphasic discourse production.

## 1.2 Aphasia

Aphasia is a language disorder due to brain damage. It is characterized by a sudden onset, usually a stroke or traumatic brain injury, but it can also be caused by a brain tumor or brain infection. Only when the brain damage arises after language has been acquired (around 12 years old), the term “aphasia” is used. Characteristic manifestations of aphasia differ depending on the site and size of the lesion.

### 1.2.1 Aphasia syndromes

Based on speech output, the two major subgroups of aphasias are non-fluent and fluent aphasias. Fluent aphasia is characterized by a relatively fluent speech, that is, a typical number of words per minute, with word-finding difficulties, whereas in non-fluent aphasia speech production is effortful and reduced, and is often grammatically simple (agrammatism). More detailed classifications of aphasia symptoms have been proposed, distinguishing different types of aphasia based on linguistic deficits experienced by patients. Two examples are Luria's (1966b) classification used by Russian aphasiologists, and variations of the scheme developed by the Boston group (Geschwind, Goodglass, Kaplan, Benson, Alexander, etc.; e.g., Benson & Geschwind, 1971) commonly accepted in the Western world Table 1.1. As the coverage of historical background of aphasia classification is out of the scope of this thesis, interested readers are referred to one of the more relevant works (e.g., Benson & Ardila, 1996; Eling, 1994; Tesak & Code, 2008). However, pertinent to the discussion unfolding in this dissertation is the fact that aphasia can be disruptive at various linguistic levels, including phonological, lexical, morphological, syntactic, and semantic processing disturbances. For example, word-finding problems in Wernicke's and milder anomic aphasia result from a deficit, in which the connection between the word meaning and the phonological form of words is disrupted. In contrast, in Broca's aphasia speech production is impeded by the inability to compose "the articulatory schema of a word" (Akhutina, 2016, p. 4) combined with a grammatical deficit, leading to various phonemic and morphosyntactic simplifications and errors. In this dissertation, these word- and sentence- level deficits will be referred to as microlinguistic, as opposed to macrolinguistic impairments, which occur on the level of discourse and pragmatic processing.

Table 1.1: Classical aphasia types in the Boston group system and their rough equivalents in Luria's system (Akhutina, 2016; Bastiaanse, 2010)

Luria, 1966	Boston group	Fluency	Comprehension	Repetition	Characteristics
Efferent motor	Broca's	-	+	-	Morphosyntactic deficits, "telegraphic speech"
Sensory	Wernicke's	+	-	-	Semantic and / or phonemic paraphasias
Acoustic-mnemonic	Anomic	+	+	+	Word-finding difficulties
Afferent motor	Conduction	+	+	-	Phonemic paraphasias
-	Transcortical sensory	+	-	+	Repetition relatively spared
Dynamic	Transcortical motor	-	+	+	Repetition relatively spared
-	Global	-	-	-	All modalities impaired

Note: '+' == relatively spared, '-' == relatively impaired

### 1.2.2 Variability in aphasia manifestations

Studies on reliability of aphasia classification also highlighted a number of problems, such as poor agreement between standardized assessment batteries, as well as between the batteries and clinical impression of aphasia types (Swindell, Holland, & Fromm,

1984; Wertz, Deal, & Robinson, 1984). It is noteworthy that, independently of the classification, individual variability in patients with aphasia is very common, and most aphasias are of the “mixed” type (Bastiaanse, 2010; Prins & Bastiaanse, 2004). The variety of aphasia symptoms represent certain difficulties for research as well as for clinical practice, since they can obscure the outcomes, and limit the possibility of generalization. Nonetheless, in the studies described in this dissertation, the existing variability can be ignored to a certain extent. Specifically, due to the diversity of linguistic deficits, aphasic speech provides a palette of micro- and macrolinguistic breakdowns, and potentially a spectrum of options between “coherent” and “incoherent” discourse. Hence, variations in aphasic speech production provide a unique opportunity to study the processes underlying language production.

### 1.3 Cross-linguistic studies of discourse production

In Chapter 2, a number of reasons to study discourse processing in aphasia crosslinguistically are discussed. In this section, the focus will be on two major reasons for the current study to be cross-linguistic in nature. First, the chosen languages – Cantonese and Russian – are typologically different. While one is analytic, the other one is synthetic, in other words, some of the meanings that are encoded morphologically in Russian are expressed through word order, context, adverbs, and particles in Cantonese. Morphosyntactic processing is often impaired in aphasic speakers (e.g., Edwards & Bastiaanse, 1998; Glosser & Deser, 1991; Prins & Bastiaanse, 2004; Thompson et al., 2013). Thus, grammatical differences between Cantonese and Russian can shed light on the relationship between micro- ad macrolinguistic processes, as well as their connection to coherence and its maintenance in aphasia.

Second, although formal approaches to discourse structure have been widely used with different languages in discourse studies, this is not the case in the research on discourse of people with language impairments. Apart from the Cantonese and the Russian studies presented in this dissertation, the only other example is the analysis of dream stories told by children with neurotic disorders by Kibrik and Podlesskaya (2009). In this dissertation, the applicability of such an approach (Rhetorical Structure Theory; Mann & Thompson, 1988) to spoken aphasic discourse in typologically different languages is explored. The outcomes can be insightful both for the understanding of discourse abilities in aphasia, and for the validation and/or adjustment of a prominent theoretical and methodological framework for studying discourse production.

### 1.4 Purposes of the current research project

The current research project has the following purposes.

1. Purpose #1. The term coherence received a number of different definitions in previous studies, due mainly to the fact that it has been used to refer to different concepts, including, among others, informativeness, the “hanging together” of a discourse, or its thematic unity. The reported results have, as a result, often not been directly comparable and/or did not converge. Before new issues are raised, pieces of evidence from earlier studies should be assembled together in an attempt to construct a clear and comprehensive picture of discourse production in aphasia. The first purpose of this dissertation is, thus, to gather and systematize the existing knowledge on discourse coherence and related phenomena in aphasia.

2. Purpose #2. The second purpose of this dissertation is to find out whether discourse of people with aphasia is constructed in a similar way to that of non-brain-damaged speakers, and if that is the case, what effects language impairments have on discourse structure. The concept of discourse structure is introduced as an operational facet of coherence (cf. Section 1.1.1). The previously expressed idea that coherence is attained through the establishment of relations between elements of discourse, or a discourse structure, is thus being explored in the context of language impairments in aphasia. The formal discourse analysis framework used in this project (RST) is by no means the only available operationalization of the concept of discourse structure (e.g., Coelho et al., 1994; Ulatowska et al., 1981) and does not presume to cover all of its aspects. Rather, it is one specific operationalization with its limitations, which has been extensively used in previous research on healthy discourse. First attempts to identify discourse structure in aphasic spoken samples were made based on the evidence from two typologically different languages to verify cross-linguistic generalization of the findings. The assumption of a strong connection between discourse structure and coherence implies that, if present, discourse structure impairments could be responsible for decreased coherence in aphasic discourse. However, rather than considering the relationship between discourse structure and coherence in isolation, it is integrated into the investigation of different aspects of coherence along with a number of other linguistic variables.
3. Purpose #3. In this dissertation, coherence is assumed to be a multifaceted construct. The third purpose is thereby to define which intrinsic qualities of discourse it is comprised of and identify the facets of coherence that are impaired in aphasic discourse.

Informativeness, understandability, connectedness, and clarity are four main correlates of coherence identified based on the review in Chapter 2 and considerations from research on unimpaired discourse (e.g., Gernsbacher & Givón, 1995; Mann & Thompson, 1985; Van Dijk & Kintsch, 1983). *Informativeness* has been linked to coherence in a number of previous studies (e.g., Christiansen, 1995; Coelho et al., 1994; Olness, 2006; Ulatowska et al., 1981) and refers to the amount of relevant information content in a discourse. *Understandability* is a measure of how well the meaning intended by the speaker could be perceived and/or interpreted by the listener. It reflects the idea expressed earlier in this chapter that coherence is co-constructed by a speaker and a listener/reader in an interaction (Section 1.1.2). A discourse can be understandable in spite of partial omissions of information content, for example, owing to pragmatic context (e.g., common ground), and vice versa, an objectively informative discourse is not necessarily understandable. Two additional aspects that play a role in the latter case are connectedness and clarity.

*Connectedness* has been partly explored in several studies and encompasses the relationship between consecutive elements of discourse (e.g., Glosser & Deser, 1991; Wright, Koutsoftas, Fergadiotis, & Capilouto, 2010) and appropriate organization of larger blocks of discourse into a unified whole through the use of semantic and pragmatic relations. To avoid confusion, it should be mentioned that in the study of Ulatowska et al. (1981), clarity was tentatively identified with discourse cohesion. Cohesion is commonly defined as local connectivity achieved explicitly by lexical and grammatical means (cf. Section 2.2.4), while implicit semantic connectedness between smaller elements of discourse is referred to as local coherence. In this project, the term connectedness is used to refer to the aspect of coherence associated with lower- as well as higher-level connectivity between elements of discourse. The notion of

connectedness brings about the idea of local and global structural connectivity briefly addressed in Section 1.1.1.

The term *clarity* is then redefined to describe the overall meaningfulness of discourse achieved through the appropriateness of its elements with respect to the overall topic of discourse or the dynamic of a given interaction. Presuming the presence of a minimal adequate information content, clarity with which this content is delivered is an essential component for semantic unity, or “the degree to which this unit [discourse], as a whole, hangs together or makes sense” (Olness & Ulatowska, 2011, p. 1397). It could be defined as the conformity of a discourse to a specific logical progression or a set of rules expected within a particular genre. For example, long or unrelated deviations from the topic in a conversation or incorrect temporal order of narrative episodes could decrease discourse clarity. Clarity has been addressed as part of such concepts as narrative superstructure (e.g., Ulatowska, Allard, & Chapman, 1990) and story grammar (Coelho et al., 1994), or else evaluated through the occurrence of elements disrupting discourse progression and meaningfulness (e.g., Andreetta, Cantagallo, & Marini, 2012; Christiansen, 1995; Glosser & Deser, 1991; Marini, Andreetta, Del Tin, & Carlomagno, 2011; Sherratt & Penn, 1990; Sherratt & Bryan, 2012; Wright et al., 2010).

In this dissertation, it is assumed that coherence can be described and evaluated in terms of the four aforementioned components or parameters. They address the amount of information (informativeness), appropriateness of content and its progression with respect to the topic of discourse or within a given interactive setting (clarity), connectivity between its elements, or its structural organization (connectedness), and its pragmatic adequacy or sufficiency (understandability). It is assumed that subjective ratings of the first three qualitative parameters should be roughly in line with the corresponding quantitative linguistic measures – for example, certain discourse structure parameters for connectivity and clarity, a measure of information content for informativeness. Understandability, however, is presumed to capture the contribution of a listener to the establishment of coherence, or the pragmatic context with such resources as common ground and world knowledge. This approach is a working operationalization of coherence which allows for further elaboration and/or reevaluation. The investigation of the relationship between these aspects of discourse and a number of micro- and macrolinguistic measures is the first step towards a better understanding of the nature and mechanisms of coherence.

4. Purpose #4. The fourth and final purpose of this project is thereby to shed light on how coherence is established linguistically. Various micro- and macrolinguistic variables are measured based on aphasic and unimpaired speech samples and factored in independently of group (e.g., high vs. low number of information content units, morphosyntactic errors, certain types of discourse relations etc.). The interactions between these variables and coherence ratings are then considered in order to identify linguistic variables potentially contributing to coherence and/or responsible for its impairment. Two features of this project should be highlighted in relation to this point. Firstly, two languages with fairly different morphosyntactic organization are considered – Russian and Cantonese. While the utilized discourse structure framework is presumed to be language-independent, it has not been extensively used with language-impaired speech and the effects of linguistic deficits on discourse structure may vary in different languages. Hence, the cross-linguistic design is especially valuable when drawing conclusions about the relationship between micro- and macrolinguistic variables and coherence (cf. Section 1.3). It must be noted that the

Russian study (Chapter 4) does not directly replicate the Cantonese study (Chapter 3) and the findings are thus not subject to direct comparison. However, similar trends are expected in the two studies, since coherence and discourse structure are language-independent phenomena.

Secondly, discourse samples of speakers with different aphasia profiles were considered. Whereas within-group variability in aphasia profiles is often a hindrance when it comes to drawing inferences, it is advantageous in this project, since the effects of a wider range of deficits on coherence can be observed and taken into account (cf. Section 1.2.1). This dissertation does not attempt to make conclusions about coherence impairments in specific types of aphasia, but rather focuses on the effects different types of linguistic impairments observed in aphasia affect coherence perception.

While it is explicitly stated here that one of the purposes of this dissertation is to investigate linguistic mechanisms of coherence, it should be stressed once again that one of the main assumptions of this dissertation is that “coherence is achieved through the integration of multiple semantic resources, e.g., information expressed in the text through a variety of linguistic and paralinguistic tools; knowledge of the world and the predictable ways in which discourse is structured; and the relevant characteristics of the context in which the discourse is being produced” (Olness & Ulatowska, 2011, p. 1397) and that language is just one of available means.

To summarize, this dissertation focuses on consolidating the existing knowledge on coherence (Chapter 2) and exploring further the multifaceted nature of coherence, linguistic mechanisms involved, and the contribution of discourse structure (Chapters 3 and 4). Findings of this project could therefore have implications for aphasia therapy, as they contain information on the ability of people with aphasia to communicate their thoughts in a clear connected manner. They could also provide insights on the communication success and conversation dynamics of people with aphasia, since in the current project, coherence is operationalized as a product of interaction between a speaker and a listener. In a more general sense, advancing our understanding of this most intrinsic property of discourse can shed light on the general mechanisms underlying connected language production and potentially, albeit not directly, also language comprehension.

## 1.5 Outline of the dissertation

The dissertation is structured as follows. In **Chapter 2**, the body of work on discourse production in aphasia is addressed. The chapter systematically reviews the key issues, constructs, and themes in the literatures on discourse aphasiology, constructing a springboard for further research in the field. Despite the substantial amount of attention coherence in aphasia received in the recent past, the findings on the matter to date often lack comparability due to theoretical, terminological and methodological differences between the existing studies. Chapter 2 demonstrates how these differences have affected research on coherence and a number of related discourse-level concepts.

Consider, for example, two studies by Ulatowska et al. (1981) and by Christiansen (1995) that both investigated the aspect coherence referred to as *clarity* in this dissertation. The first study (Ulatowska et al., 1981), in which coherence is operationalized as well-formedness of a discourse in terms of plausibility, conventionality, and conclusiveness, demonstrated that narratives of aphasic speakers were similar to the discourse of non-impaired speakers in terms of logical progression and conventional episodic structure.

In the second study (Christiansen, 1995), coherence was defined as semantic connectiveness at the propositional level and assessed in terms of relevance, progression, and completeness of propositional content. The results indicated that speakers with aphasia produced more coherence violations (e.g., relevance and progression violations, information gaps) than the control group. Despite relatively similar definitions, substantially differing methodologies made these results complementary, but not directly comparable. In the study by Ulatowska et al. (1981), coherence was also rated on a scale, and aphasic narratives received lower ratings than those of control speakers. In the rating study by Glosser and Deser (1991), however, no difference was observed between control and aphasic groups. In this case, the results are not comparable due to the fact that coherence was operationalized differently in the two studies, and its different aspects were rated. In this dissertation, it is argued that terminological and methodological differences are triggered largely by the fact that coherence is often presumed to be a unitary construct, whereas it is a complex multifaceted phenomenon.

Chapter 2 addresses a number of other discourse-level concepts in relation to coherence and potentially relevant for its different aspects. Informativeness, information structure, discourse structure, cohesion, effectiveness, and efficiency are the parameters most commonly used to characterize spoken discourse production and evaluate its impairments in aphasia. Each of them has been previously linked to the concept of coherence. For example, *information content* has been measured in several investigations of coherence under the assumption that discourse cannot be coherent unless its information content is sufficient (Andreetta et al., 2012; Capilouto, Wright, & Wagovich, 2006; Christiansen, 1995; Ulatowska et al., 1983; Ulatowska et al., 1981). Studies on functional communication are not centered around linguistic performance alone, as it has been observed that people with language impairments are often able to maintain successful communication with help of non-linguistic means. One of the key parameters commonly used to assess communication success is *effectiveness*, defined as the ability to produce sufficiently meaningful and understandable discourse in order to reach a communicative goal. One of the aspects of coherence considered in this dissertation referred to as *understandability* is a reflection of communicative effectiveness. The linguistic component of effectiveness is commonly assessed using, once again, informativeness-based measures, such as correct number of main ideas transferred between a person with aphasia and a non-aphasic partner in a conversation (Ramsberger & Rende, 2002), or content unit analysis, in which a content unit is a word or a phrase expressing a single unit of information (Yorkston & Beukelman, 1980). Going one step further, *efficiency*, which reflects not just how meaningful discourse is, but also how fluently speakers are able to produce it, is measured in information units (e.g., Nicholas & Brookshire, 1993). An adapted measure of efficiency is included in the investigation of coherence in this dissertation (cf. Chapter 3).

Information organization has also been investigated as a factor influencing discourse coherence. The relationship between *discourse structure* and coherence has been briefly discussed in Section 1.1.1 and will remain a prominent topic of this dissertation, as one of the operationalizations of the concept of coherence. Less plain is the role of *information structure*, which is concerned with correct “selection and packaging” of information in a sentence (Chafe, 1976), for example, the ability of speakers to assign prominence to certain parts of discourse. It has been reported that aphasic speakers are capable of using some aspects of information structure (e.g., Olness, Matteson, & Stewart, 2010) and their difficulties with other aspects may be a result of syntactic deficits (Bastiaanse, Koekoek, & van Zonneveld, 2003; Ulatowska et al., 1990). While it is clear that information structure is often essential for correct interpretation of a linguistic expression by listeners, its effect on coherence is unclear. Similarly, the relationship between coherence and

cohesion, which stands for semantic connectedness between elements of discourse reached through lexical and grammatic means (Halliday & Hasan, 1976), is not yet clear and requires further investigation. While the link between coherence and cohesion has been a matter of an extensive debate in healthy discourse analysis, it has also been demonstrated that the observed reduction of cohesion in aphasia is a result of impaired microlinguistic processes (cf. Section 2.2.4 for the discussion). These include lexical retrieval (Glosser & Deser, 1991) and syntactic deficits, due to which aphasic speakers tend to use simplified syntactic constructions (e.g., Akhutina, 1982; Edwards & Bastiaanse, 1998; Penn, 1988). While no assessment of information structure and cohesion were included into the studies presented in this dissertation, the relevant research is addressed in Chapter 2.

The interaction between different discourse-level phenomena has been considered in very few studies to date (e.g., Andretta et al., 2012; Marini et al., 2011; Wright & Capilouto, 2012) and little effort has been devoted to making outcomes of various available measures comparable. Bringing together relevant pieces of already existing information is an essential first step that has to be taken before new evidence is collected and interpreted. Hence, the review of different findings and methods presented in Chapter 2 is a necessary starting point and Purpose #1 (cf. Section 1.4) of this project.

**Chapter 3** introduces a study of coherence maintenance in Cantonese speakers with aphasia, in which coherence is explored as a perceived quality of discourse, evaluated by naive listeners. In this chapter, a hypothesis that originated from and was based on the studies of discourse in non-brain-damaged speakers is tested. According to this hypothesis, internal organization of discourse contributes to coherence along with, and potentially more than, its word and sentence level well-formedness. Using Rhetorical Structure Theory (RST; Mann & Thompson, 1985, 1988), a theory of the analysis of discourse organization new to aphasiology, the ability of aphasic speakers to build discourse structure and its effect on coherence is explored.

In **Chapter 4**, the interaction between different linguistic levels involved in discourse production and trained raters' perception of coherence is investigated through the analysis of semi-spontaneous narratives produced by Russian speakers with and without aphasia. In the study described Chapter 4, relatively elaborate and complex discourse samples were collected through video-induced story-telling, as opposed to procedural discourse and picture-elicited narratives in Chapter 3. Classification analysis was used to explain coherence ratings assigned by raters based on a number of micro- and macrolinguistic parameters of these samples. The outcomes of the classification analysis and the evaluation of individual variables' contribution to its accuracy are interpreted in terms of the contribution of different linguistic variables to the raters' perception of coherence.

**Chapter 5** contains the general discussion of all findings and the conclusions of the dissertation. The limitations and the some of the main challenges of this work, as well as the implications and further directions of this research are also addressed in this chapter.

## Chapter 2

# Discourse production in aphasia: a current review of theoretical and methodological challenges <sup>1</sup>

### Abstract

*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 fifteen years. Nevertheless, many questions still do not have definitive answers.

*Aims:* The aim of this paper 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 initiated by Armstrong (2000), 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.

---

<sup>1</sup>This chapter was adapted from: Linnik, A., Bastiaanse, R., & Höhle, B. (2016). Discourse production in aphasia: a current review of theoretical and methodological challenges. *Aphasiology*, 30(7), 765–800. doi:10.1080/02687038.2015.1113489

## 2.1 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 organization 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 to 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 emphasized 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 organization 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 systematically 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 multi-level approaches addressing the connection between them, and overall communicative effectiveness and efficiency. This 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 cross-linguistic 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 depends 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, 1978) to aphasic data, as well as a group of standardized 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, 2011) was dedicated to various topics in aphasic discourse research, namely, discourse level treatment techniques (Boyle, 2011; Kempler & Goral, 2011), new methodological developments (Fergadiotis, Kapantzoglou, & Wright, 2011; Marini et al., 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 toward potential refinements of study design and methodology.

## 2.2 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 et al. (1990, 1983, 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; Fergadiotis & Wright, 2011; Wright, 2011), such as production of excessive irrelevant proposition content, reduced efficiency, and low lexical informativeness (Andreetta, 2014; Andreetta & Marini, 2015; Christiansen, 1995; Nicholas & Brookshire, 1995).

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 people with aphasia despite the microstructural linguistic difficulties. However, several researchers have noted that people with aphasia 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).

### 2.2.1 Informativeness

Discourse is a flow of information put into words, organized 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 incomprehensibility of aphasic discourse (e.g., Andretta et al., 2012; Capilouto et al., 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 2.1). 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 story-telling, 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 emplotment, 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” (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 (Andretta, 2014; Andretta et al., 2012; Andretta & Marini, 2015). The information on the design and results of the above mentioned studies, along with several other works addressing informativeness in aphasic discourse, is summarized in Table 2.1.

There are a few challenges with the informativeness measurements that have been used in previous studies. Firstly, our understanding of informativeness in discourse may be constrained by the limitations of the current methods. Namely, for most of the men-

tioned 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 naturally occurring conversation (Oelschlaeger & Thorne, 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, Goda, and Spencer (1995) showed that lexical information is higher for conversational discourse, but thematic information content cannot be estimated based on the principles identified for semi-spontaneous narratives.

Table 2.1: Studies addressing informativeness in aphasic discourse: methods and findings.

Investigator(s)	Case/Multiple case/Group	Elicitation task	Units	Impaired?
Yorkston and Beukelman (1980)	50 PWA, 78 NBD	Single picture description	Content Units	No, in mild and high-moderate aphasia; Yes, in more severe aphasia
Ulatowska, Allard, and Chapman (1990)	30 PWA (10 mild, 15 moderate, 5 severe)	Personal narrative, fable retelling, picture series description, procedural discourse	Number of propositions (events), number of steps in the procedures	No, in mild and moderate aphasia; Yes, in severe aphasia;
Nicholas and Brookshire (1993)	20 PWA, 20 NBD	Single pictures, picture series, personal narratives, procedural discourse	Correct Information Units (CIU)	Yes
Doyle, Goda, and Spencer (1995), Doyle et al. (2000)	20 PWA, 15 PWA	Conversation; Story Retelling Procedure (SPR)	CIU	No
McNeil, Doyle, Fossett, Park, and Goda (2001)	15 PWA, 31 NBD	SPR (Doyle et al., 2000)	Percentage of Information Units	Yes
Capilouto, Wright, and Wagovich (2006), Wright, Capilouto, Wagovich, Cranfill, and Davis (2005)	40 healthy adults; 8 PWA, 8 NBD	Single and sequential picture description	Main events	Yes
Andreetta (2014), Andreetta, Cantagallo, and Marini (2012), Andreetta and Marini (2015), Marini, Andreetta, Del Tin, and Carlomagno (2011), Marini, Boewe, Caltagirone, and Carlomagno (2005)	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	Single and sequential picture description	Thematic Units, Lexical Information Units (%LIU)	No for thematic informativeness; Yes for lexical information content

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

Secondly, 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 paper, and refer to the level of discourse organization at the macro-level, whereas lexical informativeness is a micro-linguistic 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.

### 2.2.2 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, although IS has received considerable attention in healthy discourse analysis (Chafe, 1976; Chomsky, 1971; Féry & Krifka, 2008; Gundel, 2003; Lambrecht, 1994; Roberts, 2012; Stede, 2004, , i.a.).

The central concepts of information structure are topic, comment, focus, and givenness (see Krifka, 2008; von Heusinger, Klaus, 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 & Fretheim, 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 & Fretheim, 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., 1989). Later insensitivity of adults with aphasia to the given-new organization 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 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 non-language-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 2.2 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 people with aphasia 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 et al., 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 micro-linguistic difficulties or not.

Table 2.2: Studies addressing information structure in aphasic discourse: methods and findings.

Investigator(s)	Group/ multiple case/case	Elicitation task	Method	Impaired?
Bates et al. (1983); Wulfeck et al. (1989)	10 PWAs (5 Broca, 5 Wernicke); 5-10 PWA, native English, Italian, and German speakers	Picture series description; Biographical interview	Analysis of the devices used to express topic/focus distinction: lexicalization vs. ellipsis, pronominalization, def./indef. articles, word order it dative items, conjunctions and connective adverbs and adjectives	No
Early and VanDemark (1985)	10 PWA, 10 NBD	Picture series description	Analysis of the use of definite/indefinite markers to identify given/new information	Possibly (demonstrated pragmatic competence part of the time)
Bastiaanse et al. (2003)	8 PWA (Broca), 6 NBD	Sentence completion (in response to pictures)	Investigation of the production of constructions with moved objects (object scrambling)	Yes, but due to a syntactic rather than a pragmatic deficit
Olness, Matteson, and Stewart (2010)	17 PWA, 16 NBD	Personal narrative (frightening experience)	Analysis of narrative evaluative devices, modalising function of language in narratives	No

*Note: PWA == people with aphasia, NBD == non-brain-damaged people*

### 2.2.3 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 (cf. Cameron, 2001; Fox, 1987; Schiffrin, 1994; Van Dijk, 1997, for a discussion on the definition of discourse). Thoughts, ideas, and

information transferred through discourse are not chaotic, but organized, which lead to the idea of discourse being internally structured. Thus, the term “discourse structure” in this paper refers to the internal organization of discourse into a coherent whole. Discourse is commonly considered to have two dimensions – local and global, also referred to as micro- and macro-structure respectively (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 organization at the macro-level, and it follows that there is a corresponding possibility that no one common aspect of macro-organization 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 organization of global meaning through the use of narrative elements (e.g., setting, evaluation, coda) he then called “superstructure”. These definitions of macro- and superstructure were used in several studies investigating discourse in aphasia (Ulatowska et al., 1990; Ulatowska & Chapman, 1994; Ulatowska et al., 1983; Ulatowska et al., 1981; Ulatowska & Sadowska, 1992). Their results showed that reduction in information content and poor distribution of information disrupt 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 suprasentential 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 et al., 2010; Olness & Ulatowska, 2011). However, a few studies addressing discourse organization in terms of propositional content provided evidence against the preservation of this aspect of discourse organization in aphasia (e.g., Christiansen, 1995; Huber, 1990). Another discourse production macro-phenomenon 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 experience difficulties at different levels of manipulation of discourse structure, in particular, with abstraction and generalization (Ulatowska & Chapman, 1994; Ulatowska, Chapman, Johnson, & Branch, 1999). A brief overview of the methods and results of the studies investigating different aspects of aphasic discourse organization at the macro-level discussed in this section are presented in Table 2.3.

Different theoretical approaches to the investigation of discourse structure used in the existing literature provided insights on various aspects of discourse macro-structure. Bringing these different perspectives together can potentially lead to a more in-depth, multifaceted understanding of language processing in aphasia at the macro-level. 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 organization start affecting global connectedness and meaning of discourse, and which aspects of discourse structuring are involved in the process.

Table 2.3: Studies addressing discourse structure in aphasic discourse: methods and findings.

Investigator(s)	Group/ multiple case/case	Elicitation task	Method	Impaired?
Ulatowska et al. (1990, 1983, 1981); Ulatowska and Sadowska (1992)	Groups (5-15 PWA), several cases	Personal narratives, picture description, story retelling	Analysis of relevant/irrelevant content, distribution of information within superstructure components, distribution of important information and detail	No
Huber (1990)	72 PWA (global, Wernicke, Broca, amnesic), examples for 6	Picture description (cartoon)	Evaluation of essential/optional propositional content	Yes
Glosser and Deser (1991)	9 PWA (fluent)	Personal narratives	Associated thematic organization with global coherence and rated every verbalization's appropriateness with respect to the general topic of conversation	No
Ulatowska and Chapman (1994), Ulatowska, Chapman, Johnson, and Branch (1999)	4 case studies; 15 PWA, 15 NBD	3 Aesop's fables (written), Spontaneous speech, expository discourse; 2 picture stories, 6 verbal fables	Analysis of the ability to manipulate macrostructure, namely provide the main idea/gist and the moral of the story; Analysis of linguistic means of signaling macro-structure (e.g., use of connectors, reference)	Restricted in some aspects of macrostructure; Yes on the moral task, No on the gist task
Christiansen (1995)	15 PWA	Picture series description (4 cartoon stories)	Propositional content analysis (events, states, elaborations, comments)	Yes
Olness and Ulatowska (2011), Ulatowska and Olness (2007)	Single case examples	Personal narratives	Analysis of the event line, temporal organization, evaluative devices	No

*Note: PWA == people with aphasia, NBD == non-brain-damaged people, RBD == people with right hemisphere brain damage*

#### 2.2.4 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 analyzed semantic relations between elements (content words or their replacement) in a text, or “cohesive ties”, including pronominal reference (cf. Table 2.4). Referential function, which is a key component in establishing cohesion, is known

Table 2.4: Studies addressing cohesion in aphasic discourse: methods, findings, and the relationship between cohesion and coherence.

Investigator(s)	Group/ multiple case/case	Method	Impaired?	Is cohesion necessary for coherence?
Piehler and Holland (1984)	2 case studies	Halliday and Hasan (1976) system to study cohesive ties. Five major classes of cohesion devices: reference, substitution, ellipsis, conjunction, lexical reiteration and collocation	Recovery over time	No
Ulatowska et al. (1983, 1981)	10 PWA, 10 NBD; 15 PWA, 15 NBD	4-point rating scale of clarity as a rough measure of cohesion; Marking of verb tenses, definiteness of nominals, appropriate connectors for sequencing	No	–
Armstrong (1987)	3 case studies	Halliday and Hasan (1976) system; Cohesive Harmony Index (CHI) – the percentage of tokens in the chains to the total number of tokens	Yes	Yes
(Chapman & Ulatowska, 1989; Ulatowska, Allard, & Chapman, 1990)	Groups (5-15 PWA)	Comprehension study: ambiguous antecedents resolution task	Yes	–
Glosser and Deser (1991)	9 PWA (fluent), 9 patients with Alzheimer's disease, 9 patients with CHI	Halliday and Hasan (1976) system	Yes	No
Bloom, Borod, Santschi-Haywood, Pick, and Obler (1996)	12 PWA, 9 RBD, 12 NBD	Analysis of specific lexical devices used to establish cohesion (Gleason et al., 1980; Nicholas, Obler, Albert, & Helm-Estabrooks, 1985; Ulatowska, Hayashi, Cannito, & Fleming, 1986); Analysis of referential system (anaphora, deictic terms, indefinite terms, definite articles), and connective devices	No	No
Coelho, Liles, Duffy, Clarkson, and Elia (1994), Coelho and Flewellyn (2003)	Longitudinal case study	Halliday and Hasan (1976) system; Cohesive adequacy (Liles, 1985)	Recovery over time	No
Andreetta (2014), Andreetta, Cantagallo, and Marini (2012)	10 PWA (anomic), 10 NBD; 20 PWA, 20 NBD	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)	Yes	No

*Note: PWA == people with aphasia, NBD == non-brain-damaged people, RBD == people with right hemisphere brain damage*

to be particularly vulnerable in aphasia (Olness & Ulatowska, 2011; Ulatowska et al., 1999); PWA tend to omit antecedents of pronouns and create anaphoric ambiguity (cf. Andretta et al., 2012; Armstrong, 2000; Bloom, Borod, Santschi-Haywood, Pick, & Obler, 1996; Boyle, 2011; Glosser & Deser, 1991; Liles & Coelho, 1998; Marini et al., 2011; Ulatowska et al., 1983; Ulatowska et al., 1981). 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 non-brain-damaged participants was explored. Their findings suggested that cohesion is facilitated for people with aphasia 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 studying cohesion in aphasia are outlined in Table 2.4.

### 2.2.5 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 2.5.). 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” (Foltz, 2007; Olness, 2006; Ulatowska, Olness, Samson, et al., 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, Laakso, Vuorinen, & Rinne, 1998; Rogalski, Altmann, Plummer-D’Amato, Behrman, & Marsiske, 2010). The procedure includes segmenting samples into verbalizations, a verbalization being an independent clause with all its dependent clausal and non-clausal elements, and rating the contextual appropriateness of every verbalization by two independent trained raters. To assess global coherence, the relevance of a verbalization 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 non-brain-damaged (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, Capilouto, & Koutsoftas, 2013; Wright et al., 2010).

Although Ulatowska and colleagues claimed the discourse of their aphasic participants

to be well-structured, other studies reported higher scores for non-language-impaired speakers. Fergadiotis, Kapantzoglou, 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 (Laudauer & Dumais, 1997). Their findings indicated the existence of 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 non-brain-damaged speakers. She analyzed 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 emphasized 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 et al., 2011; Marini, Boewe, Caltagirone, & Carlomagno, 2005; Marini et al., 2008) 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 et al., 2012; Andreetta & Marini, 2015). 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 authors’ 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 2.5 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, stripped of extra-linguistic 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. Section 2.3.3).

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 extra-linguistic 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 (e.g., thematic informativeness; Ulatowska et al., 1990), can shed more light on how coherence is achieved, and what causes its disruption.

Table 2.5: Studies addressing coherence in aphasic discourse: definitions, methods, and findings.

Investigator(s)	Group/mult. case/case	Definition of coherence	Elicitation task	Method	Results
Ulatowska et al. (1983), Ulatowska et al. (1981)	10 PWA, 10 NBD; 15 PWA, 15 NBD	Well-formedness of a text in terms of plausibility, conventionality, and conclusiveness of text	Personal narrative, procedural discourse, story retelling	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;	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.
Glosser and Deser (1991) (G&D 1991)	9 PWA (fluent), 9 patients with Alzheimer's disease, 9 patients with closed head injury (CHI)	A term used to characterize conceptual organizational aspects of discourse at suprasentential level; It depends at least partially on the speaker's ability to maintain thematic unity (p. 69)	Personal narrative	5-point rating scale	No difference between aphasic and control group in ratings of either global or local coherence
Coelho et al. (1994)	Longitudinal case study	Story grammar (number of complete episodes consisting of an initiating event, an action, and a direct consequence)	Picture series (19 frames) and single picture description	Number of episodes	While cohesion improved, story grammar remained moderately depressed over 12 months
Christiansen (1995)	15 PWA (5 anomic, 5 conduction, 5 Wernicke), 20 NBD	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)	Picture series description (4 cartoon stories)	Analysis of propositional content (events, states, elaborations, and comments), coherence violations (information gaps, relevance and progression)	Discourse skills of PWA and NBD were different

*Continued on next page*

Coelho and Flewellyn (2003)	Longitudinal case study	Global c. refers to how discourse is organized 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)	Picture series (19 frames) and single picture description	3-point modification of G&D (1991) scale (Leer & Turkstra 1999 version converted into z-scores)	As severity of aphasia decreased and microlinguistic skills recovered, the coherence remained impaired and showed minimal change over 12 months
Olness (2006); Olness and Ulatowska (2011)	12 PWA	“Hanging together” of the text as a unit (p. 176) /The overall semantic unity of themes and topics in a discourse (Glosser, 1993)	Picture description; examples from personal narratives	Analysis of narrative superstructure (sequential order, setting, initiating event, complicating action, resolution, coda (Berman, 1997; Labov, 1972)); A single rating of semantic unity is used	Abbreviation of narrative event line in aphasia, aphasia severity level correlates with narrative production skills
Wright et al. (2013), Wright et al. (2010)	14 PWA, 14 NBD; 50 healthy adults	Global c. reflects how discourse relates to the overall topic (G&D 1991)	Picture series description (two wordless books)	4-point rating scale (Koutsoftas et al., 2009)	Control group received significantly higher scores than the aphasic group
Fergadiotis, Kapantzoglou, and Wright (2011)	15 PWA	Global c. reflects how discourse relates to the global theme (Kintsch & Van Dijk, 1978)	Picture series description (two picture books)	Compared the methods of G&D (1991), Koutsoftas et al. (2009), and Latent Semantic Analysis (LSA)	Strong correlations between the three measures, evidence of the impact of aphasia severity on coherence
Andreetta (2014), Andreetta et al. (2012), Andreetta and Marini (2015), Marini et al. (2011)	2 case studies; 10 PWA (anomic), 10 NBD; 20 PWA, 20 NBD	The ability to semantically relate remote utterances in the framework of a given discourse	Picture series and single picture description	Local and global coherence errors, cohesion errors	Different performance of clinical and control groups

*Note: PWA == people with aphasia, NBD == non-brain-damaged people, RBD == people with right hemisphere brain damage*

### 2.2.6 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 (e.g., Halliday & Hasan, 1976; Kehler, 1995; Martin, 2001; McNamara, Louwerse, McCarthy, & Graesser, 2010; Shapiro & Hudson, 2014; Thompson, 1994). A direct dependency has been hypothesized 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. 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., 1983; Ulatowska et al., 1981). Many researchers agree that micro- and macrolinguistic abilities are independently organized (e.g., Giora, 1985; Glosser & Deser, 1991; Lenk, 1998; Tanskanen, 2006; Ulatowska et al., 1983; Ulatowska et al., 1981). 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 macro-linguistic 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.

### 2.2.7 Multi-level 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 Jr & O’Brien, 1995; Sanders, Spooren, & Noordman, 1992). Thus, several combinations of different scales and methods have been suggested (cf. Table 2.6). 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 multi-level 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 multi-level 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 multi-level approach developed by Wright and Capilouto (2012) that combined micro- and macrolinguistic measures, including syntactic complexity, information content, lexical diversity, and global coherence.

Table 2.6: Multi-level approaches to studying discourse coherence in aphasia.

Investigator(s)	Aim	Included measures	Results
Vermeulen, Bastiaanse, and Van Wageningen (1989), Bastiaanse et al. (1996);	Analysis of spontaneous speech in aphasia	Speech rate (words/min), MLU, analysis of conjunctions, auxiliary verbs, prepositions, empty words, incomprehensible speech; TTR; number of semantic and phonetic paraphasias, neologisms; Distribution of subordinate/main clauses, distribution of lexical words/copulas;	Reduced ability to use complex grammatical structure; reduced frequency and diversity of verbs
Glosser and Deser (1991)	Investigation of discourse organization, in particular the dissociation between micro- and macrolinguistic cognitive processes, in speakers with different types of brain damage (Alzheimer's disease, TBI, and post-stroke aphasia)	Local and global coherence, cohesion, lexical measures (literal and verbal paraphasias, indefinite terms), syntactic complexity and occurrence of syntactic errors	Impairments on measures of syntactic completeness and complexity, microlinguistic aspects of discourse production, lexical-semantic abilities, inappropriate use of cohesive ties in PWA speech; Thematic coherence within normal limits
Sherratt (2007), Sherratt and Bryan (2012)	Assessment of the applicability and utility of using a multi-level discourse processing model to examine the interaction between different linguistic levels (RBD and NBD speakers)	Fluency, cohesion (Halliday & Hasan, 1976), clarity disruptors (non-specific elements and word substitutions); syntactic analysis (including the analysis of clausal structures); discourse grammar, content and fluency disruptors, relevance, productivity	Multi-level procedures provide a more realistic perspective of discourse and a more objective assessment of certain discourse features
Andreetta et al. (2012), Andreetta and Marini (2015), Marini et al. (2011)	Investigation of discourse processing in anomic aphasia, and the connection between lexical impairment and discourse organization	MLU; semantic paraphasias, phonological errors; percent of complete sentences; local and global coherence errors, index of cohesiveness; number of thematic units, percent of lexical information units (grammatically, pragmatically, and phonologically appropriate words)	Reduced lexical retrieval and informativeness, speech rate; increased amount of global coherence errors; The data suggest that PWAs don't have conceptual organization problems; their macrolinguistic deficit is an epiphenomenon of the microlinguistic problems
Wright and Capilouto (2012)	Exploring maintenance of global coherence in aphasia	Syntactic complexity index (total clauses/total independent clauses); percent of information units (intelligible words, neologisms, fillers, partial words, and commentary); Lexical diversity (voc-D); 4-point scale for global coherence	Lower global coherence scores for the aphasic group. Best microlinguistic predictors of global coherence were percent of information units and lexical diversity.

Note: PWA == people with aphasia, NBD == non-brain-damaged people, RBD == people with right hemisphere brain damage, MLU == mean length of utterance, TTR == type-token ratio

Marini and colleagues (Andreetta & Marini, 2015; Marini et al., 2011) developed a multi-level procedure specifically designed for the assessment of macro- and microlinguistic skills of people with aphasia. Some of the variables at the micro-structural 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, et al., 2005; Marini, Carlomagno, Caltagirone, & Nocentini, 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 multi-level analyses is the possibility to account for the inter-relatedness 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 multi-level approaches is blurred by the lack of consideration of the interactions between different 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 supra-sentential organization, 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 macro-linguistic 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 and 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 multi-level approaches. Some of the measures refer to the same concepts, but use different mechanisms to assess them, while others target the same concepts, but operationalize 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.

### **2.2.8 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.

Table 2.7: Communicative effectiveness and efficiency in aphasic discourse.

Investigator(s)	Group/ multiple case/case	Aim	Method
Pedersen et al. (2001)	68 PWA	Adaptation and psychometric evaluation of CETI for Danish	CETI for Danish, test-retest reliability
Irwin, Wertz, and Avent (2002)	20 PWA	Establish the relationship among language impairment, functional communication, and pragmatic performance in aphasia	Porch Index of Communicative Abilities (PICA; Porch, 1967), Rating of Functional Performance (RFP; Wertz et al., 1981), and Pragmatic protocol (Prutting & Kirchner, 1983); Correlation analysis
(Ramsberger & Rende, 2002)	14 PWA, 56 NBD	Measuring transactional success in conversation	Number of correct ideas produced in the NBD partner's retelling of a story which PWA watched at the conclusion of the conversation
Kong and Law (2004, 2009)	10 PWA, 30 NBD; 5 PWA (longitudinal)	Communication effectiveness assessment in Cantonese aphasic speakers	Cantonese Linguistic Communication Measure (CLCM), including the Index of communication efficiency (N of informative verbs per min)
McCullough, McCullough, Ruark, and Rainey (2006)	27 PWA,	Define the relationship between pragmatic performance and functional communication	Pragmatic Protocol, American Speech-Language and Hearing Association – Functional Assessment of Communication Skills (ASHA-FACS; Frattali, Thompson, Holland, Wohl, & Ferketic, 1995)
Van der Meulen et al. (2010)	122 PWA, 25 NBD	Assessment of verbal and non-verbal communication in severe aphasia	Scenario test for both verbal and non-verbal communication in daily life situations (4-point scale) and interactive settings
Ruiter et al. (2011)	10 PWA, 20 NBD	Development of a quantitative measure of verbal effectiveness and efficiency in the ANELT	Content Unit Analysis (Yorkston & Beukelman, 1980) adapted for the ANELT scenarios
Muò et al. (2015)	60 PWA, 20 patients with traumatic brain injury, 100 NBD	Adaptation and psychometric evaluation of ASHA FACS for Italian	Italian version of ASHA-FACS for Italian (I-ASHA-FACS)

*Note: PWA == people with aphasia, NBD == non-brain-damaged people, RBD == people with right hemisphere brain damage*

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. Manochiopinig, Sheard, and Reed (1992) reviewed fifteen communicative effectiveness measures and subdivided them into five types: observational profiles, communicative efficiency measures, standardized 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” (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 reevaluated, and several new methods have been proposed to complement them. Thus, the Scenario Test (Van der Meulen et al., 2010) extends the Amsterdam-Nijmegen Everyday Language Test (ANELT; Blomert, Kean, Koster, & Schokker, 1994; Blomert, Koster, & Kean, 1995), 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 non-verbal 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 and to other languages, for example, Danish (Pedersen, Vinter, & Olsen, 2001), Italian (Vernero et al., 2002; 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 (Nicholas & Brookshire, 1993, CIU/min). 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 (Gordon, 2008; 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 2.7.

The connection between linguistic competence and communicative effectiveness is not

very straightforward. For example, measures of pragmatic performance and functional communication have been shown to address different, though possibly overlapping, aspects of performance in people with aphasia (Irwin et al., 2002; McCullough et al., 2006). Armstrong and Ferguson (2010b) addressed the role of language in “functional communication” and suggested that further investigation of different behaviors 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 multi-level procedure similar to the ones described in the previous section, is a way to shed light on the dissociation between linguistic impairments and success of communication.

## 2.3 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 programs 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 has 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, sample size, genre, elicitation task, modality, and addresses cross-linguistic studies of discourse in aphasia.

### 2.3.1 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 counter-evidence, 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 organization and to provide challenging counter-examples, case-studies do not allow for making generalization 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 lengthy and complicated process. Most of the studies on aphasic discourse have been based on the analysis of small to medium size groups of participants (e.g., 3-10 PWA in Armstrong and Ulatowska (2007), Glosser and Deser (1991), Hough (1990), Ruigendijk, Vasić, and Avrutin (2006), Ulatowska et al. (1981); 11-15 in Christiansen (1995), Goodglass, Christiansen, and Gallagher (1993), Olness (2006), Ulatowska et al. (1983)), although for several studies larger numbers of participants were recruited (e.g., 20 agrammatic PWA in Miceli, Silveri, Romani, and Caramazza (1989); 28 in Olness (2007); 74 in Wagenaar, Snow, and Prins (1975); 121 in Vermeulen et al. (1989)). For the purpose of reducing the amount of time and labor 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 8 languages by May 1, 2014 (Forbes, Fromm, & MacWhinney, 2012; Fromm, Forbes, Holland, & MacWhinney, 2014; MacWhinney, 2000; MacWhinney, Fromm, Forbes, & Holland, 2011; MacWhinney, Fromm, Holland, & Forbes, 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 6 participants has been reported so far (Westerhout & Monachesi, 2006). Another project – the Moss Aphasia Psycholinguistics Project Database (MAPPD; Mirman et al., 2010) – contains a collection of behavioral 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 recognized by funding agencies. However, when choosing a corpus collected by a third party for generalized 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 video-based.

### 2.3.2 Genre

Many researchers noted that the choice of elicitation task influences quality and quantity of the discourse produced (Coelho, 2002; Olness, 2007; Olness, Ulatowska, Wertz, Thompson, & Auther, 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., 1983; Ulatowska et al., 1981; Weinrich, McCall, Boser, & Virata, 2002) and personal narratives (Armstrong & Ulatowska, 2007; Behrns, 2009; Olness & Ulatowska, 2011; Ulatowska, Olness, Samson, et al., 2004; Ulatowska et al., 2013) received special attention due to their being 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 as compared to narratives, whereas narratives of people with aphasia had lower information content, shorter story line, more errors in the order of events, and syntactically less complex language than those of control participants (Ulatowska et al., 1990; Ulatowska, Olness, & Williams, 2004).

Gernsbacher and Givón (1995) emphasized 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 people with aphasia to co-construct meaning in communication through the analysis of such phenomena 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, 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, 2007). The special issue of *Aphasiology* (Wilkinson, 2015) on Conversation Analysis (CA) application to aphasic data, recently addressed such topics as repair (Barnes & Ferguson, 2015; Laakso, 2015; Penn, Frankel, & Wilkinson, 2015), adapted behaviors of conversation partners (Klippi, 2015), and interaction-focused therapy for aphasia (Beeke et al., 2015; Damico et al., 2015; Saldert, Johansson, & Wilkinson, 2015), emphasizing the importance of studying language in interaction and taking into account the role of conversation partners in the recovery of people with aphasia.

### 2.3.3 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 to in the years after Armstrong's (2000) review. Behrns (2009) compared written and oral narrative production in people with aphasia 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 its being potentially complementary or even compensatory to language (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, 2013). Including gestures in further analyses of conversations with people with aphasia can be insightful with respect to the general mechanisms of meaning co-construction in interaction (Pritchard, Dipper, Morgan, & Cocks, 2015).

### 2.3.4 Cross-linguistic 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, cross-linguistic comparisons are extremely valuable in discourse studies. While a considerable amount of evidence comes from English, cross-linguistic studies in aphasiology started emerging as early as in the 80s 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 strikingly similar across languages. Case-studies of narratives in aphasia in 14 languages have been collected in *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 people with aphasia had linguistic deficits rather than an impairment of control of speech production. MacWhinney and Holland promoted cross-linguistic 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 cross-linguistic perspective in order 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 generalization. Studies on different languages can stimulate the development of therapeutic techniques for the language-impaired speakers of those languages. At the same time, cross-linguistic comparisons are important for the understanding of the universal mechanisms of language production and its deterioration in aphasia.

## 2.4 Conclusion

Discourse production is the most important channel for communication. Comparing discourse produced by non-brain-damaged 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 organized into a whole is not yet entirely clear. Understanding how language production functions is crucial for the understanding of 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 macro-linguistic skills of people with aphasia, who undeniably have difficulties with language-based communication. It is of yet unclear, however, if their ability to construct understandable connected discourse is impaired per se, or if it is the

result of various micro-linguistic 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 at showing 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 people with aphasia. Armstrong also noted the importance of exploring the connection between micro- and macro-linguistic 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 the issues remain unresolved. Methodological variability continues as a critical source of disparate findings. Though new methods and tests allow for a more objective and a more 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 organization at the macro-linguistic 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 extra-linguistic levels, is essential for the understanding of 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 the understanding of the role language plays in establishing communication. This could be achieved, for example, by devising a comprehensive multi-level 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.

## Chapter 3

# Evaluating discourse coherence in anomic aphasia using Rhetorical Structure Theory <sup>2</sup>

### Abstract

*Purpose:* The existing body of work regarding discourse coherence in aphasia has provided mixed results, leaving the question of coherence being impaired or intact as a result of brain injury unanswered. In this study, discourse coherence in non-brain-damaged (NBD) speakers and speakers with anomic aphasia was investigated quantitatively and qualitatively.

*Method:* Fifteen native speakers of Cantonese with anomic aphasia and 15 NBD participants produced 60 language samples. Elicitation tasks included story-telling induced by a picture series and a procedural description. The samples were annotated for discourse structure in the framework of Rhetorical Structure Theory (RST) in order to analyze a number of structural parameters. After that 20 naive listeners rated coherence of each sample.

*Result:* Disordered discourse was rated as significantly less coherent. The NBD group demonstrated a higher production fluency than the participants with aphasia and used a richer set of semantic relations to create discourse, particularly in the description of settings, expression of causality, and extent of elaboration. People with aphasia also tended to omit essential information content.

*Conclusion:* Reduced essential information content, lower degree of elaboration, and a larger amount of structural disruptions may have contributed to the reduced overall discourse coherence in speakers with anomic aphasia.

---

<sup>2</sup>This chapter was adapted from: Kong, A. P.-H., Linnik, A., Law, S.-P., & Shum, W. W.-M. (2018). Measuring discourse coherence in anomic aphasia using Rhetorical Structure Theory. *International journal of speech-language pathology*, 20(4), 406–421. Published online: 17 Mar 2017. doi:10.1080/17549507.2017.1293158 (cf. the attached contribution statement)

## 3.1 Introduction

Oral narratives are a crucial form of discourse used in everyday life. Their production involves both micro- and macro-linguistic processes. The micro-linguistic level includes lexical processing, organization of phonological information into morphological strings and words, and then into syntactic constructions (Marini & Fabbro, 2007). Macro-linguistic abilities are associated with discourse processing at the supra-sentential level. For example, cohesion, or semantic relations between contiguous utterances, established through the use of lexical and grammatical devices, such as conjunctions, coreference, and ellipsis (Halliday & Hasan, 1976), is a macro-linguistic property. These discourse properties also mediate local connections between sentences or utterances with the global relations among propositions, minimal semantic units within a text, to integrate linguistic and conceptual features (Kintsch, 1994).

### 3.1.1 Coherence

Coherence is a complex phenomenon (e.g., Givón, 1993), and the problem of defining it has been extensively addressed by a number of researchers (Foltz, 2007; Gernsbacher & Givón, 1995; Kehler, 2002, 2004; Kintsch & Van Dijk, 1978; Sanders & Spooren, 2001; Van Dijk, 1977). Independent of the definition, however, coherence is a crucial property which transforms a sequence of sentences into a discourse (e.g., Kehler, 2002; Ulatowska et al., 2003). It refers to the semantic connectedness or ‘hanging together’ of speech, or the semantic connectedness of discourse at the propositional level (Van Dijk, 1980). Discourse coherence can be further divided into local and global. Local coherence reflects a speaker’s ability to establish connection between currently processed information with the immediately preceding context (Albrecht & O’Brien, 1993), whereas global coherence refers to a speaker’s ability to semantically relate remote utterances to the theme, topic, or gist of a discourse (Kintsch & Van Dijk, 1978).

### 3.1.2 Coherence in aphasia: methods

Discourse produced by speakers with aphasia is often perceived as vague and lacking clarity (Early & VanDemark, 1985; Gleason et al., 1980; Ulatowska et al., 1981), but it has also been noted that people with aphasia preserve remarkably good functional communication skills (Holland, 1982; Huber, 1990; Olness et al., 2010; Ulatowska et al., 1990). It has been hypothesized that the decreased understandability of discourse in aphasia reflects difficulties with establishing coherence (Ulatowska et al., 1981; Van Dijk, 1980).

Research on macro-linguistic abilities of people with aphasia has offered a variety of methods to evaluate local and global coherence defined through the concepts of topic maintenance and thematic unity, as well as coherence error analysis (e.g., Christiansen, 1995; Coelho & Flewellyn, 2003; Olness, 2006; Olness & Ulatowska, 2011; Ulatowska et al., 1990; Ulatowska et al., 1981; Wright et al., 2013). These methods can be roughly divided into perceptual, based on human judgment, and factual, or data-driven, ones. The first type is represented by a number of rating scales. Ratings are intended to assess the overall level of coherence of a discourse either through the assessment of every utterance’s relatedness to the semantic unity of the discourse, or through a single rating of the degree to which it can be considered a unity. On the one hand, ratings are not able to provide an insight into the linguistic factors contributing to the maintenance of coherence. On the other hand, through the use of listeners’ judgment for the assessment of speech production, listeners’ contribution to coherence establishment is taken into

consideration. Thus, the fact that coherence is co-established by both speaker and listener (e.g., Gernsbacher & Givón, 1995; Wright & Capilouto, 2012), is not dismissed. The second type includes a collection of data-driven methods, which do not involve human judgment. This can be story grammar analysis (e.g., Coelho et al., 1994) or a similar analysis of narrative superstructure (Olness, 2006; Olness & Ulatowska, 2011), which are based on the assessment of main narrative elements (e.g., setting, initiating event, complicating action, resolution, and coda; Labov, 1972), as well as variations of analyses of propositional content and coherence violations in a sample (Andreetta et al., 2012; Christiansen, 1995).

### 3.1.3 Coherence in aphasia: previous findings

A number of studies demonstrated that discourse coherence in aphasia is impaired (e.g., Andreetta et al., 2012; Andreetta & Marini, 2015; Christiansen, 1995; Coelho & Flewellyn, 2003; Wright & Capilouto, 2012), whereas other studies provided evidence that it is within normal limits (Glosser & Deser, 1991; Ulatowska et al., 1981; Ulatowska et al., 2013). Further disparities are related to the more global question about the origins of coherence. It has been suggested that impaired micro-linguistic skills lead to macro-linguistic processing difficulties. For example, morphosyntactic deficits may cause poor cohesion, which, in turn, is correlated with global coherence (e.g., Armstrong, 1987). Contrary to this idea, several studies demonstrated that thematic coherence in oral discourse can remain relatively intact despite micro-linguistic deficits in aphasic speech (e.g., Coelho & Flewellyn, 2003; Glosser & Deser, 1991; Ulatowska et al., 1983; Ulatowska et al., 1981). This indicates that micro- and macro-level of discourse may be organized independently. The techniques discussed above are valuable for assessment and therapy outcomes evaluation, as they reliably estimate the level of coherence in discourse. Nevertheless, they do not provide an insight on the micro- and macrolinguistic factors influencing coherence impairment or preservation.

As a solution to this problem, a number of multi-level approaches have been developed and implemented to explore the relationship between the micro- and macro-linguistic abilities of aphasic speakers (e.g., Marini et al., 2011; Sherratt, 2007; Wright & Capilouto, 2012). For example, Andreetta et al. (2012) investigated the effect of lexical retrieval difficulties on macro-linguistic processing during the construction of a narrative. They concluded that impaired word finding reduced the levels of sentence completeness and the overall degree of cohesion across utterances, whereas lexical fillers and repetitions lowered the overall level of global coherence in spoken discourse. (Wright & Capilouto, 2012) also studied the effect of micro-linguistic impairments on the maintenance of global coherence in aphasia in a story-telling task. They show that reduced information content and lexical diversity had an effect on coherence in the stories of aphasic participants, which is consistent with the findings in a number of earlier studies (e.g. Christiansen, 1995; Coelho & Flewellyn, 2003).

### 3.1.4 Different aspects of coherence

Several factors have limited the generalizability of the findings to date. One of the issues is the variety of methods developed to analyze coherence. Comparing the results of studies is not a straightforward task due to theoretical and procedural differences (Linnik et al., 2016). Different aspects of coherence evaluated in previous studies include, for example, appropriateness and completeness of thematic content (e.g., Glosser & Deser, 1991; Wright et al., 2010), coherence violations (Andreetta et al., 2012; Christiansen, 1995; Marini et al., 2011), and semantic-pragmatic unity of discourse as a whole (Olness,

2006; Ulatowska, Olness, & Williams, 2004). The current study followed the latter path and addressed coherence as an intrinsic trait of discourse, specifically, its overall unity and connectedness.

### 3.1.5 Genre-related differences

Apart from different methodologies used in previous research, the variability in discourse genres is another factor contributing to the disparities in results. Elicitation tasks included picture description (e.g., Andreetta et al., 2012; Marini, Boewe, et al., 2005; Saffran, Berndt, & Schwartz, 1989), personal narratives (e.g., Glosser & Deser, 1991; Olness & Ulatowska, 2011), and procedural discourse (e.g., Ulatowska et al., 1990; Ulatowska et al., 1981). Longacre (1996) emphasized that different genres are associated with specific patterns of linguistic features and different logical and thematic organization. Different elicitation tasks also impose different cognitive and linguistic demands on aphasic speakers (Bliss & McCabe, 2006). Specifically, story-telling using pictures can be considered cognitively less demanding than a personal narrative or expository discourse because these two latter tasks require organization of several ideas in the lack of visual support. Marini, Boewe, et al. (2005) investigated discourse production of a large group of healthy adults and found a higher degree of coherence for narratives elicited using a sequential picture description task, as compared to narratives elicited by a single picture. They attributed this difference to the higher level of inter-relationship among characters in a series of pictorial stimuli with multiple themes within the description. Olness (2006) examined the difference between narratives elicited using sequences of pictures and single pictures. She concluded that the latter have limitations for discourse studies, while “what is traditionally elicited by single pictures may not be discourse” (p. 185) as it does not require connectivity, or coherence. (Wright & Capilouto, 2012; Wright, Koutsoftas, Capilouto, & Fergadiotis, 2014), in turn, reported significantly lower coherence scores for personal recounts, as compared to picture-elicited stories in an aphasic population. These findings are clinically relevant to the selection of assessment and treatment materials targeting discourse production. Based on these considerations, two types of narratives were chosen for the current study, namely procedural description elicited with a single picture and story-telling with a series of pictures. The choice of elicitation tasks was motivated by the comparability of resulting samples.

### 3.1.6 The present study

In this study, we explore macro-linguistic impairments in aphasia and the linguistic processes underlying discourse production using a combination of cognitive and structural approaches. Specifically, we apply Rhetorical Structure Theory (Mann & Thompson, 1988), a relational approach to the investigation of discourse structure. This theory has been widely used for healthy discourse analysis. Additionally, a rating scale is used to assess listener perception of global coherence. By integrating data-driven and perceptual perspectives, we expect to be able to grasp different aspects of coherence as a phenomenon, and shed light on aphasic speakers’ ability to establish coherence.

## 3.2 Rhetorical Structure Theory

It has been argued that coherent discourse has an internal structure and that this structure is hierarchical, rather than linear (e.g., Fox, 1987; Grosz & Sidner, 1986; Hobbs, 1985). One of the frameworks formalizing this hypothesis is Rhetorical Structure Theory

(RST; Mann & Thompson, 1988; Taboada & Mann, 2006b). RST provides a methodological apparatus for the description of structural organization of discourse. According to RST, discourse consists of elementary discourse units (EDUs), minimal building blocks of discourse structure which roughly correspond to clauses (Mann & Thompson, 1988). These basic segments are connected to each other with semantic, or rhetorical, relations ('Consequence', 'Cause', 'Evaluation', 'Elaboration', etc.), forming a tree-like structure (Fig. 3.1). Coherence is thus achieved through the establishment of discourse structure.

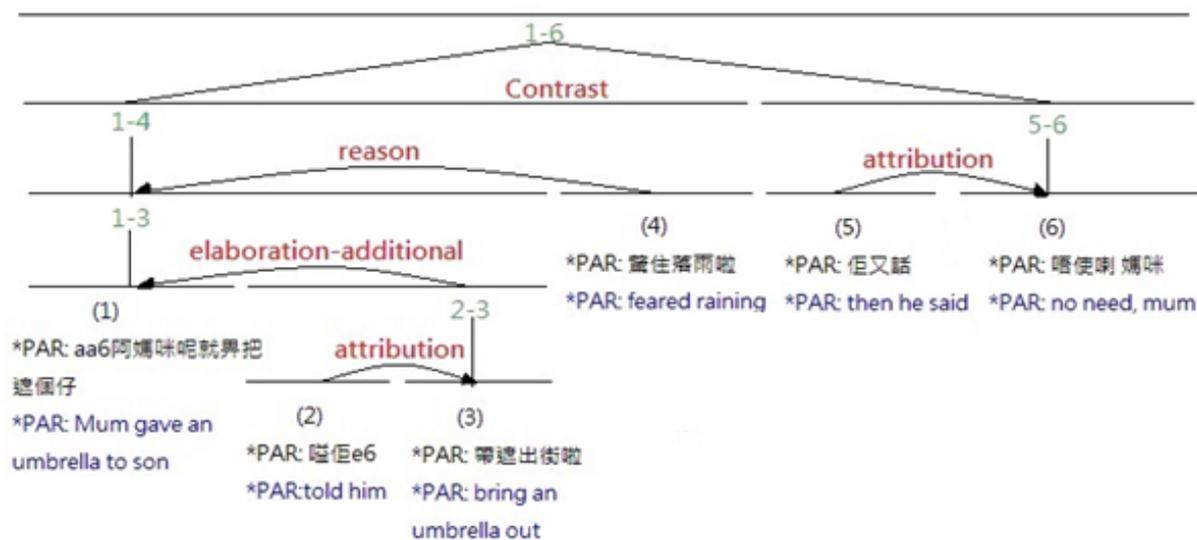


Figure 3.1: Example of the RST analysis in Cantonese

### Rhetorical Structure Theory: pros and cons

RST has been widely used in written and spoken discourse analysis, as well as in computational linguistics (see Taboada & Mann, 2006a, 2006b, for a review). A number of other approaches were based on similar theoretical considerations about discourse structure, but work within methodologically different solutions/frameworks (Cristea, Ide, & Romary, 1998; Grosz & Sidner, 1986; Lascarides & Asher, 2008; Moore & Pollack, 1992; Moser & Moore, 1996; Prasad et al., 2008; Stede, 2004; Walker, 1996; Wolf & Gibson, 2005). None of them has been as extensively tested as RST. Our choice of RST for the analysis of discourse in aphasia is motivated primarily by its plausibility with respect to certain cognitive processes involved in discourse production. Specifically, RST is based on the view that coherence relations are cognitive entities responsible for holding spans of discourse together and are playing an important role in the interpretation of discourse. Slight changes in structure can result in different interpretations of a discourse (e.g., Sanders, 1993; Taboada & Mann, 2006b). Hence, the ability to establish coherence relations or a discourse structure is crucial for communication. Psychological validity and technical adequacy of RST has been questioned, due to the constraints the framework imposes. Among them, for example, are the requirement of a single inference when linking discourse segments, and the lack of a unified opinion on the number and classification of coherence relations among the researchers using relational approaches. These points of criticism, among a number of others, have been addressed at length in previous literature (Marcu, 2000, 2003; Taboada, 2004; Taboada & Mann, 2006b), and, despite constituting an important and interesting discussion, will not be reiterated in this paper. Although RST analysis may be more laborious and complicated than the methods previously used

in aphasic discourse structure investigations, it is the first method that addresses aphasic speakers' ability to establish relational connectivity between discourse segments. It delivers a fine-grained result, thus granting an opportunity to look deeper into the processes of discourse organization. Crucially, as compared to other methods, it provides information on relations (or a lack of thereof) between discourse spans in addition to the quantitative measure of propositional content.

RST has a number of other advantages. It allows marking, or annotating, salient information distribution during discourse production (Marcu, 1999; Stede, 2008). Namely, when linking two EDUs, one has to decide on the so-called "nuclearity status" of the constituents. In a relationship nucleus is a constituent containing a more salient piece of information, which could not be eliminated without a significant loss in information content, while a satellite contains less essential, additional information. In the example on Fig. 3.1, the facts that a mother gave her son an umbrella (1) and that he refused to take it (6) are equally important in terms of content, while other EDUs provide additional information, such as the reason why she did so (4) and an elaboration specifying that she told him to take an umbrella (3). In this case, EDUs (1) and (6) are nuclei, linked with a multinuclear relation ("Contrast"). The EDUs containing supplementary information are connected to these two nuclei with mononuclear relations ("Reason" and "Elaboration-additional"). Apart from the organization of information in the discourse flow, semantic (also called "discourse" or "rhetorical") relations in RST reflect some of the cognitive processes involved in discourse production, such as reasoning, evaluation, and inference. In addition, the use of cognitive predicates or reported speech (e.g., "he thought", "she said"; EDUs (2) and (5) on Fig. 3.1) are marked through the relation of "Attribution".

Chinese, including Cantonese, is generally characterized by the lack of inflectional morphology (Packard, 2000; Wang & Sun, 2015) and frequent use of elliptical sentences (Chung, Code, & Ball, 2004). The common absence of number and gender agreement between pronouns and nouns and the usual omission of topic and grammatical subjects in sentences can present difficulties for the theories relying on signaling of relations between parts of discourse, such as through discourse markers. Although some suggestions about connectors signaling certain relations have been made for English (Carlson & Marcu, 2001), RST does not imply the surface signaling of relations, for example, through discourse markers or connectives (cf. Taboada, 2006), which makes it an appropriate choice for Cantonese (and Chinese) discourse analysis. Since RST is not language-specific, it can also be used for cross-linguistic comparisons at the discourse level (see Iruskieta, Da Cunha, & Taboada, 2015, for discussion). The findings of this work are thus valuable for the understanding of the general principles of language production.

### 3.3 Aims and research questions

In this study we aim to examine the differences in the way discourse structure and subsequently coherence are established by people with aphasia and non-brain-damaged (NBD) speakers. Based on previous literature, the following research questions are posed in the present study.

Is coherence reduced in the discourse produced by aphasic speakers?

Do aphasic speakers demonstrate impairments in the macrolinguistic organization?

If so, are these impairments related?

To answer the first question, RST analysis was performed in order to capture potential failures in the process of building a discourse structure. The macrostructures in the samples produced by aphasic and non-aphasic participants were compared on a number of variables. To answer the second question, ratings from naive listeners were collected. Finally, a correlation analysis between the RST-variables and coherence ratings was performed to answer the third question.

Various microlinguistic difficulties, specifically, word-level errors and paraphasias, or a reduced proportion of function words, which are used to express grammatical meaning, are common in aphasia. Thus, if coherence is reduced in the aphasic group, we wish to find out whether this resulted from impaired microlinguistic processing, or if the macro- and microlinguistic problems are independent. Two microlinguistic variables have been included in the correlation analysis to determine whether microlinguistic deficits are related to coherence or not. As discussed earlier, several variables, such as the availability of visual cues, presence of a thematic structure and story characters, have been suggested to influence cognitive demand and, as a result, on the linguistic performance of people with aphasia (e.g., Fergadiotis, Wright, & Capilouto, 2011; Olness, 2006; Ulatowska et al., 2003). Based on previous findings, we expect to find differences in structural properties in discourse of different genres (Longacre, 1996; Olness, 2007; Pritchard et al., 2015; Ulatowska et al., 1983; Ulatowska et al., 1981). Lastly, this study also intends to validate a quantitative structural approach to the investigation of discourse coherence based on RST through the analysis of correlations between its outcomes and naive listeners' ratings.

## 3.4 Method

### 3.4.1 Data

The data used in the current study are a part of the corpus of Cantonese aphasic discourse (Kong, Law, & Lee, 2012), for which data collection methods and stimuli were adapted for Cantonese from the AphasiaBank protocol (MacWhinney et al., 2011). All language samples collected were transcribed in the Codes for the Human Analysis of Transcripts format (CHAT; MacWhinney, 2000).

A total of 60 transcripts from 13 male and 2 female native Cantonese participants diagnosed with anomic aphasia, according to the Cantonese version of the Western Aphasia Battery (CAB; Yiu, 1992), and an equal number of non-brain-damaged participants matched in gender, age ( $\pm 5$  years), and education level ( $\pm 1$  year) were extracted from the database. The age of the aphasic group ranged between 43 and 72 years (mean = 55.2 years; SD = 9.70 years) and the aphasia quotients ranged from 77.1/100 to 99/100 (mean = 89.6; SD = 7.09). The age of NBD subjects ranged between 44 and 71 years (mean = 55.8 years; SD = 8.08 years); none of them had any previous history of psychiatric or neurological illness, learning disabilities, hearing and/or visual impairments that would affect their use of language. The education levels for both speaker groups ranged between 6 and 13 years.

Transcripts from two discourse tasks were chosen to study the effect of genre on discourse macrostructure, namely, (1) a description of the procedure of making an egg and ham sandwich, elicited using a single picture with photos of the ingredients, and (2) a narrative elicited using a series of six pictures (black and white line drawings), depicting a boy who refused to take an umbrella on a rainy day (Fig. 3.2). The pictorial stimuli were first presented to the participants, who were instructed to describe the procedure in the first task and then to tell a story with a beginning, a middle, and an end in the

second task looking at the stimuli. These two genres have been used in a large number of studies, which allows a greater comparability of the results with previous findings. These tasks were selected instead of personal recount (also available in the corpus) because they allow greater control over the content of the elicited discourse samples.



Figure 3.2: Picture sequence “Refused umbrella” from the AphasiaBank protocol used as a stimulus in the story-telling task.

### 3.4.2 Data analysis

#### Discourse segmentation

Each language sample was first segmented into elementary discourse units (EDUs), the minimal semantic building blocks of a discourse (Mann & Thompson, 1988). Discourse segmentation is known to be challenging for annotators, especially when it comes to spoken language (e.g., Artstein & Poesio, 2008). RST was initially devised for unimpaired written language, and EDUs were defined primarily based on syntactic criteria, that is, the presence of a predicate. However, segment boundaries in spoken discourse are often not as clearly defined as they are in written discourse, even less so in aphasic speech. Spontaneous and semi-spontaneous story-telling is occasionally characterized by deviations from the main story line, for example, comments and embedded discourse units used to elaborate on parts of other EDUs. These constructions were identified as “structural expansions” in the analysis. Another group of phenomena intrinsic to spontaneous speech consists of reformulations, self-corrections, repetitions, and retracing, that is, returning to an earlier part of discourse to add information that could help a listener to understand the story and its details. In this study, a combination of phonological (e.g., prosodic contours and pauses), syntactic, and semantic (e.g., semantic completeness) criteria was used. Specifically, three sets of detailed guidelines were used in the segmentation process:

- (1) RST annotation guidelines, which have been developed and tested in the process of creating the RST Discourse Treebank, a corpus of newspaper articles annotated

with RST (Carlson & Marcu, 2001);

- (2) The CHAT transcription format manual (MacWhinney, 2000)
- (3) The guidelines for segmenting spoken language developed by Kibrik and Podlesskaya (2009), focusing on disfluencies, such as repetitions, reformulations, self-interruptions, etc.

The resulting segmentation procedure was similar to the one used and described by Marini et al. (2011). First syntactic clauses were identified where possible, syntactically incomplete clauses were segmented based on phonological boundaries and semantic completeness. The hard-to-define term “semantic completeness” refers to the relative understandability of a piece of discourse and does not necessarily imply well-formedness of the EDU. Aphasic speech may lack syntactic and prosodic indicators of segment boundaries, along with “semantic completeness” in the common sense. For these cases, we took the speaker’s perspective into account and relaxed the aforementioned criteria. Specifically, semantic understandability along with phonological criteria were attributed more weight than syntactic well-formedness in the process of segmentation. Two types of ill-formed EDUs were allowed. First, an “incomplete” EDU was defined as a content-wise comprehensible clause with an omission of a critical syntactic component, such as an object of a transitive verb (e.g., as a result of a word-finding difficulty) (Example (i)).

(i) *Mother [gave] [son] an umbrella.*

Second, if not only grammatical form, but the general sense of an EDU is lost, but, for example, a phonological contour is present, the string of words is identified as a “failed” EDU. EDUs of this type are often accompanied by commentaries about the speaker’s language difficulties (Example (ii)).

(ii) *Mother eh well [...]. Ugh, don’t remember/can’t find the word.*

Additionally, syntactically well-formed discourse units that were semantically out of place, or empty information-wise (italicized in Example (iii)), were linked to the rest of the structure and the connection was marked with a question mark (“?”) relation.

(iii) *Mother gave her son an umbrella. Ehh.. well, that’s like... ehm, no, not that. He didn’t want to take it.*

These three kinds of EDUs were considered to be disruptive for discourse structure.

### **Discourse structure annotation**

After the segmentation, all EDUs were incrementally linked to each other using a set of rhetorical relations following the guidelines of Carlson and Marcu (2001). Marcu’s modification of the RST Tool (Marcu, Amorrortu, & Romera, 1999; O’Donnell, 2000b) was used to perform the annotations. Twenty-six different semantic relations (out of possible 78), such as background, consequence, condition, explanation, and means, were used to annotate the discourse samples in the current study. Examples for each relation are given in the Appendix A. Further details about the relation definitions and assignment can be found in RST Discourse Treebank annotation manual (Carlson & Marcu, 2001).

### 3.4.3 Reliability of EDU segmentation and discourse structure (RST) annotation

Discourse, like smaller elements of natural language, such as words and sentences, is open to more than a single interpretation. Despite the detailed guidelines, the RST annotations bear a certain degree of subjectivity due to the possibility of multiple analyses. Analysis of the same discourse sample is expected to sometimes yield multiple resultant structures (Mann & Thompson, 1987; Stede, 2008; Taboada & Mann, 2006b). Comparing RST annotations is not a straightforward procedure (Iruskieta et al., 2015; Marcu, 2000). The main goal of this study was not to estimate the ability of trained annotators to come up with the same annotations independently, but rather to single out a plausible interpretation agreed upon by multiple raters. Hence, the segmentation was first verified. As the decisions were not independent, percent agreement on all the annotation decisions (instead of the F-measure, i.e., precision and recall) was then obtained. In the current study, all the 60 samples were first annotated by author WWMS. The annotations were then checked by author AL, with an agreement reaching 95% for the segmentation and 85% for the discourse structure annotation. The remaining 15% of structure annotations were discussed by all the authors and an agreement was reached on the analysis.

### 3.4.4 Analyzing discourse organization using RST

Each annotated discourse sample was analyzed in terms of 12 variables possibly contributing to coherence. The variables can be divided into five groups related to certain discourse properties.

I. Speaker's efficiency in formulating complete discourse units, with respect to the possible effect of fatigue:

1. Efficiency of production in the first half of the sample: total number of EDUs produced in the first half of the recording divided by the time elapsed in minutes.
2. Efficiency of production in the second half of the sample: same method of calculation.

II. Connectivity of semantic units within a discourse and discourse complexity:

3. Total number of EDUs, or total length of a discourse sample in EDUs.
4. Number of relation types: the number of different types of rhetorical relations used in the sample.
5. Relation type frequency, or the total number of rhetorical relations of different types used.

III. Degree of elaboration and complexity of discourse:

6. Depth of the structure, or the maximum number of nodes on the tree structure between the top and the lowest node, or the number of levels the structure "branches" down (Kibrik & Podlesskaya, 2009). For example, on Figure 3.1, the depth of nodes (1-4) and (5-6) is 1, the depth of nodes (1-3), (4), (5), and (6) is 2, for nodes (1) and (2-3) it is 3, etc. Thus, the maximal depth is of the tree is 4. In narratives, main events represent main constituents in a tree and are connected at the top. Each of them represents a scene with a different number of elaborative components, which is represented as a subtree with the main constituent as the main nucleus. Hence, this parameter was included as a rough indicator of elaboration, or detail, in the story.

7. Percentage of structural expansions: the number of embedded discourse units and comments divided by the total number of EDUs.
8. Percentage of mononuclear relations: the number of mononuclear relations in the sample divided by the total number of mono- and multinuclear relations.

IV. Degree of inadequacy, impairment, or structural disruption:

9. Percentage of incomplete EDUs.
10. Percentage of failed EDUs.

V. The final two measures were included to capture the effect of micro-linguistic variables on the overall discourse coherence:

11. Percentage of errors: number of all semantic and phonemic paraphasias, morphosyntactic errors, and neologisms divided by the total number of words.
12. Percentage of function words: The ratio of all function words to the total number of words.

### 3.4.5 Coherence ratings

Twenty university students, all native speakers of Cantonese, were recruited to rate the 60 audio recordings as naive listeners. The raters were divided into four groups and the sequence of presentation of the 60 audio files was randomized across groups. The raters were asked to indicate for each audio the following:

1. The level of understandability on a 9-point scale, with “1” meaning that the rater did not understand the content at all and “9” meaning that all content was clearly understood.
2. The connectedness and completeness of the content with three options – ‘complete’, ‘incomplete, but understandable’, and ‘incomplete and hard to follow’.
3. Whether the events were presented in the correct order, and
4. Whether there was a part of the story or procedure description that was hard to understand, and if so, which part it was.

Before the study, a 30-minute briefing and a short practice session were provided to each group of raters.

### 3.4.6 Analysis of information content

In the present study, thematic content was measured through the proportion of main events, or thematic units, that is, main ideas and details in a stimulus mentioned by participants (Andretta et al., 2012; Capilouto et al., 2006; Marini et al., 2011; Marini, Boewe, et al., 2005; Wright, 2011; Wright et al., 2005). The procedure described in Capilouto et al. (Capilouto, Wright, & Wagovich, 2005, 2006) was followed to perform the main events analysis. Unlike in the original procedure, the main events in the stimuli were determined using NBD transcripts. The information – events for the narratives and steps for the procedures – present in at least 70% of the NBD transcripts was classified as ‘essential’. Other information included by participants was considered to be optional, providing additional elaborative components to the main story line or the main steps in the procedure. The original procedure by Capilouto et al. suggests evaluating speakers’

ability to convey relationships between the events in addition to producing main events. However, aphasic speakers often use relatively simple syntactic structures (e.g., Edwards & Bastiaanse, 1998), which could affect their ability to explicitly express relations, but does not necessarily indicate their inability to properly use these relations. In addition, the grammatical system of Cantonese differs from that of European languages, which makes the procedure less straightforward. Thus, in the present study, the focus of the main event analysis was on the information content, while the relational component was addressed in the discourse structure analysis. After the list of main events and elaborative components was formed based on the analysis of the NBD samples, and its final version was agreed upon by all the authors, the proportion of main events in the aphasic samples was calculated. A list of events can be found in Appendix B. Point-to-point inter-rater agreement for the aphasic samples reached 90%.

### 3.4.7 Statistical analysis

The normality of the residuals for the variables calculated based on the RST annotations was analyzed using Q-Q plots. In the non-brain-damaged group, a ceiling effect was observed on most of the measures. Non-parametric statistical analyses were implemented with non-normally distributed variables.

To investigate the differences in performance of PWA and NBD speakers in the two genres, a set of two-way mixed ANOVAs was administered and subsequent t-tests were carried out for post-hoc analyses. For the scores on the non-normally distributed variables, Mann-Whitney and Wilcoxon signed-rank tests were used to analyze the group difference and the effect of elicitation task, respectively. The significance level was adjusted for multiple comparisons using Holm-Bonferroni correction.

For the rating task, outliers with scores of absolute value greater than two standard deviations from the group means (less than 5%) were removed from the data set. A Mann-Whitney test was used to compare the between-group difference for all the four ratings of coherence as the residuals were not normally distributed. Wilcoxon signed-rank test was used to examine the effect of genre on coherence ratings.

To study potential factors contributing to discourse coherence and its impairment in aphasia, the Kendall tau rank correlation coefficients between the coherence ratings and the 12 micro- and macro-linguistic variables were calculated.

## 3.5 Results

### 3.5.1 Analysis of coherence ratings

#### Group differences

A summary of descriptive statistics of the naive listeners' subjective ratings of coherence is displayed in Table 3.1.

Consistent with previous findings (e.g., Ulatowska et al., 1981), the aphasic group's discourse was rated by a lower understandability, clarity, and completeness scores. The results of a Mann-Whitney test revealed that the difference for understandability and clarity ratings was significant on both discourse tasks (Table 3.2), while completeness was lowered in the aphasic group only in the narratives (story-telling task), but not in the procedural discourse.

Table 3.1: Descriptive statistics of coherence ratings by naive listeners.

	Anomic aphasia		Non-brain-damaged	
	Mean (SD)	Range	Mean (SD)	Range
Story-telling task				
Understandability (1-9)	4.60 (1.60)	1.40-6.70	7.95 (0.54)	7.15-9.00
Completeness (%)	42.67 (28.21)	5.00-95.00	90.95 (14.19)	50.00-100.00
Order of events (%)	86.67 (19.52)	25.00-100.00	98.67 (2.97)	90.00-100.00
Disruptions of clarity (%)	65.00 (29.28)	10.00-100.00	9.42 (10.17)	0.00-30.00
Procedural discourse task				
Coherence rating (from 1-9)	4.71 (1.74)	1.95-6.55	7.33 (0.93)	6.25-9.00
Completeness (%)	50.68 (34.87)	0.00-90.00	79.21 (21.99)	25.00-100.00
Order of events (%)	81.56 (14.45)	60.00-100.00	94.91 (7.34)	73.68 -100.00
Disruptions of clarity (%)	54.96 (28.61)	5.00-100.00	15.21 (16.25)	0.00-60.00

Table 3.2: Comparisons of speaker groups in terms of perceptual judgment by naive listeners.

	Mann-Whitney U =			
	Understandability	Completeness	Order of events	Disruptions of clarity
Story-telling	0.00 *	11.50 *	34.5 *	12.00 *
Procedural discourse	8.50 *	51.5	50.00 *	26.00 *

Note: \*  $p \leq 0.00625$

### Genre differences

Wilcoxon signed-rank test demonstrated that the effect of genre on the ratings was not significant in either speaker group (Table 3.3).

Table 3.3: Comparisons of naive listeners' judgment in the two genres by speaker group.

	Wilcoxon Signed-Rank W =			
	Understandability	Completeness	Order of events	Disruptions of clarity
PWA	104, p=.73	94, p= 0.45	145, p=0.17	139, p=0.28
NBD	160, p=0.051	159.5, p=0.048	151.5, p=0.062	94.5, p=0.46

## 3.5.2 Analysis of micro- and macrolinguistic (RST) variables

### Group differences

The descriptive statistics of linguistic measures for each group in the two elicitation tasks is provided in Table 3.4. The results demonstrated that the performance in non-brain-damaged and aphasic groups differed on all RST measures, except for the percentage of structural expansions, on both tasks (Table 3.4).

The results of the two-way ANOVA and Mann-Whitney test revealed a main effect of speaker group on a number of quantitative measures (Table 3.5). The NBD speakers produced more EDUs, and with a greater efficiency (EDU/min). They also used a larger variety (number) of relations to build discourse structure than the aphasic group. However, as shown on Figure 3.3, the distribution of RST relations of different types in the two

speaker groups was comparable, indicating that people with aphasia demonstrated relatively preserved use of coherence relations. The aphasic participants' samples contained more failed and incomplete EDUs than the samples of non-brain-damaged participants.

The NBD group tended to use more relations of attribution (i.e., complex constructions with direct and indirect speech and cognitive predicates) as well as background, explanation, and elaboration relations than the aphasic group. For storytelling, in both groups, but especially in the aphasic group, the most frequently used RST relation types were those expressing causality, while in procedural discourse, temporal relations predominated.

The aphasic group produced more errors than the NBD group in both narrative ( $U = 55.00$ ,  $z = -2.39$ ,  $p < .01$  and  $U = 19.00$ ,  $z = -4.15$ ,  $p < .001$ , respectively) and procedural discourse ( $U = 50.50$ ,  $z = -2.58$ ,  $p < .01$  and  $U = 32.00$ ,  $z = -3.77$ ,  $p < .001$ , respectively).

Table 3.4: Descriptive statistics of macro- and microlinguistic measures

		Anomic aphasia		Non-brain-damaged	
		Mean (SD)	Range	Mean (SD)	Range
Story-telling task					
RST measures	Efficiency-1st half	22.61 (13.23)	7.50-53.99	37.96 (9.56)	21.43-60.01
	Efficiency-2nd half	22.10 (10.19)	6.79-42.86	36.75 (6.80)	21.43-50.00
	Total number of EDUs	14.60 (5.18)	7.00-26.00	19.20 (4.71)	10.00-26.00
	Number of relation types	11.47 (5.80)	4.00-24.00	17.67 (4.47)	9.00-24.00
	Depth of discourse structure	5.27 (1.79)	2.00-8.00	6.80 (1.42)	5.00-10.00
	% of structural expansions	1.77 (3.93)	0.00-12.50	0.67 (1.76)	0.00-5.00
	% of mononuclear relations	74.34 (11.70)	54.55-100.00	75.11 (11.02)	44.44-90.48
	% of incomplete EDUs	5.38 (10.32)	0.00-36.36	1.59 (3.26)	0.00-11.76
	% of failed EDUs	6.47 (19.25)	0.00-75.00	0.00 (0.00)	0.00-0.00
Micro-ling. measures	% of errors	4.87 (4.44)	0.00-15.38	0.21 (0.56)	0.00-1.69
	% of function words	48.31 (7.99)	34.04-58.67	52.27 (4.76)	47.46-64.29
Procedural discourse task					
RST measures	Efficiency-1st half	22.61 (13.23)	7.50-53.99	37.96 (9.56)	21.43-60.01
	Efficiency-2nd half	22.10 (10.19)	6.79-42.86	36.75 (6.80)	21.43-50.00
	Total number of EDUs	14.60 (5.18)	7.00-26.00	19.20 (4.71)	10.00-26.00
	Number of relation types	11.47 (5.80)	4.00-24.00	17.67 (4.47)	9.00-24.00
	Depth of discourse structure	5.27 (1.79)	2.00-8.00	6.80 (1.42)	5.00-10.00
	% of structural expansions	1.77 (3.93)	0.00-12.50	0.67 (1.76)	0.00-5.00
	% of mononuclear relations	74.34 (11.70)	54.55-100.00	75.11 (11.02)	44.44-90.48
	% of incomplete EDUs	5.38 (10.32)	0.00-36.36	1.59 (3.26)	0.00-11.76
	% of failed EDUs	6.47 (19.25)	0.00-75.00	0.00 (0.00)	0.00-0.00
Micro-ling. measures	% of errors	4.87 (4.44)	0.00-15.38	0.21 (0.56)	0.00-1.69
	% of function words	48.31 (7.99)	34.04-58.67	52.27 (4.76)	47.46-64.29

Note: EDU == elementary discourse unit

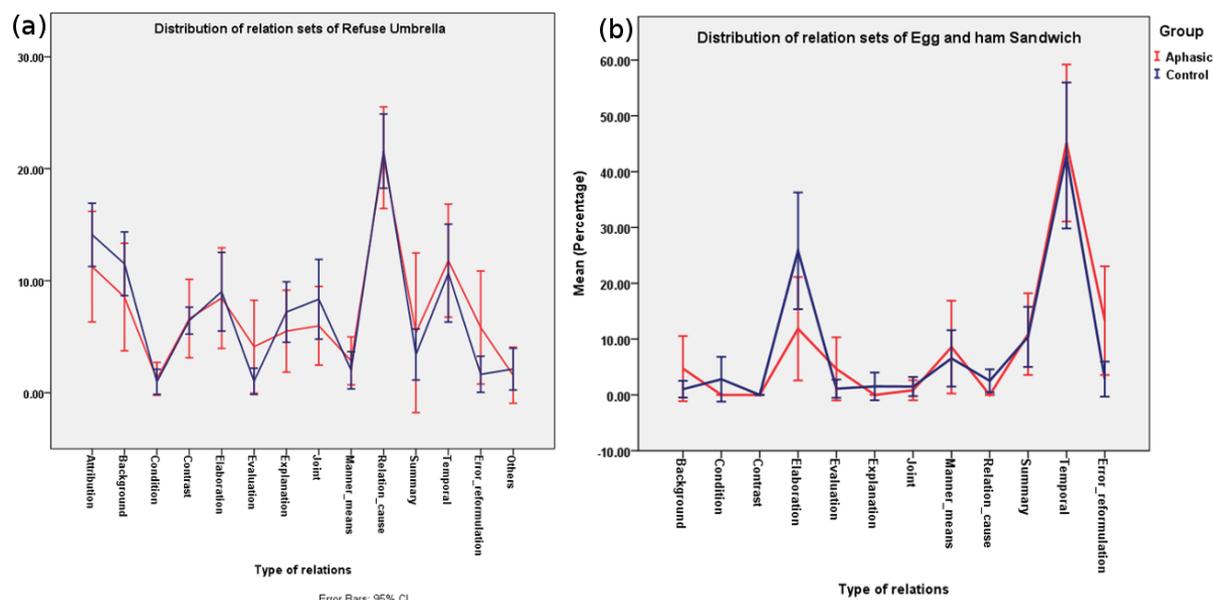


Figure 3.3: Distribution of RST relations in the (a) story-telling and (b) procedural description tasks. (Note: PWA == people with aphasia, NBD == non-brain-damaged people, RBD == people with right hemisphere brain damage)

### Genre differences

There was a main effect of genre on the RST measures (Table 3.5). The story-telling task yielded higher efficiency, that is, faster discourse production in EDUs ( $F(1,58)=28.5$ ,  $p = .0002$ ;  $F(1,58)=7.23$ ,  $p=.01$ ) and a larger number of RST relations than the procedural discourse task ( $F(1,58)=48.56$ ,  $p < .0001$ ). In addition, there was an effect of genre on the type of structures used, namely, mono- versus multi-nuclear relations ( $U=383.5$ ,  $Z=3.106$ ,  $p=.009$ ,  $r=.4$ ). Specifically, procedural discourse contained a higher proportion of mono-nuclear relations. Speakers with aphasia also produced significantly more EDUs ( $T = 4$ ,  $z = -3.19$ ,  $p < .001$ ) and a greater depth of resultant discourse structure ( $T = 2$ ,  $z = -3.19$ ,  $p < .001$ ) in the story-telling task than in procedural description.

### Interaction effects

There was an interaction effect of group and genre on efficiency (Table 3.5). Independent t-tests revealed that the non-brain-damaged participants produced a larger number of EDUs per minute in both the first and second halves ( $t(28) = -3.64$ ,  $p < .001$  and  $t(28) = -4.63$ ,  $p < .001$ ) of the story-telling task than the aphasic group.

## 3.5.3 Correlations between the linguistic variables and coherence ratings

### Macro-linguistic measures

Most of the correlations between measures based on the RST analysis and subjective coherence ratings were significant at the .05 level or higher (Table 3.6). The degree of elaboration, the number of relation types used, and the total number of EDUs produced were correlated with the coherence ratings, whereas no significant correlation was found between the ratings and the proportion of structural expansion. There was an inverse correlation between the proportions of failed and incomplete elementary discourse units and all four subjective ratings.

Table 3.5: Statistical comparisons between performances of aphasic and non-brain-damaged groups.

	Effect of group	Effect of genre	Interaction effect
	Two way ANOVA, F(1, 58)		ANOVA, F(3,56)
Efficiency-1st half	14.81 ***	28.50 ***	258.30 *
Efficiency-2nd half	18.15 ***	7.23 *	220.19 *
Total number of EDUs	11.74 **	14.53 **	10.97, p=0.73
Number of relation types	5.554 *	48.56 ***	22.33, p=0.66
Depth of discourse structure	11.62 **	21.09 ***	14.45 , p=0.59
	Mann-Whitney U	Wilcoxon Signed-Rank W	—
% of incomplete EDUs	340, p=0.17	25, p=0.112	—
% of failed EDUs	330 *	13, p=0.67	—
% of structural expansions	423, p=0.76	13, p=0.57	—
% of errors	105 ***	37, p=0.34	—
% of mononuclear relations	509.5, p=0.76	383.5 **	—

Note: \* $p < .05$ ; \*\* $p \leq .01$ ; \*\*\*  $p \leq .001$ . *P-values were adjusted using the Holm-Bonferroni method.*

Table 3.6: Correlation between macro- and micro-linguistic variables and subjective coherence ratings.

	Kendall's tau			
	Understandability	Completeness	Order of events	Disruptions of clarity
<b>Story-telling task</b>				
Efficiency-1st half	0.425***	0.408**	0.301*	-0.433***
Efficiency-2nd half	0.544***	0.489***	0.384**	-0.484***
Total number of EDUs	0.474***	0.482***	0.291*	-0.307*
Number of relation types	0.570***	0.583***	0.418**	-0.386**
Depth of the resulting discourse structure	0.472***	-0.454***	-0.360*	-0.301*
% of structural expansions	-0.017	0.013	0.066	-0.077
% of incomplete EDUs	-0.032	0.017	0.031	0.161
% of failed EDUs	-0.389**	-0.362*	-0.486**	0.403**
% of errors	-0.709 ***	-0.728***	-0.456**	0.563***
% of function words	0.188	0.096	0.053	-0.059
<b>Procedural discourse task</b>				
Efficiency-1st half	0.263*	0.104	0.131	-0.196
Efficiency-2nd half	0.161	0.212	0.151	-0.147
Total number of EDUs	0.362**	0.306*	0.202	-0.125
Number of relation types	0.344**	0.297*	0.169	-0.137
Depth of the resulting discourse structure	0.431**	0.361**	0.213	-0.289*
% of structural expansions	-0.023	-0.109	-0.062	0.101
% of incomplete EDUs	-0.449**	-0.293*	-0.403**	0.516***
% of failed EDUs	-0.369*	-0.303*	-0.318*	0.403**
% of errors	-0.623***	-0.455**	-0.382*	0.555***
% of function words	0.241	0.161	0.121	-0.133

Note: \* $p < .05$ ; \*\* $p \leq .01$ ; \*\*\*  $p \leq .001$ .

### Micro-linguistic measures

Percentage of word-level errors was significantly correlated with all four coherence measures, especially clarity/understandability in both tasks. No significant correlations were found between the percentage of function words, and coherence ratings.

#### 3.5.4 Information content

In the procedural description, nearly all aphasic speakers mentioned all essential steps, but not the two optional elements (elaborative materials). In the story-telling task, only half of the total amount of essential information was given by the majority ( $\geq 80\%$ ) of the aphasic speakers, and elaborative material was largely omitted in their samples. The results of Mann-Whitney tests indicated that there was a group effect for story-telling ( $U = 119$ ,  $p < .001$ ) but only marginally significant difference for procedural discourse ( $U = 24.5$ ,  $p < .05$ ).

## 3.6 Discussion

The present study pioneered the application of a formal discourse analysis framework, Rhetorical Structure Theory (RST), to the investigation of coherence in spoken discourse in aphasia. Narrative and procedural discourse samples in 15 Cantonese speakers with anomic aphasia and 15 non-brain-damaged (NBD) participants were examined. Discourse structure was annotated in the samples using RST. Then the coherence of the samples was rated by naive listeners. The ratings were contrasted with a number of variables extracted from the RST annotations in order to explore the relation between macro and microlinguistic properties of discourse and coherence.

### Macrolinguistic variables and coherence

According to RST, discourse consists of elementary meaningful information units (EDUs), interconnected by rhetorical, or coherence, relations. Our group of speakers with aphasia produced significantly fewer EDUs, with a lower efficiency (EDU/min) than the non-aphasic group. Discourse structures in aphasic narratives also contained a smaller variety of semantic, or rhetorical, relations, and had smaller depths. These macrostructural differences corresponded to significantly lower coherence ratings, as judged by naive native speakers. These results indicate that the complexity of discourse, expressed by its length, depth, and the variety of relations used to build it, influences the perception of discourse coherence. Importantly, in contrast with previous studies, linguistic complexity is measured at the discourse level, rather than the syntactic level (e.g., Korpijaakko-Huuhka & Lind, 2012; Ulatowska et al., 1983; Ulatowska et al., 1981).

### Microlinguistic variables and coherence

Our results with respect to the relationship between microlinguistic impairments and coherence were inconclusive. The percentage of word-level errors was higher in the aphasic samples. This parameter was found to be correlated with the coherence judgment. In Cantonese, unlike in synthetic, or highly inflected, languages, grammatical and logical connections between words are mostly established through function words and word order. However, no correlation was found between coherence ratings and the percentage of function words. Coelho and Flewellyn (2003) reported a longitudinal study of an anomic speaker's discourse production abilities, in which they found that the participant's

microlinguistic processing improved over time, while macrolinguistic level remained moderately impaired. In the present study, discourse structures in the narratives produced by aphasic speakers were less well-formed than those in non-aphasic ones. The indicators of structural well-formedness – proportions of failed and incomplete EDUs – were correlated with the coherence ratings. Two crucial observations are pertinent to these results. First, these two types of ill-formed EDUs often result from grammatical deficits and lexical access problems. Second, failed and incomplete EDUs are disruptive for discourse structure per se, but also due to the use of repair strategies that they trigger (e.g., reformulations or corrections). Based on these observation and keeping in mind the findings of previous studies (Andretta et al., 2012; Coelho & Flewellyn, 2003), we suggest that micro-linguistic impairments contribute to poor coherence through their negative impact on discourse structure, rather than directly. This hypothesis should be tested in future studies addressing the relationship between micro- and macrostructure and coherence.

### **Information content**

The information content analysis demonstrated that essential thematic information content in the story-telling was reduced in our aphasic group’s narratives, but not in procedural descriptions. These results are consistent with previous studies reporting a reduction of thematic content in narratives (e.g., Capilouto et al., 2006) and preservation of essential information in procedural discourse (Ulatowska et al., 1983) produced by people with different types of aphasia. However, Andretta et al. (2012) found that thematic informativeness, unlike lexical informativeness, remained within normal limits in the picture-elicited narratives produced by anomic speakers.

As the methodology, elicitation tasks, and aphasia types were comparable in the study of (Andretta et al., 2012) and the present work, these inconsistencies in findings can be attributed to the individual variability between aphasic speakers. Another explanation stems from the differences in language structures. Italian anomic speakers in Andretta et al.’s study and our Cantonese aphasic group produced mild grammatical disturbances (see also Bastiaanse et al., 2011; Edwards & Bastiaanse, 1998). Morphosyntactic information is encoded differently in the two languages, which means that the consequences of the grammatical difficulties can differ as well. Christiansen (1995) found that anomic speakers tended to skip whole propositions to compensate for their word-finding problems. While in Italian relations between words are signaled grammatically, in Cantonese the same is done through lexical means and word order. It can be hypothesized that for speakers with anomia, the defining feature of which is impaired lexical access, inflectional morphology was easier to “forgo” than the lexicalized way of expressing morphological meanings in Cantonese. As a result, Cantonese speakers with aphasia reduced language production in general as a coping strategy, while Italian speakers did not. A comparative study of thematic informativeness in languages with different structures and in different types of aphasia could shed more light on the matter.

In the present study, reduced thematic informativeness, corresponded to lower coherence ratings in the aphasic group. Informativeness can thus be considered to be one of the factors potentially influencing coherence perception.

### **Genre-related differences**

As it was mentioned earlier, several factors have been shown to influence cognitive demand of an elicitation task, and, consequently, linguistic performance of aphasic speakers. Our results quantitatively confirmed that there is a difference between macro-linguistic patterns of the two genres (e.g., Ulatowska et al., 1990). That is, procedural discourse

represents an enumeration of steps, or even a listing of nouns in case of verb-production difficulties in the aphasic group, which in structural terms means a simpler structure with equally distributed information. Story-telling involves more complicated relations between different parts of discourse, making the ability to establish mono-nuclear semantic and pragmatic relations essential for producing this type of discourse. Hence, the importance of careful selection of elicitation tasks, be it in clinical or experimental settings, should be stressed once again.

The RST analysis, as compared to a more traditional view on discourse organization in terms of narrative structure (setting, complicating action, and resolution), delivered a more fine-grained representation of the macro-structural patterns of the two genres. It demonstrated that discourse structures in narratives produced by aphasic and non-aphasic speakers differed on a number of parameters, and were not all well-formed. Our results also suggest that factors other than discourse organization may be critical to the perception of coherence. Word-level errors and reduced information content potentially negatively impact coherence. While further investigations of these variables' impact is necessary, the clinical implication of this finding is that discourse-level treatment may not be efficient or complete without additional therapy targeting word finding.

### **Limitations**

The limitations of this study include a relatively limited sample size restricted to a one single aphasia type. The choice of correlation analysis for this study, motivated by the ceiling effects and the number of observations, unlike regression analysis or similar methods, imposes certain constraints on the interpretation of our findings. In addition, variations in language production in different aphasia types may provide additional information to complete the picture drawn in this study. Further investigations building on the current work should extend the investigation to other spoken discourse genres. Despite these limitations, the present findings provide a solid foundation for further studies investigating the contribution of linguistic variables to coherence.

### **Conclusion**

The RST-based coherence analysis presented in this study has offered a novel, fine-grained, systematic way of examining the macro-structural features of oral discourse in aphasia. This is evidenced by the strong correlations between the RST measures and the subjective coherence ratings of narrative production by speakers with anomic aphasia. By using ratings, coherence has been addressed as a perceived quality of discourse, rather than its solely linguistic well-formedness. A number of linguistic variables potentially required for a discourse to be perceived as coherent have been identified. However, the present study did not address the question whether coherence is established through linguistic means alone, or whether comprehension-related variables are involved as well. Future extension of the current investigation should involve a larger number of speakers with different types and severity of aphasia to determine whether coherence is or is not dependent on the micro-linguistic performance of a speaker.

To our knowledge, apart from the work of Kibrik and Podlesskaya (2009) on spontaneous speech in children with neurosis, RST has not been applied to the investigation of clinical data before. Further RST-based analysis of discourse in other adult language-impaired populations, such as individuals with dementia, who have been reported to have problems with producing coherent discourse (e.g., Bourgeois & Hickey, 2011), can inform language production theory as well as therapy. Lastly, we have briefly discussed that aphasic speakers of languages with different grammatical structures may choose different

coping strategies to compensate for their language impairments. This is only one example of the cross-linguistic differences in discourse-level language production. The analysis of discourse structures with RST implemented in the present study offers a possibility for conducting cross-linguistic analyses in order to unveil differences between languages at the macrolinguistic level, but also the general principles of language production, the crucial linguistic mechanisms behind this essential communicative process.

## Chapter 4

# Linguistic mechanisms of coherence in aphasic and non-aphasic discourse.<sup>3</sup>

### Abstract

*Purpose:* ‘Coherence’ is the quality distinguishing discourse from a random collection of sentences. This paper presents a cross-methodological investigation of coherence in discourse of Russian speakers with and without aphasia. The purpose of this study was to determine the linguistic mechanisms responsible for discourse coherence and its impairment in aphasia.

*Method:* Twenty participants were asked to retell the content of a short movie. The retellings were first rated for coherence. Then they were annotated using Rhetorical Structure Theory (RST), an approach formalizing the idea of internal organization of discourse. A number of macrostructural variables were calculated from the annotations. A classification analysis was performed to determine whether the raters’ choice between “coherent” and “incoherent” can be explained based on the RST parameters and a set of microlinguistic variables, and to identify the features contributing to raters’ perception of coherence.

*Results:* The results suggest that listeners’ coherence judgment is directly related to macrostructural well-formedness of discourse, which, in turn, can be affected by a number of microlinguistic variables.

*Conclusions:* In most cases, different aspects of coherence are associated with a number of macro- or microlinguistic features of a discourse. However, our findings also suggest that other, possibly interpretation-related, effects influence coherence judgment.

---

<sup>3</sup>This chapter was adapted from Linnik, A., Bastiaanse, R., Khudyakova, M., & Stede, M., Linguistic mechanisms of coherence in aphasic and non-aphasic discourse. Submitted on June 15th, 2016 (Journal of Speech, Language, and Hearing Research, manuscript JSLHR-L-16-0245)

## 4.1 Introduction

A number of structural-discourse-analysis theories, in the traditions of the formal approach, considered discourse to be a unit of language above sentences. Discourse is a semantic unit, not a grammatical one (Halliday & Matthiessen, 1994). The current study addresses perhaps the most important property of discourse: its internal overall semantic or pragmatic unity, and how it is established. In this paper we present a cross-methodological analysis of spoken discourse in people with aphasia (PWA), a language impairment resulting from stroke, and non-brain-damaged speakers (NBD). The study was designed and implemented to answer the following questions:

1. Is the ability to produce/construct coherent discourse impaired in aphasia?
2. Which linguistic components contribute to coherence?

### 4.1.1 Coherence

The problem of defining coherence has been extensively addressed by a number of researchers (Foltz, 2007; Gernsbacher & Givón, 1995; Kehler, 2002, 2004; Kintsch & Van Dijk, 1978; Sanders & Spooren, 1999; Van Dijk, 1977). Most of the existing definitions are based on the notions of thematic relatedness, topic maintenance (Glosser & Deser, 1991; Halliday & Hasan, 1976), or a general notion of semantic unity, where the parts of discourse “hang together” (Foltz, 2007; Olness & Ulatowska, 2011; Olness, Ulatowska, Carpenter, Williams-Hubbard, & Dykes, 2005; Ulatowska, Olness, & Williams, 2004). One prominent line of discourse analysis defines coherence through the relations (e.g., causal or contrastive ones) between the minimal elements of discourse, its internal organization (Hobbs, 1985; Knott, Sanders, & Oberlander, 2001; Mann & Thompson, 1985, 1988; Marcu, 2000; Polanyi, 1988). In general, coherence can be divided into local coherence, established at the level of pairs of sentences or propositions, and global coherence, referring to the semantic unity of a discourse as a whole. Perhaps the most well-known theory of local coherence, the Centering Theory (Grosz, Weinstein, & Joshi, 1995; Joshi & Weinstein, 1981), suggests that coherence can be modeled based on the relative salience of entities – referents – in a discourse segment and the choice of referring expressions. The present study, however, mainly focuses on global coherence, henceforth referred to as “coherence”.

Connectedness has been argued to be a characteristic of the mental representation of discourse, as opposed to the discourse as a linguistic entity alone (Renkema, 2009; Sanders & Pander Maat, 2006). Furthermore, Gernsbacher and Givón (1995) considered coherence to be a property emerging during speech production and comprehension, which allows a listener to reconstruct discourse as the speaker had it in mind, or the mental representation of it. Hence, the listener’s perspective has to be taken into account when investigating coherence. The current paper describes a study of listeners’ perception of coherence and a set of micro- and macrolinguistic means by which it is established in discourse of aphasic and non-aphasic speakers.

### 4.1.2 Coherence in aphasia

A substantial body of research has documented the disturbances in phonological, lexical-semantic, and syntactic aspects of language production in aphasia (Bastiaanse, Hugen, Kos, & Van Zonneveld, 2002; Blumstein, 1998; Edwards, 1995; Glosser & Deser, 1991, i.a.). Although language production at the discourse level received considerable attention during the last years, coherence remains relatively poorly understood. While some

studies suggested discourse coherence in aphasia is impaired (Andreetta et al., 2012; Christiansen, 1995; Wright & Capilouto, 2012), others reported its preservation (e.g., Glosser & Deser, 1991; Ulatowska et al., 1990). One of the reasons the findings have been inconsistent is the difference in approaches. Specifically, previous studies focused on various aspects of coherence and used different methods to analyze them. The most relevant ones are discussed in this section. More information on previous investigation of discourse production in aphasia can be found in the comprehensive review by Armstrong (2000) and the recent review by Linnik et al. (2016).

Several coherence measures commonly used in language-impairment research and in aphasiology in particular, are based on the notion of topic maintenance, and employ different rating scales. Glosser and Deser (1991) suggested segmenting discourse into utterances and rating every utterance's relatedness to the content of the preceding unit (local coherence) and the overall topic (global coherence) on a five-point scale. The authors reported that their aphasic group obtained coherence scores comparable to those of the NBD group. Because the scores 2 and 4 were rarely used by their raters, a three-point version of Glosser and Deser's (1991) scale was proposed by Van Leer and Turkstra (1999). They studied coherence in speech after traumatic brain injury, but their three-point scale has also been used to measure coherence in aphasic and non-aphasic stroke survivors (Coelho & Flewellyn, 2003; Rogalski et al., 2010). Wright and colleagues (Wright et al., 2013; Wright et al., 2014; Wright et al., 2010) used a four-point scale, also based on the Glosser and Deser (1991) method, which was claimed to be more reliable for the evaluation of coherence in aphasia than the original five-point version.

Whereas Glosser and Deser (1991) style scales evaluated coherence through rating each utterance, in the current study we were interested in coherence as the overall "quality" of discourse, which "can be operationalized only through a single rating of the semantic unity of the whole discourse-situated-in-context, that is, the degree to which a unit [discourse], as a whole, hangs together or makes sense" (Olness & Ulatowska, 2011, p. 1397). Two scales have previously been used to evaluate overall coherence in this way: one by Ulatowska et al. (1981) and one by Armstrong (1987). Armstrong (1987) introduced the notion of "cohesive harmony" and demonstrated that it correlates with listeners' perception of coherence rated on a 4-point scale as the ability to "make sense" of a text. "Cohesion" refers to the local connectivity between sentences and clauses, which is reached through a variety of surface linguistic means, such as coreference, conjunction, and lexical relations (e.g., collocation). Although cohesion and its relation to coherence are outside the scope of this paper, it should be mentioned that in a number of other works, cohesion was studied as a correlate of discourse coherence (Coelho et al., 1994; Halliday & Hasan, 1976; Pehler & Holland, 1984). The three-point scale devised by Ulatowska et al. (1981) addresses the quality of content and clarity as tentative correlates of the overall coherence and cohesion, respectively. By disentangling coherence and cohesion, Ulatowska et al. (1981) took a big step towards a more comprehensive approach to connected-speech analysis in aphasia.

A different approach was suggested by Christiansen (1995), who analyzed the discourse of fluent aphasic speakers from the perspective of propositional coherence. The texts were divided into propositions that were analyzed in terms of coherence violations (information gaps, progression and relevance). Similarly, Marini and colleagues (e.g., Marini, Boewe, et al., 2005; Marini, Carlomagno, et al., 2005; Marini et al., 2008) assessed coherence through the analysis of coherence errors, for example, semantically unrelated and filler utterances. Several multi-level procedures combining word and sentence-level measures, information content, fluency, global and/or local coherence assessment have been developed recently to explore the interaction between linguistic variables involved in discourse production

(Andreetta et al., 2012; Andreetta & Marini, 2015; Marini et al., 2011; Sherratt, 2007; Wright & Capilouto, 2012).

The connection between linguistic variables and coherence perception has hardly been addressed before. Ulatowska and colleagues (Ulatowska et al., 1990; Ulatowska et al., 1983; Ulatowska et al., 1981), for example, reported that the occurrence and length of narrative elements, such as setting, complicating action, and resolution, in aphasic narratives were comparable to those in their control group's samples, whereas aphasic samples were rated as less coherent in terms of content and clarity. The current study taps into the linguistic mechanisms behind coherence. We explore how different micro- and macrolinguistic parameters of discourse relate to the overall coherence ratings. We use Ulatowska et al.'s (1981) scale, and additionally report the results for Van Leer and Turkstra's (1999) three-point scale analysis for comparability reasons.

### 4.1.3 Discourse structure

Discourse in healthy population has been extensively investigated within a variety of frameworks. Ever since the claim was made that natural discourse has a more complex internal organization than a simple linear sequence of utterances/sentences (e.g., Fox, 1987), it has been argued that discourse structure can be formalized (Grosz & Sidner, 1986; Mann & Thompson, 1985, 1988; Marcu, 1998; Miltsakaki, Prasad, Joshi, & Webber, 2004; Moore & Pollack, 1992; Moser & Moore, 1996; Poesio, Patel, & Di Eugenio, 2006; Walker, 1998; Wolf & Gibson, 2005; Wolf, Gibson, Fisher, & Knight, 2005).

#### Rhetorical Structure Theory

Rhetorical Structure Theory (RST) has been introduced as a descriptive theory of natural text organization, "an abstract set of conventions" (Mann & Thompson, 1988, p. 247), but has evolved over time into a rather fine-grained framework of discourse analysis. The basic assumption of RST is that discourse is internally organized into a hierarchical structure rather than a linear sequence (e.g., Fox, 1987; Grosz & Sidner, 1986; Polanyi, 1988). In RST, textual coherence is described in terms of semantic and pragmatic relations, also called "discourse" or "coherence" relations. Discourse consists of elementary units, each representing a complete piece of information, which are connected to each other through coherence relations. The inter-connected elementary discourse units (EDUs) form larger spans that are in turn linked to each other, building up into a tree representation. In every pair of spans, both units are assigned a nuclearity status. In a pair, the so-called nucleus contains the more important part of the discourse and is independently comprehensible, while the satellite either elaborates the information presented in the nuclei, or is less pertinent to the core of the discourse, or the purpose of its author. Consider the construed example in (i) (Fig. 4.1). There are five EDUs in this example. EDUs (2), (4), and (5) are crucial in terms of information content, they are the nuclei in the scheme. Some additional information is presented in the satellites (1) and (3). One level up in the tree, the segment in the span (4-5) can be considered a consequence (i.e., a satellite) of the situation described in the span (1-3), although the nuclearity status of these EDUs may change in context.

(i) (1) *Although she was not hungry anymore,* (2) *Mary decided to try the cake* (3) *her friend made.* (4) *She took a piece,* (5) *but could only eat a half of it.*

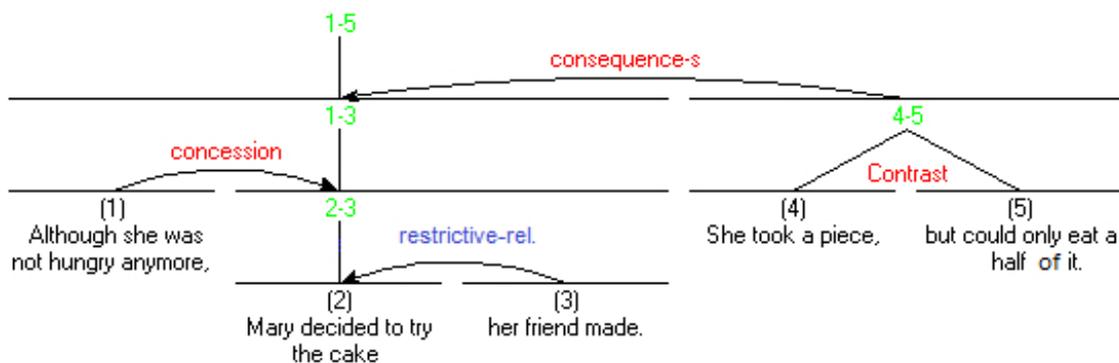


Figure 4.1: Example of an RST tree with a multi-nuclear relation (Contrast), mono-nuclear relations (Concession, Consequence), and an embedded EDU (a restrictive relative clause)<sup>5</sup>

A detailed description of the basic principles of RST can be found in the original papers by Mann and Thompson (1985, 1988), as well as in a series of works by e.g., Taboada (2004), and reviews of Taboada and Mann (2006a, 2006b).

RST is a powerful tool for discourse analysis that has been widely applied to natural language data (e.g., Carlson, Marcu, & Okurowski, 2003; Fawcett & Davies, 1992; Juliano & Cassim, 2012; Kibrik, Khudyakova, Dobrov, & Linnik, 2013; Kibrik, Khudyakova, Dobrov, Linnik, & Zalmanov, 2016; Marcu, 1997; Marcu et al., 1999; Stede, 2004; Stent, 2000; Taboada, 2004). While its primary theoretical application has been discourse-coherence research (e.g., Mann & Thompson, 1988; Redeker, 2000; Sanders, 1997), RST also found practical applications in a number of computational linguistics tasks (see Hovy, 1993; Taboada & Mann, 2006a, for an overview).

RST has been criticized for some of its aspects (cf. Taboada & Mann, 2006b), such as tree-structure representation (Miltakaki et al., 2004; Wolf & Gibson, 2005), relational ambiguity, and possibility of multiple analyses (e.g., Moore & Pollack, 1992; Stede, 2008). The latter issue derives from the more global problem of discourse ambiguity: just like a word or a sentence can be ambiguous, a piece of discourse may have multiple interpretations. Certain disambiguation solutions have been suggested to minimize this problem (Marcu, 1998; Stede, 2008). The criticism of the tree-shaped structure focuses on the following restriction: only one link and one relation can be chosen to establish the connection between a segment and the rest of the discourse, while in some cases, more than one such connection can be recognized. Due to this restriction, RST analysis does not reflect some of the connections simultaneously present between different discourse segments (Marcu et al., 1999; Wolf & Gibson, 2005). However, the advocates of the tree-structure representation argue that under-constrained representations lead to an overload of links, often doubling each other, and resulting in the possibility of linking any two discourse segments due to the “the human tendency to interpret texts as coherent even when they are not” (Egg & Redeker, 2010; Marcu, 2003).

Despite the above-mentioned objections, there were two major reasons for choosing RST to study discourse in aphasia. First, it offers a detailed annotation scheme – a set of clearly formulated discourse-tagging principles (Carlson & Marcu, 2001) which has been previously used in creating a large corpus of written annotations (Carlson et al., 2003). Den Ouden, van Wijk, Terken, and Noordman (1998) reported that the inter-annotator agreement for RST was higher than for a less detailed intention-based analysis proposed by Grosz and Sidner (1986), possibly as a result of the more explicit definitions of relations in RST. Second, although RST is not an explanatory theory of discourse

<sup>5</sup>The “restrictive-rel” relation is not a part of the original RST set, but it is included in the extended set of relations devised by Carlson and Marcu (2001).

production (Mann & Thompson, 2001), it is a widely applied descriptive framework containing a number of cognitively plausible aspects (e.g., Taboada & Mann, 2006a). For example, according to RST's basic assumptions about discourse structure, every part of a coherent discourse should have a place in the structure. This requirement appeals to Grice's maxims of quantity and relevance and thus reflects the same idea that is behind propositional content or analyses of coherence violations (e.g., Christiansen, 1995; Marini et al., 2011), in which irrelevant propositions are penalized. Another example is the restriction to choose a single most plausible analysis among the existing possibilities. It was introduced because of the ambiguity at discourse level, not uncommon at other linguistic levels. It mimics the natural process of listeners' selection of a single interpretation during discourse processing (e.g., Mann & Thompson, 1988). The fact that manipulations of discourse structure can result in different interpretation of a discourse further supports the argument that discourse structure plays an important role in discourse processing (Sanders, 1993). RST provides an apparatus for the analysis of alterations of discourse structure in aphasia. Specifically, it is possible that coherence impairment in aphasia is a reflection of discourse-structure alterations (e.g., disintegration), which results from the deficits at different linguistic levels experienced by aphasic speakers. Spontaneous speech, as opposed to written or preplanned oral discourse, provides access to studying the immediate cognitive processes, such as inferencing, attention allocation, and reasoning, related to language production. Analyzing spoken discourse can be very informative in terms of the principles underlying language organization. Structural organization of healthy, spoken discourse has been explored in a relatively small number of studies (Antonio & Cassim, 2012; Daradoumis, 1996; Kibrik & Podlesskaya, 2009; Nakatani, Hirschberg, & Grosz, 1995; Taboada, 2001, 2004). Although the original RST was developed for written text, and lacked the means to represent underdeveloped structures, such as corrections, paraphrasing, repetitions (Antonio & Cassim, 2012), it was demonstrated that RST can be implemented in spoken-discourse analysis (Daradoumis, 1996; Taboada, 2004). Moreover, Kibrik and Podlesskaya (2009) showed that RST can be used to investigate narrative-structuring abilities in populations with atypical speech production.

#### 4.1.4 The present study

We hypothesize that aphasic speakers are able to reach understandability despite microlinguistic deficits because their ability to maintain the core structure in their discourse is preserved. However, because of multiple disfluencies, or "gaps", in their discourse structures, their narratives are perceived as less coherent. The term "disfluencies" refers to the occurrences of false-starts, repairs and reformulations, and unintelligible parts, which potentially disrupt discourse flow. As stated earlier, we consider coherence to be a product of both speaker and listener perspectives, and the global meaning to be co-constructed by both discourse organization by a speaker and its perception by a listener. Ulatowska et al. (1981) among others hypothesized that "there should be some relationship between the objective measures of sentence or discourse structure, and the rating of the scorer" (p. 367), but postponed investigation of this issue. The focus of our study is the investigation of the relationship between different linguistic parameters of discourse and the perception of coherence in terms of overall coherence ratings. A number of micro- and macrolinguistic parameters, such as word-level error rate, type-token ratio, and proportion of structural disfluencies and main events (see Table 4.2), were collected from a set of semi-spontaneous story retellings by people with and without aphasia, which were annotated for discourse structure with RST (cf. the "Methods" section). Coherence as a

perceived quality of discourse was measured using the scale developed by Ulatowska et al. (1981). For comparability purposes, the samples were also evaluated using the 3-point version of Glosser and Deser’s (1991) scale (Van Leer & Turkstra, 1999). Classification analysis was performed to test whether coherence ratings could be explained by micro- and/or macrolinguistic variables, and to identify which of the variables were the most informative.

## 4.2 Methods

### 4.2.1 Participants

Ten people with aphasia resulting from stroke (mean age 56.4, range 40-73; 6 male and 4 female) and ten non-brain-damaged native Russian speakers (58.7, range 42-84; 5 male and 5 female) participated in the study. The groups were matched for age (8 pairs with  $\pm 5$  years, 2 pairs with 10-11 years of age difference). All aphasic participants were recruited at the Center for Speech Pathology and Neurorehabilitation in Moscow, Russia. Five of the aphasic participants were diagnosed with fluent, five others with non-fluent aphasia using Luria’s Neuropsychological Investigation (Luria, 1966a), the severity level varying from mild to severe, and time post onset ranging from 3 months to 7 years and 9 months. The non-brain-damaged participants were recruited at a national research institution in Moscow and through personal contacts. None of them had been diagnosed with neurological impairments or psychiatric disorders. All of the control participants were highly educated, whereas the educational background in the aphasic group varied from secondary (in one case incomplete) to higher education. The individual characteristics of aphasic participants are provided in the Appendix C. The participants signed an informed consent form that was formulated according to the Declaration of Helsinki.

### 4.2.2 Materials

The Pear Film (Chafe, 1980), a six-minute silent movie, was used to elicit spoken discourse samples. Following Chafe’s original procedure, the participants were instructed to tell to someone who had not seen the movie before what happened in the movie. The audio recordings of the retellings have been added to the Russian Clinical Pear Stories corpus (CliPS). Their discourse samples were transcribed using the Codes for the Human Analysis of Transcripts (CHAT) format (MacWhinney, 2000). The information about the number of words, utterances, EDUs and the duration is provided in Table 4.1.

Table 4.1: Results of the quantitative analysis of the discourse samples produced by aphasic (PWA) and non-brain-damaged (NBD) participants of the current study. Means (ranges).

	PWA	NBD
Number of words	234 (124-509 )	321 (146-551)
Number of utterances	28.7 (15-50)	29.4 (9-56)
Number of EDUs	43.4 (21-104)	62.7 (24-98)
Time in minutes	6.95 (2.4-18.4)	3.1 (2.1-4.97)

*Note: EDU = elementary discourse unit; PWA == people with aphasia, NBD == non-brain-damaged people*

### 4.2.3 Procedure

#### Coherence ratings

The rating scale developed by Ulatowska et al. (1981) contains four questions assessing the quality of content (Q1-Q4), and one for clarity of language (Q5). Both content and clarity were rated on a three-point scale.

- (Q1) Do you know what is happening in the story?
- (Q2) Does the sequence of events make sense?
- (Q3) Is the story unambiguous as to what each of the participants did?
- (Q4) Is the story accurate in terms of the stimulus?
- (Q5) Clarity

According to the original procedure suggested by Ulatowska et al. (1981), ratings on all the questions addressing content (Q1-Q4) should be added up to form a final score. This resulting score may, however, be difficult to interpret, because every parameter assessed by these questions receives the same “weight” in the final score. It has been argued that referential cohesion is not a pre-requisite for appropriate coherence (e.g., Glosser & Deser, 1991; Keenan et al., 1984; Ulatowska et al., 1983; Ulatowska et al., 1981). Hence, the variables may not equally contribute to the overall coherence. Arguably, the concept of clarity already implies an appropriate sequence of events (Q2), and accuracy with respect to the stimulus (Q4) may not be essential for a retelling to be coherent. Considering these arguments, in the current study we focused on Content (Q1) and Clarity (Q5) as representative correlates of coherence.

Two five-point scales, one for Understandability (q1) and one for Connectedness (q2), were added to the original scale.

- (q1) How understandable is the retelling you’ve listened to?
- (q2) How connected is it?

Adjusting task instructions is not a new idea (e.g., Wright & Capilouto, 2009), and we expected that in this particular case, using these two additional questions result in a better interpretability of the outcomes. That is, in addition to asking if one can make sense of a story at all, we wanted to know how understandable it was. If the first question focuses on the ability of a listener to interpret what he or she heard, the second one rather is about the well-formedness and appropriateness of the product of a speaker’s effort.

#### Inter-rater reliability

All the samples were rated independently by two of the authors (AL and MK). There is an extensive discussion about the correct way to calculate inter-rater agreement in linguistics (e.g., Artstein & Poesio, 2008). Percent agreement is arguably a rather unreliable indicator of inter-rater reliability, especially with increasing number of categories. However, it is still used in many studies, and for comparability purposes it is noteworthy that inter-rater percent agreement for Ulatowska’s (1981) and Glosser and Deser’s (1991) rating scales in this study reached 89%. The results of three chance-corrected inter-rater agreement measures indicated a high level of agreement between the raters on all questions (ICC .85-.95, Krippendorff’s alpha .77-.85, Cohen’s kappa .8-.9) but Q2 and Q4 (around chance-level agreement). After the inter-rater reliability scores were calculated, all disagreements were resolved in a discussion, and a single, overall coherence score for every sample for each scale was agreed upon.

### Segmentation

The samples were segmented into *elementary discourse units*, following the RST annotation guidelines (Carlson & Marcu, 2001; Kibrik & Podlesskaya, 2009) and the CHAT manual (MacWhinney, 2000). The segmentation principles were based on syntactic, semantic, and phonologic criteria resembling the ones used by Marini et al. (2011). The samples were also segmented into *utterances* – also called “verbalizations” (Glosser & Deser, 1991), communicative units (C-units; e.g., Armstrong et al., 2011; Wright & Capilouto, 2009, 2012) or terminal units (T-units; e.g., Coelho et al., 1994; Hunt, 1970; Van Leer & Turkstra, 1999) – for the Glosser and Deser (1991) rating procedure. An utterance consisted of a main clause and all subordinate clauses or non-clausal structures attached or embedded in it. Unlike an utterance, an EDU “encodes a *minimum* of meaning and/or discourse function interpretable relative to a set of contexts” (Stede, 2011, p. 87), hence, for example, clausal complements and infinitival clauses (cf. Carlson & Marcu, 2001) were treated as separate EDUs. In the example on Figure 4.2(A) below, three EDUs in the (3-5) span constitute a single utterance.

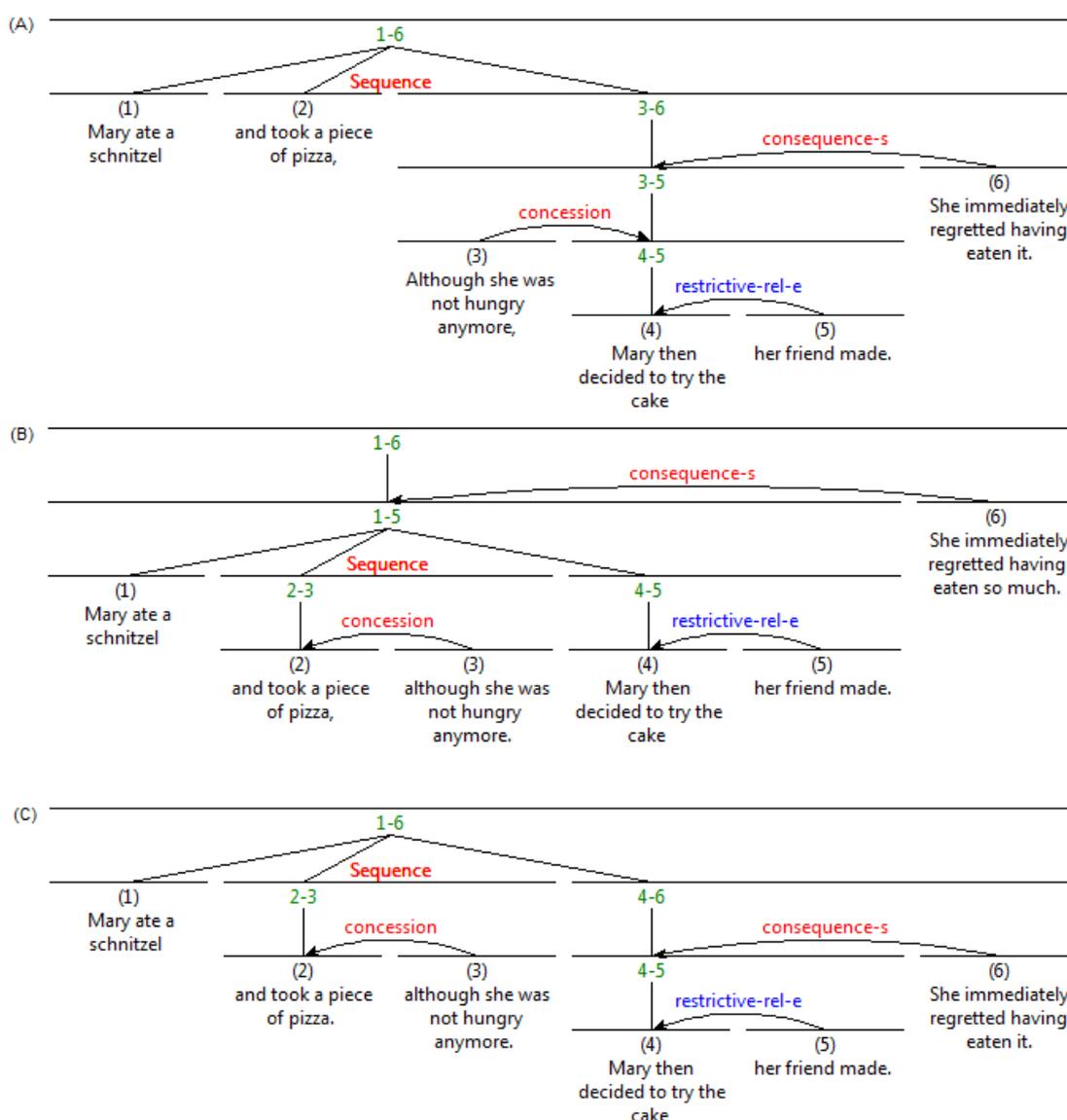


Figure 4.2: Example of discourse structures demonstrating different attachment points (A-B) and different constituents (A-C).

### Reliability of segmentation

The segmentation was performed by two of the authors (AL and MK). 10% of the samples from both groups were segmented independently ( $N=2$ ,  $NEDU=93$ ). The inter-coder agreement, calculated as the proportion of segment boundaries identified by both coders out of the sum of all independent boundary placement decisions made by the two coders, was found to be 90%. Occasional near-misses in boundary placement with a difference of one word were ignored, as it has been noted that coders often agree on the core of a segment, but tend to place exact boundaries differently (Artstein & Poesio, 2008). The segmentation of the remaining samples was performed by one of the coders, and verified by the other coder. All disagreements were resolved in a discussion prior to further analyses.

### Discourse structure annotation and inter-annotator agreement

Discourse structure annotation was performed by two of the authors (AL and MK), who had prior experience with RST annotation, using the annotation software RST-Tool (Marcu et al., 1999; O'Donnell, 2000a), with the extended set of discourse relations suggested by Carlson and Marcu (2001). Ten percent of the samples from both groups were excluded from the further analysis and used to calculate inter-annotator agreement (cf. Results section below). Estimating reliability of discourse-structure tagging is not straightforward (Artstein & Poesio, 2008; Iruskieta et al., 2015; Marcu et al., 1999, , i.a.). Two methods have been suggested to compare rhetorical structures, a quantitative (Marcu, 2000), and a qualitative one (Iruskieta et al., 2015). Following Iruskieta, da Cunha and Taboada, the inter-annotator agreement was assessed in terms of the four types of decision involved in the annotation, that is, the choice of a discourse relation; the attachment point of the EDU; the constituent which it is attached to; and nuclearity. In the example on Figure 4.2, EDU (3) has different attachment points (central component of a span) in (A) and (B), whereas constituent (6) has the same attachment point in (A) and (C), but the constituent it is attached to is different, (4-5) and (3-5) respectively.

The inter-annotator agreement on the two samples including 74 EDUs and 296 annotation decisions was 77%. After a discussion of the resulting annotations, the remaining discourse samples were annotated by one of the annotators and subsequently validated by the other annotator. All disagreements were resolved in a discussion. The annotations were reexamined several weeks later, minor mistakes were corrected. Our goal was to obtain a single, most plausible annotation for each sample, which could be used for further analyses. The described annotation procedure was thus considered acceptable for the purposes of the current study, since RST has not been extensively used with language-impaired data. Although the golden standard of discourse tagging is the double-blind annotation, at this stage we decided against it. While this strict procedure is important for verifying construct validity, before this step is taken, the annotation principles used in this study should be tested and validated on different data sets.

### Linguistic parameters

Microlinguistic measures consisted of a number of lexical and morphological parameters (Table 4.2). Type-token ratio, the most commonly used measure of lexical diversity, was calculated separately for lexical verbs and nouns. This distinction was made to account for the differences in verb and noun usage demonstrated by aphasic speakers (e.g., Bastiaanse et al., 1996; Bastiaanse & Jonkers, 1998). Mean length of EDU in words was calculated (similarly to the mean length of utterance) as a measure of sentence elaboration. Word-level phonological, semantic, and morphological errors were collapsed

into a single parameter, the ‘error rate’. Additionally, the proportion of fillers (e.g., “em”, “hm”), which are often used in place of pauses, and empty words (e.g., “like”, “what’s its name”, “you know”), frequently produced when word-finding problems arise, was included. Ungrammatical EDUs were defined as discourse units lacking grammaticality and/or essential information content. Ungrammatical EDUs were included in the microlinguistic set of variables as they result from linguistic difficulties.

In the macrolinguistic set, two types of reasons for the ill-formedness of discourse structure were defined: incomplete EDUs and structural disfluencies. An “incomplete EDU” was defined as a clause with a missing obligatory syntactic constituent. Based on the prosodic contour and the context, the phrase in the example (ii) is an incomplete EDU (the lengths of filled and empty pauses are specified in parentheses, missing pieces of information are added in square brackets).

- (ii) такие вот м с- нет не сумки а большие такие как  
 these well m b- no not bags but big these like  
 ‘[There were] big bags, oh no, not bags but big, like [baskets]’

Incomplete EDUs were considered macro-linguistic events and, thus, included in the macrolinguistic variable set, because, for example, an omission of a subject or an obligatory object can result in disrupted referential coherence. Example (ii) demonstrates that incomplete EDUs remain understandable, that is, they can be correctly interpreted by listeners.

Disfluencies included unintelligible or meaningless, often also ungrammatical EDUs disrupting the narrative line (e.g., (2) in Example (iii)), reformulations, retracings, false-starts, and corrections.

- (iii)  
 (1) А девочка эм а ре-э мальчик э шляпа у мальчика м упала  
 and girl em and de-eh boy eh hat of boy m fell  
 ‘And the girl... em... the boy’s hat fell’

- (2) и э э кажется что э ну девочка м ну э м ну м эм  
 and eh eh seems like eh well girl m well eh m well m em  
 ‘...and it seems like the girl... well, em...’

- (3) глагол не могу подобрать  
 verb not can find  
 ‘[I] can’t find the right verb’

Several other parameters collected from the RST annotations were added to the macrolinguistic set. Production measures included length in EDUs; depth of the resulting RST structure, or the maximal number of levels (edges) in the tree structure counted from the root to an EDU; and the number of different types of relations used in building the RST structures (the size of relation set). The proportion of multinuclear relations (structures with two or more EDUs equally central to the speaker’s purpose in a discourse) in the set was used to capture differences in use between the two types

of structures, multi- and mono-nuclear. In multinuclear relations, such as Sequence or Contrast, both EDUs contain equally “important” information, as opposed to one EDU elaborating or complementing the other in mononuclear relations. At the level of information structure, multinuclear constructions do not involve manipulations such as assigning prominence. This variable was included because this type of structures could thus be easier for participants who experience difficulties producing coherent discourse. The comments speakers made outside of the story, for example, about their discourse abilities, or to share a related memory, were called “meta-comments”. A separate relation was introduced to mark meta-comments in the RST trees for the purposes of this study, and the proportion was factored in the model. Attributions – the EDUs reflecting the use of cognitive predicates and reported or direct speech in discourse – potentially also have an effect on overall perception of discourse (e.g., Groenewold, Bastiaanse, Nickels, & Huiskes, 2014). Embedded EDUs were included as a measure of elaboration, as they usually are relative clauses containing additional elaborative information (Carlson & Marcu, 2001). Furthermore, the “Number of main events” variable developed by Wright et al. (2005) was used to account for the information content in the samples. Comparing the events included by all the participants in the NBD group identified a set of main events in the stimulus story, and the final set was agreed upon by two of the authors. A raw number was included as a variable in the analysis (in the original paper a proportion was used).

Table 4.2) summarizes the linguistic variables used to analyze the samples. They are divided into three sets for the classification analysis.

Table 4.2: Linguistic variables used for the analysis at the micro- and macrolinguistic levels.

Set	Features
I. Microlinguistic variables	<ol style="list-style-type: none"> <li>1. Error rate (% or word-level errors)</li> <li>2. Ungrammatical EDUs (%)</li> <li>3. TTR for verbs</li> <li>4. TTR for nouns</li> <li>5. Mean length of EDU (in words)</li> <li>6. Fillers (rate per EDU)</li> </ol>
II. Macrolinguistic variables	<ol style="list-style-type: none"> <li>1. Total number of EDUs</li> <li>2. Relation set size (N of relations of different types used to build discourse structure)</li> <li>3. Depth of discourse structure</li> <li>4. Embeddings (%)</li> <li>5. Multinuclear relations (proportion)</li> <li>6. Meta-comments (%)</li> <li>7. Incomplete EDUs (%)</li> <li>8. Disfluencies in EDUs (%)</li> <li>9. Number of main events (information content)</li> <li>10. Attributions (%) (cognitive predicates/direct speech)</li> <li>11. Word-finding comments (%)</li> </ol>
III. All variables	All macro- and microlinguistic variables

*Note: TTR = type-token ratio; EDU = elementary discourse unit*

## Data analysis

Experimental datasets often suffer from ceiling/floor effects, non-normal distributions, and relatively small numbers of observations. All of this applies to the data available for this study. Our goal was not to investigate the interactions and effects of different pairs of factors, but their ability to function as a combination in explaining coherence. The analysis was performed using the random forest (RF; Breiman, 2001) classifier, a machine learning technique based on random sampling of observations and features, or variables (bootstrapping). It has several advantages crucial for this study. It does not have any formal distribution assumptions (i.e., it is non-parametric), it works on small data samples (e.g., Dettling & Bühlmann, 2002; Díaz-Uriarte & De Andres, 2006; Pomeroy et al., 2002), and it is generally robust against overfitting (Breiman, 2001; Segal, 2004). RF also offers an internal mechanism for the evaluation of the variables' contribution to classification, called "feature selection".

## 4.3 Results

### 4.3.1 Ratings

#### Correlations between ratings

Spearman's rho ( $\rho$ ) coefficient was used to estimate correlations between the rating scales, Kendall's tau ( $\tau_b$ ) is also provided for ordinal scales. The results are presented in Table 4.3. Whereas no strong correlations were found between Glosser and Deser's (1991) global coherence scores and other ratings, local coherence scores were strongly correlated with Clarity and Connectedness.

Table 4.3: Correlations between different coherence rating scales.

Coefficient	I	II	III	IV	V
I. Global coherence	$\rho$				
II. Local coherence	0.34	P			
III. Content	0.44	0.35	$\rho / \tau_b$		
IV. Clarity	0.24	0.72**	0.62/0.59*	$\rho / \tau_b$	
V. Understandability	0.31	0.49	0.63/0.59*	0.84/0.76**	$\rho / \tau_b$
VI. Connectedness	0.36	0.71**	0.60/0.54	0.83/0.74**	0.75/0.66**

#### Between-group comparison

*G<sup>ℓ</sup>D scale.* Following the procedure described by Van Leer and Turkstra (1999), the percentages of high, medium, and low local and global coherence ratings were calculated. A one-way ANOVA was conducted, and a Bonferroni adjustment was applied to the p-values ( $.05/3 = .017$ ) for each of the measures. Global coherence levels in people with aphasia did not differ significantly from those of non-brain-damaged participants ( $F(1,18)=.5964$ ,  $p=.45$  for high,  $F(1,18)=2.043$ ,  $p=.17$  for medium,  $F(1,18)=2.043$ ,  $p=.17$  for low coherence scores). The difference in local coherence was not significant ( $F(1,18)=7.97$ ,  $p=.01126$  for high,  $F(1,18)=4.586$ ,  $p=.04617$  for low scores, Mann-Whitney  $W=30$ ,  $p=.14$ ,  $r=-.33$ , Mdn .115 in NBD, 147 in PWA for medium scores). These findings are consistent with the results reported by Glosser and Deser (1991).

*Overall coherence ratings.* The scale developed by Ulatowska et al. (1981) extended with two additional items was used to evaluate overall coherence of the samples. Due to the ceiling effect in the control group, Mann-Whitney tests were conducted to compare the

ratings between the two groups. Aphasic speakers received lower ratings than participants without aphasia on all scales (Content:  $U=124$ ,  $p=.16$ ,  $r=.37$ ; clarity:  $U=37.5$ ,  $p < .0001$ ,  $r=.77$ ; understandability:  $U=26$ ,  $p < .0001$ ,  $r=.78$ ; connectedness:  $U=57.5$ ,  $p < .0001$ ,  $r=.63$ ).

### 4.3.2 Classification analysis with random forests

The ratings were transformed into binary variables, that is, the samples from both groups were labeled as coherent/incoherent. After that the random forest classifier was trained with different combinations of linguistic variables. Default values were used for the  $m_{try}$  parameter, indicating the number of randomly sampled variables, and  $n_{tree}$ , number of decision trees to grow. The model was run with several different random seeds, leave-one-out cross-validation was applied. The results for classification are shown in Table 4.4. The table presents the classification accuracy (ACC), kappa statistic, and the area under the ROC curve (AUC) (see e.g., Fawcett, 2006; Mitchell, 2011), which in the current study can be interpreted as the probability that the model will correctly classify a pair of randomly picked observations as “coherent” and “incoherent”.

For variable importance estimates, built-in feature selection was cross-validated through selection by filter (SBF) method. A description of the feature-selection mechanisms can be found in the documentation of R packages *randomForest* (Liaw & Wiener, 2002) and *caret* (Kuhn, 2008). Figure 4.3 presents an example outcome of the built-in feature selection for the classification based on macro-structural variables. The results of the feature selection are summarized in Table 4.5.

Table 4.4: Results of the group and coherence ratings classification using random forests and different sets of predictors (variables).

Set of variables	ACC	Kappa	AUC
Content			
All	0.72	0.4	0.77-0.85
Microlinguistic	0.66-0.72	0.3-0.4	0.75-0.76
Macrolinguistic	0.77	0.5	0.84-0.88
Clarity			
All	1	1	1
Microlinguistic	0.94	0.88	0.97
Macrolinguistic	0.94	0.88	0.98
Understandability			
All	0.94	0.88	0.98-1.0
Microlinguistic	1	1	1
Macrolinguistic	0.83	0.6	0.87-0.88
Connectedness			
All	0.94-1.0	0.88	0.88-0.1.0
Microlinguistic	0.83	0.6	0.86
Macrolinguistic	0.94-1.0	0.88-1.0	1
Group			
All	0.94-1.0	0.88-1.0	0.97-1.0
Microlinguistic	0.94-1.0	1	1
Macrolinguistic	0.88	0.77	0.87-0.9

Note: ACC = (classification) Accuracy; AUC = Area Under the ROC Curve

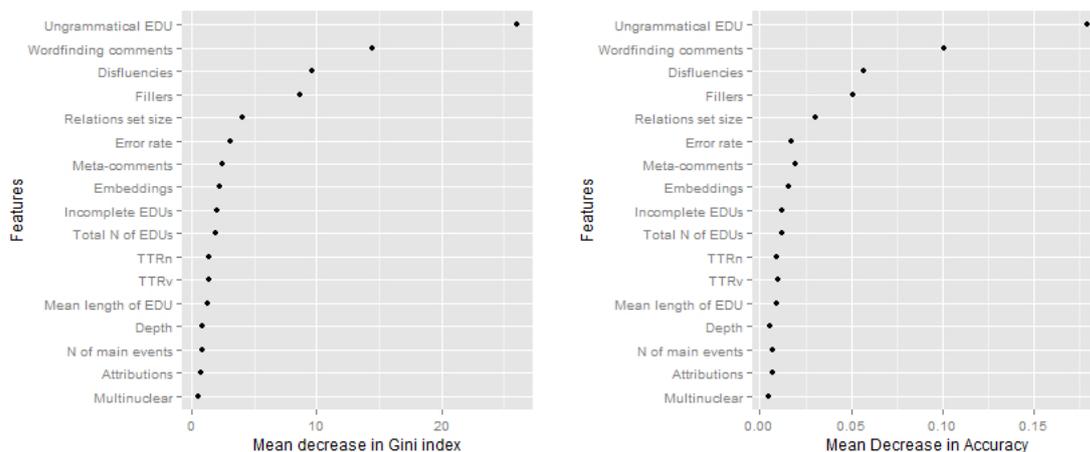


Figure 4.3: Random forest built-in feature selection plot.

*Predicting Group.* The distinction between the PWA and NBD was highly accurately predicted based on the microlinguistic variables (set I). This result is not surprising, since many aphasia assessment and classification tools include microlinguistic measures, and previous studies reported differences between PWA and NBD performance on the microlinguistic level (e.g., Andretta & Marini, 2015).

*Predicting Coherence ratings.* Although high and low coherence groups roughly correspond to the NBD/PWA groups respectively, the correspondence is not one-to-one, as shown in Figure 4.4.

Table 4.5: Variable importance for each set of predictors. (I – microlinguistic, II – macrolinguistic variables, III – all variables; A tick mark ✓ indicates the importance of a variable in the first column for predicting the corresponding dependent variable in the header.)

Set	Variable	Group			Content			Clarity			Understandability			Connectedness		
		I	II	III	I	II	III	I	II	III	I	II	III	I	II	III
Micro-ling. variables	Error rate	✓		✓				✓		✓	✓					
	Ungram. EDUs				✓		✓	✓		✓		✓		✓		✓
	MLUw			✓												
	Fillers				✓						✓		✓	✓		
Macro-ling. variables	Disfluencies		✓	✓		✓	✓		✓	✓		✓	✓		✓	✓
	Incomplete EDUs		✓	✓					✓						✓	
	Relation set		✓	✓								✓				
	N main events					✓	✓									
	Wordfinding comments		✓						✓	✓		✓	✓		✓	✓

Note: EDU == elementary discourse unit, MLU == mean length of utterance

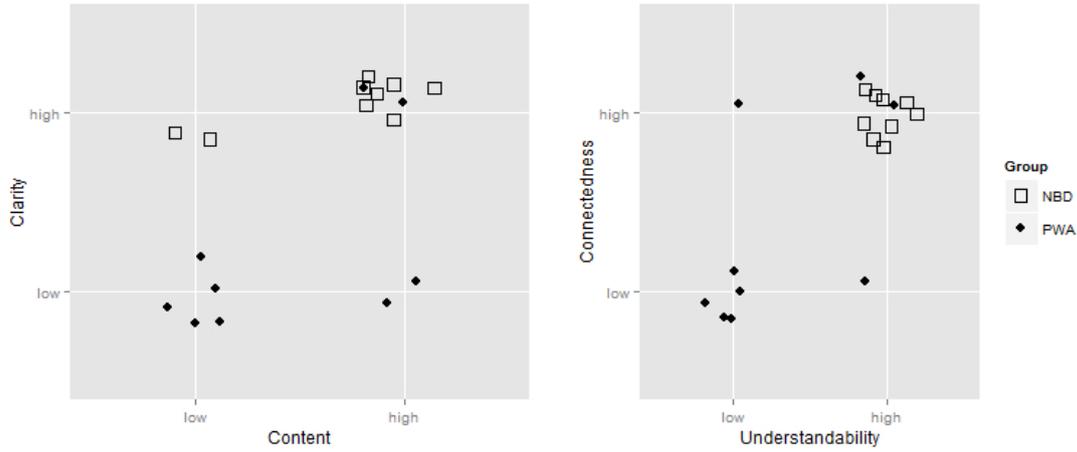


Figure 4.4: Distribution of ratings in the NBD and PWA groups.

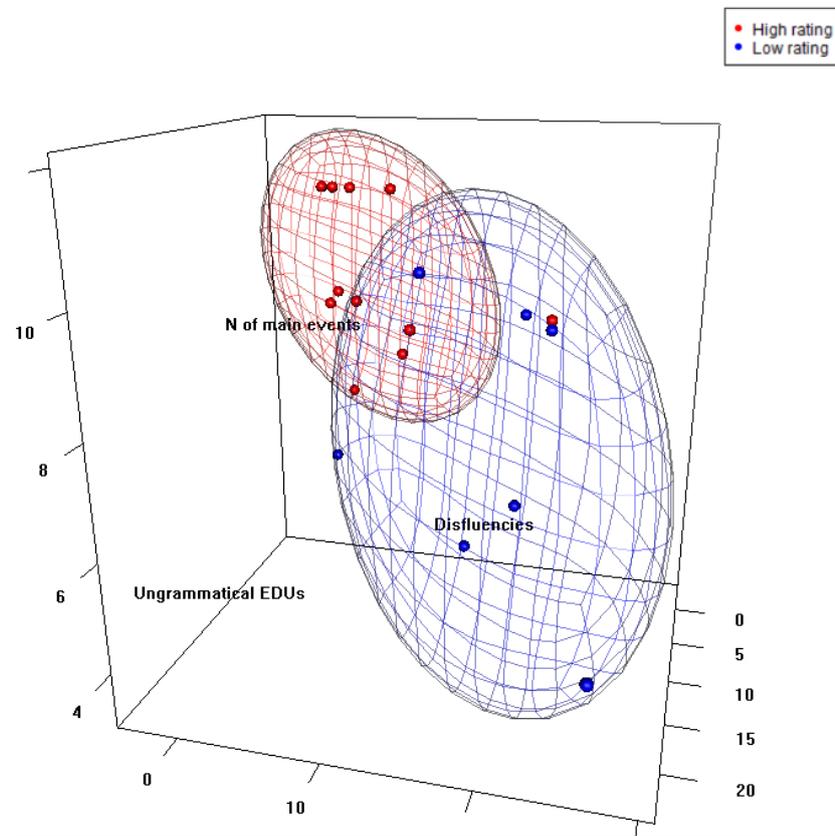


Figure 4.5: Distribution of Content ratings based on top 3 selected features.

Classification for Content was most accurate when the macrolinguistic variables (set II) were used. Feature selection demonstrates that the number of main events is a crucial predictor for this rating. Figure 4.5 illustrates the distribution of Content ratings with respect to the top-three selected linguistic variables. The overlap between the two classes observed in Figure 4.5 explains the decrease of classifier’s performance with the Content scale.

Table 4.6 provides two examples of observations problematic for classification in the class overlap. One of the samples received high Content ratings (19), the other one (7) was rated as less coherent, despite the fact that the analysis of selected features

shows that (20) contains substantially higher proportions of incomplete EDUs and error rate. This example demonstrates that the Content ratings can hardly be fully explained based on linguistic variables alone. Although such cases are scarce, they are the reason classification into “coherent” and “incoherent” is not a hundred percent accurate.

Table 4.6: Examples of cases problematic for classification into low/high coherence groups.

Participant	Group	Error rate	Ungrammatical EDUs	Fillers	Relation set	Incomplete EDUs	Disfluencies	N main events	Word-finding comments	Content rating
7	NBD	1.7	3.4	5.1	16	0	18.6	9	0	Low
19	PWA	26.1	4.5	1.8	19	12.6	20.7	9	3	High

*Note: PWA == people with aphasia, NBD == non-brain-damaged people*

Clarity, Understandability, and Connectedness ratings were more accurately predicted than Content, which indicates that the perception of these coherence aspects correlates is strongly linked to the linguistic parameters of discourse. The best accuracy for the Clarity ratings was reached with both micro- and macrolinguistic variables. While micro-linguistic variables were more informative for the Understandability classification, the perception of Connectedness was better explained through macrolinguistic parameters of the samples. The variable-selection procedure identified the proportions of ungrammatical EDUs, disfluencies, as the most informative predictors for all the scales. Additionally, for every scale different variables were selected. Thus, the proportion of word-finding comments and the word-level error rate were important for Clarity, whereas syntactic and referential well-formedness reflected by the percentage of incomplete EDUs was relevant for Connectedness.

## 4.4 Discussion

The present study focuses on the relationship between perceptual ratings of discourse coherence and various micro- and macrolinguistic parameters of discourse. Different linguistic variables have been reported to have an effect on or to be correlated with coherence in aphasic discourse (e.g., Andreetta & Marini, 2015; Marini et al., 2011; Wright & Capilouto, 2012). Most of the global coherence measures used in previous studies addressed coherence as a linguistic fact through such correlates as proportion of global coherence errors, for example, tangential utterances or propositional repetitions (Andreetta et al., 2012; Christiansen, 1995), and every utterance’s relatedness to the overall topic/theme (Glosser & Deser, 1991; Wright & Capilouto, 2012). In the current study, however, coherence was conceived as an overall perception of discourse “quality” by human raters, its intrinsic properties of understandability, clarity, and inner connectedness (Armstrong, 1987; Ulatowska et al., 1981).

The rating scale used in this study for the assessment of overall coherence was sensitive to the differences in discourse produced by people with aphasia and non-brain-damaged participants. The overall discourse coherence operationalized through the combination of understandability, clarity, adequacy of content, and connectedness, was on average decreased in the aphasic group. Our findings are consistent with previous reports of reduced coherence in aphasic speech (e.g., Christiansen, 1995; Ulatowska et al., 1981;

Wright & Capilouto, 2012). The classification analysis summarized below demonstrated that the distinction between aphasic and non-aphasic discourse samples can be rather accurately (Table 4.4) based on their linguistic properties.

The main goal of the classification analysis was to identify which linguistic factors contribute to the overall coherence of discourse, that is, due to which linguistic deficits discourse in aphasia is perceived as less coherent than that of NBD speakers. Seventeen linguistic variables were divided into micro- and macrolinguistic sets and used as predictors of coherence ratings. Based on recent findings demonstrating that microlinguistic processes are tightly related to the macrolinguistic well-formedness of discourse (e.g., Andreetta & Marini, 2015; Wright & Capilouto, 2012), we hypothesized that a combination of micro- and macrolinguistic variables plays a role in establishing coherence. Our results demonstrate that different sets of variables were more relevant for predicting each of the ratings. Content ratings were more accurately predicted by the macrolinguistic variables, the number of main events being one of the best predictors. The potential relationship between thematic informativeness and discourse coherence has been more or less directly addressed in previous studies (e.g., Andreetta & Marini, 2015; Capilouto et al., 2006). Our findings, however, indicate that only one of the aspects of coherence relies substantially on the information content and its relevance, whereas the other aspects refer to the speakers' ability to establish semantic and pragmatic connections between elements of discourse (e.g., Capilouto et al., 2006; Giora, 1997; Hobbs, 1978; Redeker, 1990; Ulatowska & Olness, 2000). For the Understandability, Connectedness, and Clarity ratings, high classification accuracy was reached using micro-, macro-, and all of the linguistic variables respectively. Ulatowska's Content rating (Do you know what is happening in the story?) tackles the minimal requirement for coherence, namely, adequate information content necessary to transfer the message. A crucial variable for predicting Content ratings was the proportion of main events included in the participants' retellings. The Understandability scale (How understandable is the retelling you have listened to?) addressed the same property of discourse as the Content scale, but used a more general, less content-focused formulation. It was more accurately predicted than Content. There are two possible reasons for the observed increase in accuracy: either the five-point scale offers a more fine-grained distinction between the "high" and "low" understandability classes than the three-point scales (see Fig. 4.4), or Understandability is conceptually closer to Clarity than to Content.

Listeners' perception of Connectedness and Clarity were also quite accurately explained by several linguistic variables, especially the variables describing micro- and macrostructural well-formedness of discourse. This finding provides an insight in the role of discourse structure for coherence. The classification results (Tables 4.4 and 4.5) demonstrated that both micro- and macrolinguistic predictors delivered generally high classification accuracy for the coherence ratings, indicating a strong interaction between macro- and microlinguistic processes. An illustration of this connection can be offered based on the two variables identified as universally important in predicting all of the coherence ratings: the proportions of ungrammatical units and structural disfluencies. Specifically, discourse units which are ungrammatical or contain a large number of word-level errors result in a ripple of disfluencies. In turn, they create gaps in the narrative line, disrupt discourse structure, and, inevitably, affect coherence.

Despite the overall high classification accuracy, our findings demonstrated that one of the aspects of coherence cannot be fully explained by linguistic variables alone. It is possible that there are linguistic variables not included in this study which influence the ratings. Another, more likely, possibility is that the key to a better understanding of coherence lays in the listeners' perception of coherence. The Ulatowska's Content rating

scale to a certain extent targets the listeners’ ability to make sense of a discourse. It has been noted before that people “naturally assume coherence and tend to interpret text [discourse] in light of this assumption” (Brown & Yule, 1983, p. 66; Asher & Lascarides, 2003). Consider an example from Hobbs (1979) in (iv).

(iv) *John took a train from Paris to Istanbul. He likes spinach.*

Hobbs (1979) argues that readers would be tempted to interpret these two sentences as related, primarily due to the common assumption that successive utterances usually refer to the same entities, supported by the fact that in this particular case, “he” can only refer to John. Many examples can be given to illustrate that natural conversations’ success relies heavily on similar assumptions about communication dynamics, and on the interlocutors’ common-ground knowledge. It is safe to assume that these and other discourse comprehension and interpretation-related phenomena are at play when evaluating discourse as coherent or not. A future direction of this study that we would like to suggest is an exploratory investigation of naive listeners’ ratings of coherence, similar to the study on narrative quality by Olness et al. (2005).

The sample size in this study was limited to 20 participants. A larger data set would make the observed patterns clearer. For example, a hypothesis to be tested within a larger scale study is that coherence represents a continuum, rather than a binary distinction (Fig. 4.6), in which case regression analysis could be a better choice than classification. Some of the misclassified cases could fall into a gray zone of “medium” coherence. In the current study, the classification is limited to the binary distinction between “coherent” and “incoherent”, largely due to the size of the available data sample. This initial effort will hopefully encourage further explorations on the continuous nature of coherence.

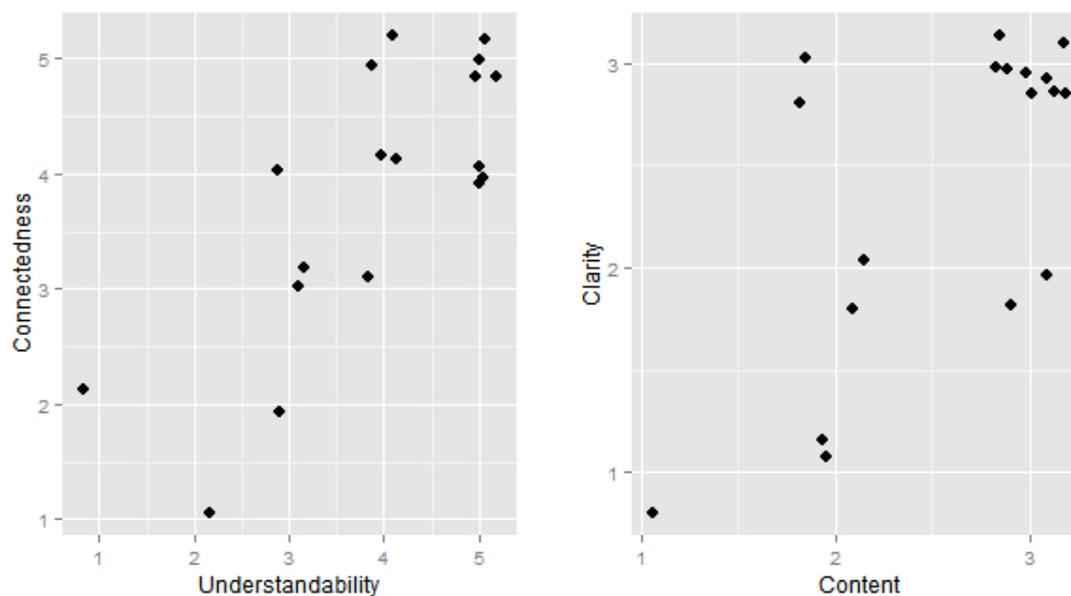


Figure 4.6: The continuum of coherence ratings.

We focused on one genre of spontaneous discourse, video-induced story retelling. This restriction was dictated mainly by the comparability requirements and the choice of methodology, since Rhetorical Structure Theory has not been widely used to study language-impairments before (Kibrik & Podlesskaya, 2009, with the exception of). Although the procedure and the stimulus are designed to elicit as much discourse as possible

in a reasonably naturalistic situation, it would be insightful to include other types of narratives, such as personal stories, in further investigations.

The results of this study have to be replicated with different language-impaired populations in the future. Nonetheless, they provide answers to some of the questions raised more than twenty years ago and extend previous findings on discourse coherence. Our findings suggest that different aspects of coherence are grounded in micro- and macrolinguistic well-formedness of discourse, as well as its informativeness. However, our results also indicate that the multifaceted concept of coherence potentially also contains a comprehension-related pragmatic component, which allows even severely linguistically-impaired aphasic speakers to be understood. Further in-depth investigations of coherence as a perceived quality of discourse can shed more light on the mechanisms involved in language production and comprehension.

The RST analyses of the discourse samples, as well as the R code, are available to speech and language researchers upon request.

## Chapter 5

# General discussion

In the introduction of this dissertation, the concept of coherence was presented. Coherence, especially in the context of spoken discourse, was argued to be a property co-constructed by a speaker and a listener or reader in an interaction. As such, it relies to a varying extent – depending on a situation – both on semantics and its linguistic manifestations, and pragmatics (e.g., Redeker, 1990). The question of how coherence, the internal unity and connectedness of discourse, is established and represented linguistically, has been persistently addressed in the literature on healthy discourse production. It is also the topic of this dissertation.

One of the general approaches briefly discussed in the introduction is grounded in the idea that coherence is established as a result of a structured organization of discourse. Words are bound by morphosyntactic dependencies and rules into clauses and sentences, which are linked to each other with coherence relations, successively forming larger, more complex meaningful groupings of information. Similarly to the way syntactic composition of sentences is formalized in different theories, this approach, often referred to as relational (e.g., Sanders & Spooren, 1999; Taboada & Mann, 2006b), explains discourse coherence through the establishment of a discourse structure, the elements of which are connected with coherence, or rhetorical, relations. In the studies presented in Chapters 3 and 4 of this dissertation, this approach was implemented to factor in a number of discourse structure parameters along with several microlinguistic parameters on the one hand, and listeners' perception of coherence on the other hand.

This chapter presents a brief recapitulation and a discussion of this dissertations' findings on discourse structure and coherence in speakers with and without aphasia. Following a concluding discussion of the results, the main limitations are addressed. An outline of future directions of this work concludes this chapter.

### 5.1 Major Findings

In the introduction of this dissertation, four issues were raised. The following subsections of this chapter contain a brief discussion on the outcomes of this dissertation project with respect to each of these issues.

#### 5.1.1 Coherence deficit in aphasic discourse

The first issue addressed in this dissertation was the disparity in previous findings on discourse production in aphasia. In her 2000 review, Armstrong noted that the arsenal of different methods developed for the investigation of discourse in aphasia, was a double-edged sword. While aphasiologists were provided with many useful tools, the methodological variability made it difficult to connect the findings of different studies.

The body of work on the matter has substantially grown in the past decade and a half. However, along with the large amount of new information came certain confusion about the conclusions that could be drawn from it. The review in Chapter 2 aimed at explaining the origins of the existing inconsistencies in previous findings on discourse production in aphasia. The findings on a number of crucial concepts underlying discourse production, including coherence, were explored, demonstrating that the inconsistencies mostly stemmed from the lack of theoretical and methodological comparability between different approaches to studying them.

An important example of this problem was the research on coherence in aphasic discourse. Initially, the motivation for this dissertation project arose from the contradictory outcomes of previous studies. While some researchers claimed that the ability to produce coherent discourse was intact in aphasia (e.g., Armstrong & Ulatowska, 2007; Glosser & Deser, 1991; Ulatowska et al., 1981), others found evidence of its decrease (e.g., Andreetta & Marini, 2015; Christiansen, 1995; Fergadiotis, Kapantzoglou, & Wright, 2011; Wright & Capilouto, 2012). To a certain extent, this situation was a consequence of the elusive nature of coherence. Considerable attention has been drawn to the definition of the concept in theoretical linguistics (e.g., Gernsbacher & Givón, 1995; Kehler, 2002, 2004; Kintsch & Van Dijk, 1978; Sanders & Spooren, 1999). Nevertheless, in the existing literature, the word “coherence” still referred to different phenomena, or rather, different aspects of this construct. Hence, the methodologies used to investigate coherence differed, and the results were often difficult to connect. While the multitude of approaches reviewed in Chapter 2 can be confusing, their organized compilation provided a clearer depiction of the inner organization of the phenomenon of coherence. Specifically, previous studies focused on different aspects of coherence, mainly related to propositional content and its relevance with respect to the main theme, or topic. Coherence, however, as it has been discussed in the beginning of this section, has been argued to hinge on relations between elements of discourse, its internal structure. The longer and the more informative the message, the more important it is to understand how different parts of it are connected to each other. Hence, the more obvious is the role of relations linking different parts of discourse. Different parts of a discourse are not necessarily connected linearly, as in the case of a simple sequence of events. For example, semantic and pragmatic relations play an important role in the understanding of a narrative containing not uncommon elaborative, explanatory, and expository elements or deviations from the main story line.

Furthermore, in addition to the semantic, or ideational, core, coherence has a second, pragmatic, component (e.g., Redeker, 1990; Sanders, 1997). While not many studies to date have addressed pragmatics in aphasia (e.g., Armstrong & Ferguson, 2010a, 2010b; Dronkers, Ludy, & Redfern, 1998; Wulfeck et al., 1989), it may just be the case that speakers experiencing difficulties with semantic processing rely more on the pragmatic component when producing discourse. In other words, aphasic speakers may rely on the interlocutor’s comprehension, his or her ability to fill the gaps, to compensate for their linguistic deficits. Based on these considerations, the next step of this dissertation was to explore the interaction between linguistic and interactive pragmatic aspects of coherence.

### 5.1.2 Discourse structure in aphasia

The conclusions of the review in Chapter 2 led us to the second issue raised in this dissertation, related to discourse structure and aphasic speech. Whereas linguistic deficits characteristic of aphasia have been identified at phonological lexical-semantic, and/or grammatical levels, it is not clear whether reduced coherence is a result of one or more

of these deficits or of it is a separate deficit at the level of macrolinguistic processing. The idea of discourse structure developed from the awareness of the fact that elements of discourse do not necessarily represent a sequence, but that they are quite often connected in a more elaborate way. Discourse structure is a theoretical construct, and its implementation and representation have been questioned numerous times. One of the frameworks formalizing the concept of discourse structure and offering methodological apparatus for the analysis thereof is Rhetorical Structure Theory (RST; Mann & Thompson, 1988). According to RST, discourse consists of elementary units, roughly equal to clauses, interconnected with coherence relations (also called “discourse” or “rhetorical” relations), forming a complete discourse structure. The first question to be addressed in Chapter 3 was whether this approach was applicable to impaired spoken language, in which syntactic and word-level impairments could present difficulties for the identification of discourse segments and relations between them. So far, no analyses of discourse structure in language-impaired speech using a relational theory such as Rhetorical Structure Theory or similar approaches, has been performed, apart from the study of discourse coherence in children with neurosis by Kibrik and Podlesskaya (2009).

In the study presented in Chapter 3, discourse structure in two types of discourse a narrative elicited by a series of pictures and a procedural description, produced by a group of native Cantonese speakers with and without anomic aphasia was analyzed. The results showed that by and large people with aphasia structured discourse similarly to the non-brain-damaged speakers. The relations between discourse units, the smallest of them roughly equal to clauses, were mostly identifiable (85% reliability), and the structures were similar in the samples of aphasic and non-brain-damaged participants. Hence, on the one hand, this study demonstrated that the ability of people with aphasia to establish discourse structure was largely preserved. Although it has been claimed previously that aphasic speakers maintained their ability to produce well-structured discourse (e.g., Olness & Ulatowska, 2011; Ulatowska et al., 1990; Ulatowska et al., 1981), there have been no investigations of formal structure in aphasic discourse to support this claim prior to the current study. Conversely, the existence of discourse structure as a concept, which was originally developed for and based on non-impaired written discourse, was substantiated by aphasic data.

Certain structural differences were, however, observed between the two groups. Namely, such parameters as the length of the samples in EDUs, the number of different rhetorical relations used to build discourse structure, and the proportion of gaps and disfluencies in the structures, indicated that non-brain-damaged participants produced structurally more complex discourse than aphasic speakers. Genre-specific differences, such as structural patterns, types of relations used, and length in EDUs, were observed as well. Despite the relative simplicity of the global structure in procedural discourse (i.e., enumeration of steps) and the optionality of elaborative components, as in the narrative task, the aphasic group performed worse than the non-aphasic group and received lower coherence ratings. This observation indicates that structural impairments occur in aphasic discourse independently of complexity. However, the study in Chapter 3 was not able to provide an insight on whether the impairments result from word- or sentence-level deficits or not, as the aphasic group was homogeneous and presented with similar deficits. This issue was the focus of the study presented in Chapter 4.

The importance of cross-linguistic studies in aphasiology was stressed in Chapter 2. The studies included in this dissertation are based on the data from two typologically different languages – Cantonese (Chapter 3) and Russian (Chapter 4). The main idea behind this methodological choice consisted in the possibility to tease apart micro- and macrolinguistic levels in discourse production. While Russian has a highly inflectional

morphology with a flexible word order, Cantonese is an analytic language with a relatively fixed word order and without inflectional morphology. Hence, aphasic speakers in these two languages experience different deficits and exploit different coping strategies. These microlinguistic differences, however, were expected to have no effect on discourse structure, as the concept of discourse structure is language-independent. Although the differences in genres and aphasia types in the groups do not allow to compare Cantonese and Russian data, it was demonstrated in this dissertation that similar kinds of discourse structures can be identified in the samples of aphasic and non-brain-damaged speakers of both Cantonese and Russian.

### 5.1.3 On the multifaceted nature of coherence

The third point addressed in this dissertation was the multifaceted nature of coherence. The analysis provided in the review in Chapter 2 revealed that several aspects of coherence have been suggested/addressed in previous research on discourse in aphasia. Specifically, in many studies coherence was practically equated with the appropriateness and integrity of propositional, or thematic, content. These studies focused on the speakers' ability to include and manipulate information in connected speech production. In a small number of works, however, coherence was considered to be a perceptual quality – semantic and pragmatic unity of discourse. It was also referred to as the overall narrative “quality” (Olness et al., 2005). In this case, the notion of coherence comprises thematic content along with other linguistic and extra-linguistic variables contributing to the “hanging together”, or connectedness and understandability of discourse. Considering that coherence is co-constructed in a speaker-listener interaction, these two approaches differ in the following way. While the first strategy appeals to the production component of coherence, the second one accounts for the comprehension as well.

For narratives, which are normally structured around an event line, propositional content is often crucial for coherence. Coherence in conversational discourse, however, can be less dependent on explicitly verbalized content due to pragmatic knowledge available to interlocutors. The question remains which events really constitute the “backbone” event line of a discourse, and how big of a loss in content should be to render the narrative less coherent. Moreover, contextual observations and inferences can contribute to maintaining coherence even in linguistically defective narratives. Hence, it is important to consider coherence a quality which is construed by linguistic means, but is to a large extent perceptual in nature.

A study of coherence is, therefore, complicated not only because of the difficulties related to the definition of this concept. Having determined what “coherence” stands for, one has to decide which of its aspects to assess in experimental (and clinical) settings. In both the Cantonese and the Russian study presented in this dissertation, coherence was addressed as “perceived quality”. Its different correlates, which were used in the rating tasks, were selected based on previous literature. They included a combination of content, clarity, connectedness, understandability, and order of events.

The classification analysis in the Russian study described in Chapter 4 aimed at identifying which of the pre-selected objectively calculated linguistic variables contributed most to establishing and maintaining coherence. The perception of these aspects of coherence – as rated by listeners – was largely explained by linguistic well-formedness of discourse samples. However, the results also suggested that other, possibly comprehension-related parameters influence the perception of (some of the aspects) of coherence. It is also possible that the representation of coherence through the aforementioned correlates was incomplete.

#### 5.1.4 How is coherence established linguistically?

All this being said, the relationship between discourse coherence and linguistic parameters contributing to its establishment was the fourth and the last issue addressed in this dissertation. The relationship between micro- and macrolinguistic parameters of discourse, including the appropriateness and completeness of propositional content has been investigated before, revealing the connection between lexical and sentence-level disturbances in aphasia and reduced or inadequate propositional content (e.g., Andreetta et al., 2012; Andreetta & Marini, 2015; Ulatowska et al., 1990; Wright & Capilouto, 2012). Although the influence of objective linguistic parameters of discourse on coherence ratings assigned by scorers has been of interest for a long time (e.g., Ulatowska et al., 1981), it has not been studied previously.

The relationship between micro- and macrolinguistic variables of discourse samples produced by aphasic and non-brain-damaged participants was explored in Chapter 3 with Cantonese and in Chapter 4 with Russian aphasic and non-brain-damaged speakers. In the first study, two types of discourse were included, namely picture-induced narratives and procedural description. The second study was based on story retellings elicited by a short movie. The collected samples were annotated for discourse structure, and a number of macro-structural variables were obtained based on the annotations.

In the Cantonese study presented in Chapter 3, a number of parameters characterizing the macrolinguistic organization of aphasic discourse were included in the correlation analysis between coherence ratings by naive listeners and a number of linguistic parameters of the discourse samples. The correlation analysis was performed to determine linguistic variables potentially contributing to coherence perception. It demonstrated that a number of discourse structure variables strongly correlated with coherence ratings, suggesting that there is a connection between discourse structure and coherence. This connection and the role of microlinguistic variables in establishing and maintaining it were further investigated in the study in Chapter 4.

To examine whether, or to what extent, it was possible to predict coherence ratings based on linguistic parameters of the discourse samples, a classification analysis was implemented in the Russian study in Chapter 4. It included a procedure for variable selection, through which the best/most informative micro- and/or macrolinguistic predictors of different coherence ratings were identified. The results demonstrated that coherence to a great extent depends on linguistic well-formedness of discourse, namely, its morphosyntactic grammaticality and structural integrity. Interestingly, different aspects of coherence had different sets of best predictors. For example, one of the main predictors of the content ratings was the number of main events, that is, the measure of propositional content included in the retellings. However, for clarity, the proportion of structural disfluencies and disruptions, word-level errors, and word-finding comments were found to be more important.

The obtained high accuracies for ratings' prediction indicated that coherence was largely established through linguistic means. However, not all of the correlates were equally well predicted based solely on linguistic variables. The Content ratings, following the original procedure in Ulatowska et al. (1981), addressed listeners' ability to understand what was happening in the story. In Armstrong's (1987) terms it can be formulated as "the ability to make sense of" a piece of discourse. Crucially, these ratings were the most difficult to predict based on the linguistic variables. The explanation suggested in Chapter 4 was rooted once again in the dual nature of coherence. Specifically, as it was emphasized before, the approach chosen in this dissertation project was to consider coherence to be a combination of both speakers' and listeners' efforts. The Content rating

scale was formulated to address the listener’s contribution in the process of building up coherence. Hence, the fact that these ratings were not particularly accurately predicted based on linguistic parameters alone supports the presupposition about the listeners’ part in the process of establishing coherence, or in this case, one of its aspects – their ability to make sense of a discourse.

## 5.2 General conclusions

In this section the main outcomes of this dissertation will be briefly summarized. Firstly, it was determined and demonstrated that coherence is a complex phenomenon consisting of several aspects. Different linguistic variables/features were found to be relevant for establishing and maintaining different aspects of coherence. For example, word-level errors were impeding the clarity, but not connectedness of discourse. Neither did they influence the raters’ ability to understand what happened in the story.

The classification analysis of samples into “coherent” and “incoherent” demonstrated that clause level grammaticality and discourse structure integrity were important predictors of discourse coherence. As it was argued in Chapter 4, the impact of these two linguistic variables can be additive/intertwined. Specifically, coherence impairment is a result of the pipeline effect: since clauses represent elementary building blocks of discourse structure, a breakdown at clausal level can result in discourse structure disintegration, which, in turn, leads to reduced coherence.

The ability of aphasic speakers to build discourse structure was found to be similar to that of non-brain-damaged people. Just as in non-brain-damaged speakers, discourse structure could be identified in the samples of all of the aphasic participants included in the studies. However, structural damage was higher in aphasic discourse samples, which reflected in the lower coherence ratings they obtained. To summarize, aphasic speakers preserve the ability to establish discourse structure, but their structures are often breached by/incomplete due to multiple micro-linguistic deficits. As a result, aphasic discourse is perceived as less coherent than that of non-brain-damaged speakers. Two observations substantiate this conclusion, though they require additional testing and validation. First, a small number of non-aphasic discourses were poorly structured and received lower coherence ratings, in line with the idea about the connection between coherence and discourse structure. Second, a small number of cases, in which high coherence ratings were assigned to linguistically impaired discourses, supported the assumption about listeners’ contribution to coherence. A larger number of discourse samples with ranging degrees of linguistic well-formedness is necessary to justify these considerations.

## 5.3 Further directions

The ultimate goal of this dissertation was to lay the groundwork for the studies of perhaps the most crucial concept in discourse production – coherence – in aphasia. Having marshaled previous findings on discourse production in aphasia in Chapter 2, we moved on to exploring discourse structure in aphasia and the connection between coherence – semantico-pragmatic unity – of discourse and its linguistic features. Spoken discourse is indispensable for daily communication. Its production is in the core of aphasia assessment and treatment (e.g., Prins & Bastiaanse, 2004). Coherence is what constitutes a large part of communicative success, and it is impaired in aphasia. The findings of this dissertation provide new insights on the contribution of linguistic components to the communicative effectiveness of people with aphasia, rehabilitation of which is one of the main goals of aphasia treatment. These outcomes are thus informative for therapy. This

work, however, has implications for research on both aphasic and non-brain-damaged spoken language. The study of coherence in aphasia is also insightful in terms of the general understanding of the processes behind human language production and comprehension. The variety of linguistic deficits observed in different aphasia types offers a unique possibility to tease apart the effects of different linguistic levels/variables on discourse production. This section provides a summary of the limitations to the research presented in this dissertation that will allow the readers to evaluate the strength of the findings and conclusions of this work, as well as several suggestions on further directions for the research on coherence in aphasic and healthy discourse.

The main limitation of the studies presented in this dissertation is the relatively small number of participants. Comparable to many studies in language pathology research, the sample sizes were limited to two or three dozens of people. The findings of this dissertation can be refined and expanded in a follow-up study with a larger number of participants, including aphasic speakers with a wider variety of language production difficulties and severity. Larger samples will be helpful in making the patterns identified in this dissertation more transparent.

Whereas the group of anomic aphasic participants in the Cantonese study was homogeneous, the Russian aphasic group was diverse, with varying symptoms and ranging from moderate to severe aphasia. The effects of different types of deficits and severity levels on discourse structure are yet to be explored systematically. Since discourse structure is a non-language-specific construct, the investigation of this issue can be based on data from different languages. While the people with aphasia who participated in this dissertation project exhibited the ability to establish discourse structure, it was not equally well preserved in and available to all of the participants. It is possible that more severe aphasic speakers experience more expressed difficulties with discourse organization, in which case it would also be less obvious whether listeners would be able to reconstruct discourse structure and make sense of their discourse.

It is also probable that less topic-constrained, more spontaneous discourse production is characterized by less apparent discourse structure and more dispersed coherence ratings. Continuing the exploratory work in these two directions can shed substantially more light on the contribution of particular linguistic features to coherence, the conversational dynamics in aphasia, and the interaction between production and comprehension in language-based communication in general.

Another factor constraining the outcomes of this dissertation is the limited choice of genres. It has been emphasized that genre-specific differences in linguistic and cognitive load imposed on speakers have an effect on discourse production (e.g., Olness, 2006, 2007; Olness et al., 2002; Ulatowska et al., 1990). Due to the novelty of the chosen approach for aphasia research and the related comparability requirements, the elicitation tasks had to be restricted to those producing the most comparable discourse samples in sufficiently naturalistic settings. Two different genres of discourse were elicited in the studies presented in this dissertation: procedural description and two types of narratives elicited using a series of pictures and a short movie. Different genres of discourse impose different cognitive functions. For example, procedural and expository discourse does not involve as much inferencing as retelling of fables, and not as much chronological sequencing or coreference as story-telling and personal narratives (e.g., Olness & Ulatowska, 2011; Ulatowska et al., 1999; Ulatowska et al., 1981). Especially interesting and informative would be such unrestricted genres as personal narratives and topic-centered conversation. Whilst they were not included in the current dissertation for comparability reasons, further studies should comprise other genres, especially those involving more spontaneous speech production. It is paramount to account for genre-related variability for a better understanding

of how coherence is established in aphasia and in healthy discourse. Determining the role of linguistic abilities of a speaker and interlocutors' contribution to communicative success could be important for the research on cognitive abilities in people with brain damage.

Among the planned extensions of this work is also a cross-linguistic comparison of discourse structuring abilities in aphasia. The advantages of cross-linguistic research, such as the possibility to make generalized conclusions, have also been discussed in Chapter 2. Comparing speech samples in different languages elicited using the same task will help to disentangle the effects of micro- and macrolinguistic levels in discourse production, shedding light on the idea of the “pipeline” impairment in aphasic should be addressed in this section as well.

A crucial methodological step taken in this dissertation was introducing the Rhetorical Structure Theory (RST), a theory explaining coherence “by postulating a hierarchical, connected structure of texts, in which every part of a text has a role, a function to play” (Taboada & Mann, 2006b, p. 425). Several points of criticism of the RST have been mentioned in Chapters 3 and 4. For example, according to RST, a single discourse/rhetorical relation is to be chosen when interconnection elementary units of discourse (EDUs). This constraint leading to the tree representation of discourse structure in RST has been questioned and argued against multiple times (e.g., Miltsakaki et al., 2004; Wolf & Gibson, 2005). However, the same constraint asserts that “the human tendency to interpret texts as coherent even when they are not” (Egg & Redeker, 2010; Marcu, 2003) does not stimulate an overload of unnecessary links between discourse segments.

On the technical side, the annotations of the discourse samples in this project were performed by the researchers co-authoring the journal papers included in this thesis. The desirable standard for estimating the reliability of discourse level analysis is independent double, or even triple, blind annotation. Whereas this approach is usually recommendable, it requires the annotation process to be straightforward and well-tested. It was, however, too strict for this first attempt of the kind to annotate aphasic speech. While working on the annotations, my colleagues and I were developing a set of additional guidelines for the challenges we were facing annotating aphasic spoken discourse (Kibrik & Podlesskaya, 2009, based on). For example, several new relations were added to the original extended RST relation set (Carlson & Marcu, 2001) to mark disfluencies in aphasic speech. Discourse segmentation guidelines had to be rethought as well to account for such common for aphasic speech cases as incomplete or semantically empty EDUs, unnatural pauses and intonation. Hence, the methodological choice to involve two or three members of the research teams participating in this dissertation project was motivated by the necessity to adjust the approach to aphasic data. Nevertheless, the novelty of this approach for aphasiology guaranteed the absence of bias in the annotation process, since no particular expectations could be formed based on previous findings. The annotations were critically reviewed by the members of research teams, and questionable cases were discussed and documented.

Along with the RST analysis of discourse samples, rating studies were conducted, in which coherence of the samples was assessed by listeners. This element of the studies' design was meant to tap into the pragmatic aspects of coherence. A follow-up study should compare scores assigned by raters familiar and unfamiliar with the elicitation stimulus (e.g., a video fragment). Such phenomena as common ground, world knowledge, and shared visual domain, related to meaning co-construction in interaction are most likely at play when linguistic components are malfunctioning. They are expected to contribute to the understandability and clarity – in other words, coherence – of spoken discourse in aphasia.

Since the results of the two rating studies presented in this dissertation were not comparable due to the differences in discourse elicitation tasks, a more systematic exploration of raters' perception of coherence is in order in the future. A study comparing ratings assigned by naive listeners versus trained raters can shed more light on the role of listeners in establishing coherence. While naive listeners' perception is potentially more ecologically valid, and the reliability may be sufficiently high, the construct validity of this method is more questionable than in the case of trained raters' evaluation. Future investigations focusing on this issue, following Olness et al. (2005), should consider open-end characterizations of such seemingly self-explanatory concepts as clarity or "the ability to make sense of" a piece of discourse.



# References

- Akhutina, T. V. (1982). Syntactic deficits in aphasia. *Paper in Linguistics*, 15(4), 263–274. doi:10.1080/08351818209370576
- Akhutina, T. V. (2016). Luria’s classification of aphasias and its theoretical basis. *Aphasiology*, 30(8), 878–897. doi:10.1080/02687038.2015.1070950
- Albrecht, J. E. & O’Brien, E. J. (1993). Updating a mental model: Maintaining both local and global coherence. *Journal of experimental psychology: Learning, memory, and cognition*, 19(5), 1061. doi:10.1037/0278-7393.19.5.1061
- Andreetta, S. (2014). *Features of narrative language in fluent aphasia* (Doctoral dissertation, Università degli Studi di Udine, Italy).
- Andreetta, S., Cantagallo, A., & Marini, A. (2012). Narrative discourse in anomic aphasia. *Neuropsychologia*, 50(8), 1787–1793. doi:10.1016/j.neuropsychologia.2012.04.003
- Andreetta, S. & Marini, A. (2015). The effect of lexical deficits on narrative disturbances in fluent aphasia. *Aphasiology*, 29(6), 705–723. doi:10.1080/02687038.2014.979394
- Antonio, J. D. & Cassim, F. T. R. (2012). Coherence relations in academic spoken discourse. *Linguistica*, 52(1), 323–336. doi:10.4312/linguistica.52.1.323-336
- Armstrong, E. (1987). Cohesive harmony in aphasic discourse and its significance in listener perception of coherence. *Clinical aphasiology*, 17, 210–215. Retrieved from <http://eprints-prod-05.library.pitt.edu/936/1/17-23.pdf>
- Armstrong, E. (1992). Clause complex relations in aphasic discourse: A longitudinal case study. *Journal of neurolinguistics*, 7(4), 261–275. doi:10.1016/0911-6044(92)90018-R
- Armstrong, E. (2000). Aphasic discourse analysis: The story so far. *Aphasiology*, 14(9), 875–892. doi:10.1080/02687030050127685
- Armstrong, E., Ciccone, N., Godecke, E., & Kok, B. (2011). Monologues and dialogues in aphasia: Some initial comparisons. *Aphasiology*, 25(11), 1347–1371. doi:10.1080/02687038.2011.577204
- Armstrong, E. & Ferguson, A. (2010a). Interacting with difficulty: The case of aphasia. *New adventures in language and interaction*, 196, 199. doi:10.1075/pbns.196.09arm
- Armstrong, E. & Ferguson, A. (2010b). Language, meaning, context, and functional communication. *Aphasiology*, 24(4), 480–496. doi:10.1080/02687030902775157
- Armstrong, E. & Ulatowska, H. K. (2007). Making stories: Evaluative language and the aphasia experience. *Aphasiology*, 21(6-8), 763–774. doi:10.1080/02687030701192364
- Artstein, R. & Poesio, M. (2008). Inter-coder agreement for computational linguistics. *Computational Linguistics*, 34(4), 555–596. doi:10.1162/coli.07-034-R2
- Asher, N. & Lascarides, A. (2003). *Logics of conversation*. Cambridge: Cambridge University Press.
- Bakheit, A., Carrington, S., Griffiths, S., & Searle, K. (2005). High scores on the Western Aphasia Battery correlate with good functional communication skills (as measured with the Communicative Effectiveness Index) in aphasic stroke patients. *Disability and Rehabilitation*, 27(6), 287–291. doi:10.1080/09638280400009006
- Barch, D. M. & Berenbaum, H. (1997). The effect of language production manipulations on negative thought disorder and discourse coherence disturbances in schizophrenia. *Psychiatry Research*, 71(2), 115–127. doi:10.1016/S0165-1781(97)00045-0
- Barnes, S. & Ferguson, A. (2015). Conversation partner responses to problematic talk produced by people with aphasia: Some alternatives to initiating, completing, or pursuing repair. *Aphasiology*, 29(3), 315–336. doi:10.1080/02687038.2013.874547
- Bastiaanse, R. (2010). *Afasie*. Houten: Bohn Stafleu van Loghum.
- Bastiaanse, R. (2013). Why reference to the past is difficult for agrammatic speakers. *Clinical linguistics & phonetics*, 27(4), 244–263. doi:10.3109/02699206.2012.751626

- Bastiaanse, R., Bamyaci, E., Hsu, C.-J., Lee, J., Duman, T. Y., & Thompson, C. K. (2011). Time reference in agrammatic aphasia: a cross-linguistic study. *Journal of Neurolinguistics*, *24*(6), 652–673. doi:10.1016/j.jneuroling.2011.07.001
- Bastiaanse, R., Edwards, S., & Kiss, K. (1996). Fluent aphasia in three languages: Aspects of spontaneous speech. *Aphasiology*, *10*(6), 561–575. doi:10.1080/02687039608248437
- Bastiaanse, R., Hugen, J., Kos, M., & Van Zonneveld, R. (2002). Lexical, morphological, and syntactic aspects of verb production in agrammatic aphasics. *Brain and language*, *80*(2), 142–159. doi:10.1006/brln.2001.2585
- Bastiaanse, R. & Jonkers, R. (1998). Verb retrieval in action naming and spontaneous speech in agrammatic and anomie aphasia. *Aphasiology*, *12*(11), 951–969. doi:10.1080/02687039808249463
- Bastiaanse, R., Koekkoek, J., & van Zonneveld, R. (2003). Object scrambling in Dutch Broca's aphasia. *Brain and Language*, *86*(2), 287–299. doi:10.1016/S0093-934X(02)00545-X
- Bates, E., Hamby, S., & Zurif, E. (1983). The effects of focal brain damage on pragmatic expression. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *37*(1), 59. doi:10.1037/h0080695
- Bates, E. & Wulfeck, B. (1989). Crosslinguistic studies of aphasia. In B. MacWhinney & E. Bates (Eds.), *The crosslinguistic study of sentence processing*. (pp. 328–371). New York: Cambridge University Press.
- Bates, E., Wulfeck, B., & MacWhinney, B. (1991). Cross-linguistic research in aphasia: An overview. *Brain and language*, *41*(2), 123–148. doi:10.1016/0093-934X(91)90149-U
- Beeke, S., Beckley, F., Johnson, F., Heilemann, C., Edwards, S., Maxim, J., & Best, W. (2015). Conversation focused aphasia therapy: Investigating the adoption of strategies by people with agrammatism. *Aphasiology*, *29*(3), 355–377. doi:10.1080/02687038.2014.881459
- Beeke, S., Maxim, J., & Wilkinson, R. (2007). Using conversation analysis to assess and treat people with aphasia. In *Seminars in speech and language* (Vol. 28, 02, pp. 136–147). Thieme Medical Publishers. doi:10.1055/s-2007-970571
- Beeke, S., Wilkinson, R., & Maxim, J. (2003). Exploring aphasic grammar 2: do language testing and conversation tell a similar story? *Clinical linguistics & phonetics*, *17*(2), 109–134. doi:10.1080/0269920031000061786
- Beeke, S., Wilkinson, R., & Maxim, J. (2007). Grammar without sentence structure: A conversation analytic investigation of agrammatism. *Aphasiology*, *21*(3-4), 256–282. doi:10.1080/02687030600911344
- Behrns, I. (2009). *Aphasia and the challenge of writing* (Doctoral dissertation, University of Gothenburg, Sweden).
- Benson, D. F. & Ardila, A. (1996). *Aphasia: A clinical perspective*. New York: Oxford University Press.
- Benson, D. F. & Geschwind, N. (1971). Aphasia and related cortical disturbances. In A. B. Baker & L. H. Baker (Eds.), *Clinical neurology* (pp. 112–140). New York: Harper & Row.
- Berman, R. A. (1997). Narrative theory and narrative development: The Labovian impact. *Journal of Narrative and Life History*, *7*(1), 235–244. doi:10.1075/jnlh.7.29nar
- Beveridge, M. E. & Bak, T. H. (2011). The languages of aphasia research: Bias and diversity. *Aphasiology*, *25*(12), 1451–1468. doi:10.1080/02687038.2011.624165
- Bliss, L. S. & McCabe, A. (2006). Comparison of discourse genres: Clinical implications. *Contemporary Issues in Communication Science and Disorders*, *33*(33), 126–137.
- Blomert, L., Kean, M., Koster, C., & Schokker, J. (1994). Amsterdam–Nijmegen everyday language test: construction, reliability and validity. *Aphasiology*, *8*(4), 381–407. doi:10.1080/02687039408248666
- Blomert, L., Koster, C., & Kean, M.-L. (1995). *Antat: Amsterdam-Nijmegen Test voor Alledaagse Taalvaardigheden (Amsterdam-Nijmegen Everyday Language Test: Manual)*. Lisse: Swets & Zeitlinger.
- Bloom, R. L. (1994). Hemispheric responsibility and discourse production: Contrasting patients with unilateral left and right hemisphere damage. In R. L. Bloom, L. K. Opler, S. De Santi, & J. S. Ehrlich (Eds.), *Discourse analysis and applications: studies in adult clinical populations* (pp. 81–94). Hillsdale, NJ: Erlbaum.
- Bloom, R. L., Borod, J. C., Santschi-Haywood, C., Pick, L. H., & Opler, L. K. (1996). Left and right hemispheric contributions to discourse coherence and cohesion. *International Journal of Neuroscience*, *88*(1-2), 125–140. doi:10.3109/00207459608999818
- Blumstein, S. E. (1998). Phonological aspects of aphasia. In M. T. Sarno (Ed.), *Acquired aphasia* (pp. 157–185). San Diego, CA: Academic Press.
- Boles, L. (1998). Conversational discourse analysis as a method for evaluating progress in aphasia: a case report. *Journal of Communication Disorders*, *31*(3), 261–274. doi:10.1016/S0021-9924(98)00005-7

- Bos, L. S., Dragoy, O., Avrutin, S., Iskra, E., & Bastiaanse, R. (2014). Understanding discourse-linked elements in aphasia: A threefold study in Russian. *Neuropsychologia*, *57*, 20–28. doi:10.1016/j.neuropsychologia.2014.02.017
- Bourgeois, M. S. & Hickey, E. (2011). *Dementia: From diagnosis to management - A functional approach*. New York: Psychology Press.
- Boyle, M. (2011). Discourse treatment for word retrieval impairment in aphasia: The story so far. *Aphasiology*, *25*(11), 1308–1326. doi:10.1080/02687038.2011.596185
- Breiman, L. (2001). Random forests. *Machine learning*, *45*(1), 5–32. doi:10.1023/A:1010933404324
- 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, Language, and Hearing Research*, *37*(2), 399–407. doi:10.1044/jshr.3702.399
- Brown, G. & Yule, G. (1983). *Discourse analysis*. Cambridge: Cambridge university press.
- Cameron, D. (2001). *Working with spoken discourse*. London, Thousand Oaks, New Delhi: Sage Publications.
- Cannito, M. P., Jarecki, J. M., & Pierce, R. S. (1986). Effects of thematic structure on syntactic comprehension in aphasia. *Brain and Language*, *27*(1), 38–49. doi:10.1016/0093-934X(86)90003-9
- Capilouto, G. J., Wright, H. H., & Wagovich, S. A. (2005). CIU and main event analyses of the structured discourse of older and younger adults. *Journal of communication disorders*, *38*(6), 431–444. doi:10.1016/j.jcomdis.2005.03.005
- Capilouto, G. J., Wright, H. H., & Wagovich, S. A. (2006). Reliability of main event measurement in the discourse of individuals with aphasia. *Aphasiology*, *20*(02-04), 205–216. doi:10.1080/02687030500473122
- Carlson, L. & Marcu, D. (2001). *Discourse tagging reference manual* (tech. rep. No. ISI-TR-545). Information Science Institute, University of South Carolina. Los Angeles, CA. Retrieved from ftp://128.9.176.20/isi-pubs/tr-545.pdf
- Carlson, L., Marcu, D., & Okurowski, M. E. (2003). Building a discourse-tagged corpus in the framework of rhetorical structure theory. In J. C. van Kuppevelt & R. W. Smith (Eds.), *Current and new directions in discourse and dialogue. text, speech and language technology* (Vol. 22, pp. 85–112). Dordrecht: Springer. doi:10.1007/978-94-010-0019-2\_5
- Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and topic* (pp. 27–55). New York: Academic Press.
- Chafe, W. L. (Ed.). (1980). *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production*. Norwood, NJ: Ablex.
- Chapman, S. B. & Ulatowska, H. K. (1989). Discourse in aphasia: Integration deficits in processing reference. *Brain and language*, *36*(4), 651–668. doi:10.1016/0093-934X(89)90092-8
- Chapman, S. B., Ulatowska, H. K., King, K., Johnson, J. K., & McIntire, D. D. (1995). Discourse in early Alzheimer's disease versus normal advanced aging. *American Journal of Speech-Language Pathology*, *4*(4), 124–129. doi:10.1044/1058-0360.0404.124
- Chomsky, N. (1971). Deep structure, surface structure, and semantic interpretation. In D. D. Steinberg & L. A. Jakobovits (Eds.), *Semantics: an interdisciplinary reader in philosophy, linguistics and psychology* (pp. 183–216). Cambridge: Cambridge University Press.
- Christiansen, J. A. (1995). Coherence violations and propositional usage in the narratives of fluent aphasics. *Brain and Language*, *51*(2), 291–317. doi:10.1006/brln.1995.1062
- Chung, K. K., Code, C., & Ball, M. J. (2004). Lexical and non-lexical speech automatisms in aphasic Cantonese speakers. *Journal of Multilingual Communication Disorders*, *2*(1), 32–42. doi:10.1080/1476967031000154213
- Coelho, C. A. (2002). Story narratives of adults with closed head injury and non-brain-injured adults: Influence of socioeconomic status, elicitation task, and executive functioning. *Journal of Speech, Language, and Hearing Research*, *45*(6), 1232–1248. doi:10.1044/1092-4388(2002/099)
- Coelho, C. A. & Flewellyn, L. (2003). Longitudinal assessment of coherence in an adult with fluent aphasia: A follow-up study. *Aphasiology*, *17*(2), 173–182. doi:10.1080/729255216
- Coelho, C. A., Liles, B. Z., Duffy, R. J., Clarkson, J. V., & Elia, D. (1994). Longitudinal assessment of narrative discourse in a mildly aphasic adult. *Clinical Aphasiology*, *22*, 145–155. Retrieved from http://eprints-prod-05.library.pitt.edu/164
- Cristea, D., Ide, N., & Romary, L. (1998). Veins theory: A model of global discourse cohesion and coherence. In *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics* (Vol. 1, pp. 281–285). Association for Computational Linguistics.

- Damico, J. S., Oelschlaeger, M., & Simmons-Mackie, N. (1999). Qualitative methods in aphasia research: Conversation analysis. *Aphasiology*, *13*(9-11), 667–679. doi:10.1080/026870399401777
- Damico, J. S., Tetnowski, J., Lynch, K., Hartwell, J., Weill, C., Heels, J., & Simmons-Mackie, N. (2015). Facilitating authentic conversation: An intervention employing principles of constructivism and conversation analysis. *Aphasiology*, *29*(3), 400–421. doi:10.1080/02687038.2014.945388
- Daradoumis, T. (1996). Towards a representation of the rhetorical structure of interrupted exchanges. In *Trends in natural language generation an artificial intelligence perspective* (pp. 106–124). Springer. doi:10.1007/3-540-60800-1\_26
- de Riesthal, M. R. (2011). Influence of the stimulus on verbal and nonverbal responses by individuals with and without aphasia. In *Proceedings of the Clinical Aphasiology Conference*. Fort Lauderdale, FL. Retrieved from <http://eprints-prod-05.library.pitt.edu/id/eprint/2301>
- den Ouden, J. N., van Wijk, C., Terken, J. M., & Noordman, L. (1998). Reliability of discourse structure annotation. *IPO Center for Research on User-System Interaction Annual Progress Report*, *33*, 129–138.
- Dettling, M. & Bühlmann, P. (2002). Supervised clustering of genes. *Genome biology*, *3*(12), research0069–1. doi:10.1186/gb-2002-3-12-research0069
- Díaz-Uriarte, R. & De Andres, S. A. (2006). Gene selection and classification of microarray data using random forest. *BMC bioinformatics*, *7*(1), 3. doi:10.1186/1471-2105-7-3
- Dipper, L., Cocks, N., Rowe, M., & Morgan, G. (2011). What can co-speech gestures in aphasia tell us about the relationship between language and gesture?: A single case study of a participant with conduction aphasia. *Gesture*, *11*(2), 123–147. Retrieved from 10.1075/gest.11.2.02dip
- Ditman, T. & Kuperberg, G. R. (2007). The time course of building discourse coherence in schizophrenia: an ERP investigation. *Psychophysiology*, *44*(6), 991–1001. Retrieved from 10.1111/j.1469-8986.2007.00565.x
- Ditman, T. & Kuperberg, G. R. (2010). Building coherence: A framework for exploring the breakdown of links across clause boundaries in schizophrenia. *Journal of neurolinguistics*, *23*(3), 254–269. Retrieved from 10.1016/j.jneuroling.2009.03.003
- Doyle, P. J., Goda, A. J., & Spencer, K. A. (1995). The communicative informativeness and efficiency of connected discourse by adults with aphasia under structured and conversational sampling conditions. *American Journal of Speech-Language Pathology*, *4*(4), 130–134. doi:10.1044/1058-0360.0404.130
- 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. doi:10.1080/026870300401306
- Dronkers, N. F., Ludy, C. A., & Redfern, B. B. (1998). Pragmatics in the absence of verbal language: Descriptions of a severe aphasic and a language-deprived adult. *Journal of Neurolinguistics*, *11*(1-2), 179–190. doi:10.1016/S0911-6044(98)00012-8
- Early, E. A. & VanDemark, A. A. (1985). Aphasic speakers' use of definite and indefinite articles to mark given and new information in discourse. *Clinical Aphasiology*, *15*, 248–254. Retrieved from <http://eprints-prod-05.library.pitt.edu/id/eprint/861>
- Edwards, J. (1981). *A cross-linguistic study of broca's aphasia* (Master's thesis, Massachusetts Institute of Technology). Retrieved from <http://hdl.handle.net/1721.1/15853>
- Edwards, S. (1995). Profiling fluent aphasic spontaneous speech: A comparison of two methodologies. *International Journal of Language & Communication Disorders*, *30*(3), 333–345. doi:10.3109/13682829509021446
- Edwards, S. & Bastiaanse, R. (1998). Diversity in the lexical and syntactic abilities of fluent aphasic speakers. *Aphasiology*, *12*(2), 99–117. doi:10.1080/02687039808250466
- Egg, M. & Redeker, G. (2010). How complex is discourse structure? In *Proceedings of the 7th International Conference on Language Resources and Evaluation, LREC 2010* (pp. 1619–1623). Valetta, Malta. Retrieved from <http://www.let.rug.nl/~rederker/EggRedekerLREC2010.pdf>
- Eling, P. (1994). *Reader in the history of aphasia: From Franz Gall to Norman Geschwind*. John Benjamins Publishing.
- Ellevåg, B., Foltz, P. W., Weinberger, D. R., & Goldberg, T. E. (2007). Quantifying incoherence in speech: an automated methodology and novel application to schizophrenia. *Schizophrenia research*, *93*(1-3), 304–316. doi:10.1016/j.schres.2007.03.001
- Fawcett, R. P. & Davies, B. L. (1992). Monologue as a turn in dialogue: Towards an integration of exchange structure and rhetorical structure theory. In *Proceedings of the 6th International Workshop*

- on *Natural Language Generation: Aspects of automated natural language generation* (pp. 151–166). Trento, Italy: Springer.
- Fawcett, T. (2006). An introduction to ROC analysis. *Pattern recognition letters*, 27(8), 861–874. doi:10.1016/j.patrec.2005.10.010
- Fergadiotis, G., Kapantzoglou, M., & Wright, H. H. (2011). Discourse coherence in adults with aphasia. In *Proceedings of the Clinical Aphasiology Conference*. Fort Lauderdale, FL. Retrieved from <http://aphasiology.pitt.edu/id/eprint/2277>
- Fergadiotis, G. & Wright, H. H. (2011). Using latent semantic analysis to capture global coherence in aphasic discourse. Platform presentation at the New Trends in Experimental Psycholinguistics Conference. Madrid, Spain.
- Fergadiotis, G., Wright, H. H., & Capilouto, G. J. (2011). Productive vocabulary across discourse types. *Aphasiology*, 25(10), 1261–1278. doi:10.1080/02687038.2011.606974
- Ferguson, A. & Harper, A. (2010). Contributions to the talk of individuals with aphasia in multiparty interactions. *Aphasiology*, 24(12), 1605–1620. doi:10.1080/02687038.2010.493296
- Féry, C. & Krifka, M. (2008). Information structure. In P. van Sterkenburg (Ed.), *Unity and diversity of languages*. John Benjamins Publishing Company. doi:10.1075/z.141.13kri
- Foka-Kavalieraki, P., Kakavoulia, M., Economou, A., Varlokosta, S., Routsis, C., Kasselimis, D., & Protopapas, A. (2008). A comprehensive approach to the analysis of narrative discourse production by Greek speakers with aphasia. Poster presented at the Science of Aphasia IX Conference. Retrieved from [http://users.uoa.gr/~aprotopapas/CV/pdf/Foka\\_etal\\_2008\\_SOA9.pdf](http://users.uoa.gr/~aprotopapas/CV/pdf/Foka_etal_2008_SOA9.pdf)
- Foltz, P. W. (2007). Discourse coherence and LSA. In T. K. Landauer, D. S. McNamara, S. Dennis, & W. Kintsch (Eds.), *Handbook of latent semantic analysis* (pp. 167–184). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.
- Forbes, M. M., Fromm, D., & MacWhinney, B. (2012). AphasiaBank: A resource for clinicians. In *Seminars in speech and language* (Vol. 33, 3, pp. 217–222). NIH Public Access.
- Fox, B. A. (1987). Anaphora and the structure of discourse. *Cambridge: Cambridge University Press*.
- Frattali, C. M., Thompson, C. K., Holland, A. L., Wohl, C. B., & Ferketic, M. (1995). *The FACS of life ASHA facts—a functional outcome measure for adults*. Rockville, MD: ASHA.
- Friedmann, N. (2006). Speech production in Broca’s agrammatic aphasia: Syntactic tree pruning. *Broca’s region*, 63–82. doi:10.1093/acprof:oso/9780195177640.001.0001
- Fromm, D., Forbes, M. M., Holland, A., & MacWhinney, B. (2014). Aphasiabank: 7-year interest rate index and yield. In *Proceedings of the Clinical Aphasiology Conference*. St. Simons Island, GA. Retrieved from <http://aphasiology.pitt.edu/id/eprint/2524>
- Fromm, D., Holland, A., Armstrong, E., Forbes, M., MacWhinney, B., Risko, A., & Mattison, N. (2011). “better but no cigar”: Persons with aphasia speak about their speech. *Aphasiology*, 25(11), 1431–1447. doi:10.1080/02687038.2011.608839
- Gernsbacher, M. A. & Givón, T. (1995). *Coherence in spontaneous text*. Amsterdam / Philadelphia: John Benjamins Publishing.
- Giora, R. (1985). Notes towards a theory of text coherence. *Poetics today*, 6(4), 699–715. doi:10.2307/1771962
- Giora, R. (1997). Discourse coherence and theory of relevance: Stumbling blocks in search of a unified theory. *Journal of pragmatics*, 27(1), 17–34. doi:10.1016/0378-2166(95)00065-8
- Givón, T. (1993). Coherence in text, coherence in mind. *Pragmatics & Cognition*, 1(2), 171–227. doi:10.1075/pc.1.2.01giv
- Gleason, J. B., Goodglass, H., Obler, L., Green, E., Hyde, M. R., & Weintraub, S. (1980). Narrative strategies of aphasic and normal-speaking subjects. *Journal of Speech, Language, and Hearing Research*, 23(2), 370–382. doi:10.1044/jshr.2302.370
- Glosser, G. (1993). Discourse production patterns in neurologically impaired and aged populations. In H. Brownell & Y. Joanette (Eds.), *Narrative discourse in neurologically impaired and normal aging adults* (pp. 191–211). Singular San Diego, CA: Singular Publishing Group, Inc.
- Glosser, G. & Deser, T. (1991). Patterns of discourse production among neurological patients with fluent language disorders. *Brain and language*, 40(1), 67–88. doi:10.1016/0093-934X(91)90117-J
- Goodglass, H., Christiansen, J. A., & Gallagher, R. (1993). Comparison of morphology and syntax in free narrative and structured tests: Fluent vs. nonfluent aphasics. *Cortex*, 29(3), 377–407. doi:10.1016/S0010-9452(13)80250-X
- Goodwin, C. (2000). Gesture, aphasia, and interaction. In D. McNeill (Ed.), *Language and gesture* (Vol. 2, pp. 84–98).

- Gordon, J. K. (2006). A quantitative production analysis of picture description. *Aphasiology*, 20(02-04), 188–204. doi:10.1080/02687030500472777
- Gordon, J. K. (2008). Measuring the lexical semantics of picture description in aphasia. *Aphasiology*, 22(7-8), 839–852. doi:10.1080/02687030701820063
- Grimes, J. E. (1975). *The thread of discourse*. Mouton Publishers.
- Groenewold, R., Bastiaanse, R., Nickels, L., & Huiskes, M. (2014). Perceived liveliness and speech comprehensibility in aphasia: the effects of direct speech in auditory narratives. *International journal of language & communication disorders*, 49(4), 486–497. doi:10.1111/1460-6984.12080
- Grosz, B. J. & Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational linguistics*, 12(3), 175–204. Retrieved from <http://aclweb.org/anthology/J86-3001>
- Grosz, B. J., Weinstein, S., & Joshi, A. K. (1995). Centering: a framework for modeling the local coherence of discourse. *Computational linguistics*, 21(2), 203–225. Retrieved from <http://www.aclweb.org/anthology/J95-2003>
- Gundel, J. K. (2003). Information structure and referential givenness/newness: How much belongs in the grammar? In *Proceedings of the 10th International Conference on Head-Driven Phrase Structure Grammar (HPSG)* (pp. 122–142). Michigan State University. Stanford, CA: CSLI Publications.
- Gundel, J. K. & Fretheim, T. (2004). Topic and focus. In L. Horn & G. Ward (Eds.), (Vol. 175, p. 196). Blackwell London. doi:10.1002/9780470756959.ch8
- Halliday, M. A. K. (1967). Notes on transitivity and theme in English: Part 2. *Journal of linguistics*, 3(2), 199–244. doi:10.1017/S0022226700016613
- Halliday, M. A. K. & Hasan, R. (1976). *Cohesion in English*. London: Routledge.
- Halliday, M. A. K. & Matthiessen, C. M. (1994). *An introduction to functional grammar* (2nd ed.). London: Edward Arnold.
- Hasan, R. (1985). The texture of a text. In M. A. K. Halliday & R. Hasan (Eds.), *Language, text and context: aspects of language in a social-semiotic perspective* (pp. 70–96). Geelong, Victoria: Deakin University.
- Hengst, J. A. (2003). Collaborative referencing between individuals with aphasia and routine communication partners. *Journal of Speech, Language, and Hearing Research*, 46(4), 831–848. doi:10.1044/1092-4388(2003/065)
- Hobbs, J. R. (1978). *Why is discourse coherent* (tech. rep. No. 176). AI Center, SRI International. Menlo Park, CA.
- Hobbs, J. R. (1979). Coherence and coreference. *Cognitive science*, 3(1), 67–90. doi:10.1207/s15516709cog0301\_4
- Hobbs, J. R. (1985). *On the coherence and structure of discourse* (tech. rep. No. CSLI-85-37). Stanford, CA: Center for the Study of Language and Information, Stanford University.
- Holland, A. L. (1982). Observing functional communication of aphasic adults. *Journal of Speech and Hearing Disorders*, 47(1), 50–56. doi:10.1044/jshd.4701.50
- Holland, A. L., Frattali, C., & Fromm, D. (1999). *Communication activities of daily living – Second edition (CADL-2)*. Austin, TX: Pro-Ed.
- Hough, M. S. (1990). Narrative comprehension in adults with right and left hemisphere brain-damage: theme organization. *Brain and language*, 38(2), 253–277. doi:10.1016/0093-934X(90)90114-V
- Hovy, E. H. (1993). Automated discourse generation using discourse structure relations. *Artificial intelligence*, 63(1-2), 341–385. doi:10.1016/0004-3702(93)90021-3
- Huber, W. (1990). Text comprehension and production in aphasia: Analysis in terms of micro-and macro-processing. In Y. Joannette & H. Brownell (Eds.), *Discourse ability and brain damage* (pp. 154–179). New York, NY: Springer. doi:10.1007/978-1-4612-3262-9\_7
- Hunt, K. W. (1970). Syntactic maturity in schoolchildren and adults. *Monographs of the society for research in child development*, 35(1), iii–67. doi:10.2307/1165818
- Im, E.-J., Kwon, M., & Sim, H.-S. (2001). The informativeness and efficiency of the connected speech samples in Korean fluent aphasics. *Communication Sciences & Disorders*, 6(2), 374–391. Retrieved from [https://www.e-csd.org/upload/6\(2\)\\_5.pdf](https://www.e-csd.org/upload/6(2)_5.pdf)
- Iruskieta, M., Da Cunha, I., & Taboada, M. (2015). A qualitative comparison method for rhetorical structures: identifying different discourse structures in multilingual corpora. *Language resources and evaluation*, 49(2), 263–309. doi:10.1007/s10579-014-9271-6
- Irwin, W. H., Wertz, R. T., & Avent, J. R. (2002). Relationships among language impairment, functional communication, and pragmatic performance in aphasia. *Aphasiology*, 16(8), 823–835. doi:10.1080/02687030244000086

- Jacobs, B. J. (2001). Social validity of changes in informativeness and efficiency of aphasic discourse following linguistic specific treatment (LST). *Brain and Language*, 78(1), 115–127. doi:10.1006/brln.2001.2452
- Joshi, A. K. & Weinstein, S. (1981). Control of inference: role of some aspects of discourse structure-centering. In *Proceedings of the 7th International Joint Conference of Artificial Intelligence (IJCAI)* (pp. 385–387). Vancouver, BC.
- Juliano, D. A. & Cassim, F. T. R. (2012). Coherence relations in academic spoken discourse. *Linguistica*, 52(1), 323. doi:10.4312/linguistica.52.1.323-336
- Keenan, J. M., Baillet, S. D., & Brown, P. (1984). The effects of causal cohesion on comprehension and memory. *Journal of Verbal Learning and Verbal Behavior*, 23(2), 115. doi:10.1016/S0022-5371(84)90082-3
- Kehler, A. (1995). *Interpreting cohesive forms in the context of discourse inference* (tech. rep. No. TR-11-95). Retrieved from <http://nrs.harvard.edu/urn-3:HUL.InstRepos:25104741>
- Kehler, A. (2002). *Coherence, reference, and the theory of grammar*. Stanford, CA: CSLI publications.
- Kehler, A. (2004). Discourse coherence. In L. Horn & G. Ward (Eds.), *The handbook of pragmatics* (Vol. 26, pp. 241–265). Oxford, UK: Blackwell Publishing.
- Kempler, D. & Goral, M. (2011). A comparison of drill-and communication-based treatment for aphasia. *Aphasiology*, 25(11), 1327–1346. doi:10.1080/02687038.2011.599364
- Kibrik, A. A., Khudyakova, M. V., Dobrov, G. B., & Linnik, A. (2013). Referential choice: A cognitively based modeling study. In *Proceedings of the PRE-CogSci 2013 Workshop on the Production of Referring Expressions*. Berlin. Retrieved from <https://pre2013.uvt.nl/pdf/kibrik-khudyakova-dobrov-linnik.pdf>
- Kibrik, A. A., Khudyakova, M. V., Dobrov, G. B., Linnik, A., & Zalmanov, D. A. (2016). Referential choice: Predictability and its limits. *Frontiers in psychology*, 7, 1429. doi:10.3389/fpsyg.2016.01429
- Kibrik, A. A. & Podlesskaya, V. I. (2009). *Rasskazy o snovidenijach: korpusnoe issledovanie usnogo russkogo diskursa [Night Dream Stories: A corpus study of spoken Russian discourse]*. Moscow: Jazyki Slavjanskich Kultur.
- Kintsch, W. (1994). Text comprehension, memory, and learning. *American Psychologist*, 49(4), 294. doi:10.1037/0003-066X.49.4.294
- Kintsch, W. & Van Dijk, T. A. (1978). Toward a model of text comprehension and production. *Psychological review*, 85(5), 363. doi:10.1037/0033-295X.85.5.363
- Klippi, A. (2015). Pointing as an embodied practice in aphasic interaction. *Aphasiology*, 29(3), 337–354. doi:10.1080/02687038.2013.878451
- Knott, A., Sanders, T., & Oberlander, J. (2001). Levels of representation in discourse relations. *Cognitive Linguistics*, 12(3), 197–210. doi:10.1515/cogl.2002.001
- Kong, A. P.-H. & Law, S.-P. (2004). A Cantonese linguistic communication measure for evaluating aphasic narrative production: normative and preliminary aphasic data. *Journal of Multilingual Communication Disorders*, 2(2), 124–146. doi:10.1080/14769670310001659708
- Kong, A. P.-H. & Law, S.-P. (2009). A linguistic communication measure for monitoring changes in Chinese aphasic narrative production. *Clinical linguistics & phonetics*, 23(4), 255–269. doi:10.1080/02699200802673234
- Kong, A. P.-H., Law, S.-P., & Lee, A. (2012). The construction of a corpus of Cantonese-aphasic-discourse: a preliminary report. Poster presented at the American Speech-Language-Hearing Association (ASHA) Convention. New Orleans, LA.
- Kong, A. P.-H., Linnik, A., Law, S.-P., & Shum, W. W.-M. (2018). Measuring discourse coherence in anomic aphasia using Rhetorical Structure Theory. *International journal of speech-language pathology*, 20(4), 406–421. Published online: 17 Mar 2017. doi:10.1080/17549507.2017.1293158
- Korpijaakko-Huuhka, A.-M. & Lind, M. (2012). The impact of aphasia on textual coherence: Evidence from two typologically different languages. *Journal of Interactional Research in Communication Disorders*, 3(1), 47. doi:10.1558/jircd.v3i1.47
- Koutsoftas, A. D., Wright, H. H., & Capilouto, G. J. (2009). Discourse coherence in healthy younger and older adults. Poster presented at the American Speech-Language-Hearing Association (ASHA) Convention. New Orleans, LA.
- Krifka, M. (2008). Basic notions of information structure. *Acta Linguistica Hungarica*, 55(3-4), 243–276. doi:10.1556/ALing.55.2008.3-4.2
- Kuhn, M. (2008). Building predictive models in R using the caret package. *Journal of Statistical Software*, 28(5), 1–26. doi:10.18637/jss.v028.i05

- Laakso, M. (2015). Collaborative participation in aphasic word searching: Comparison between significant others and speech and language therapists. *Aphasiology*, *29*(3), 269–290. doi:10.1080/02687038.2013.878450
- Labov, W. (1972). *Language in the inner city: Studies in the Black English vernacular*. Philadelphia, PA: University of Pennsylvania Press.
- Laine, M., Laakso, M., Vuorinen, E., & Rinne, J. (1998). Coherence and informativeness of discourse in two dementia types. *Journal of Neurolinguistics*, *11*(1-2), 79–87. doi:10.1016/S0911-6044(98)00006-2
- Lambrecht, K. (1994). *Information structure and sentence form: topic, focus, and the mental representations of discourse referents*. Cambridge studies in linguistics. New York, NY: Cambridge university press.
- Landauer, T. K. & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological review*, *104*(2), 211. doi:10.1037/0033-295X.104.2.211
- 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
- Lascarides, A. & Asher, N. (2008). Segmented discourse representation theory: Dynamic semantics with discourse structure. In H. Bunt & R. Muskens (Eds.), *Computing meaning* (pp. 87–124). Springer. doi:10.1007/978-1-4020-5958-2\_5
- Lenk, U. (1998). *Marking discourse coherence: Functions of discourse markers in spoken English*. Language in performance. Tübingen: Gunter Narr Verlag.
- Liaw, A. & Wiener, M. (2002). Classification and regression by randomForest. *R News*, *2*(3), 18–22. Retrieved from <http://CRAN.R-project.org/doc/Rnews/>
- Liles, B. Z. (1985). Cohesion in the narratives of normal and language-disordered children. *Journal of Speech, Language, and Hearing Research*, *28*(1), 123–133. doi:10.1044/jshr.2801.123
- Liles, B. Z. & Coelho, C. A. (1998). Cohesion analysis. In L. R. Cherney, C. A. Coelho, & B. B. Shadden (Eds.), *Analyzing discourse in communicatively impaired adults* (pp. 65–84). Gaithersburg, MA: Aspen.
- Linebaugh, C. W., Kryzer, K. M., Oden, S., & Myers, P. S. (2006). Reapportionment of communicative burden in aphasia: A study of narrative interaction. *Aphasiology*, *20*(1), 84–96. doi:10.1080/02687030500369452
- Linnik, A., Bastiaanse, R., & Höhle, B. (2016). Discourse production in aphasia: a current review of theoretical and methodological challenges. *Aphasiology*, *30*(7), 765–800. doi:10.1080/02687038.2015.1113489
- Lomas, J., Pickard, L., Bester, S., Elbard, H., Finlayson, A., & Zoghaib, C. (1989). The communicative effectiveness index: Development and psychometric evaluation of a functional communication measure for adult aphasia. *Journal of speech and hearing disorders*, *54*(1), 113–124. doi:10.1044/jshd.5401.113
- Longacre, R. E. (1996). *The grammar of discourse*. Topics in Language and Linguistics. Boston, MA: Springer.
- Lorch Jr, R. F. E. & O'Brien, E. J. (1995). *Sources of coherence in reading*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Luria, A. R. (1966a). *Higher cortical functions in man*. Springer Science & Business Media.
- Luria, A. R. (1966b). *Human brain and psychological processes*. New York, NY: Harper & Row.
- MacWhinney, B. (2000). *The CHILDES project: Tools for analyzing talk* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum Associates.
- MacWhinney, B., Fromm, D., Forbes, M., & Holland, A. (2011). AphasiaBank: Methods for studying discourse. *Aphasiology*, *25*(11), 1286–1307. doi:10.1080/02687038.2011.589893
- MacWhinney, B., Fromm, D., Holland, A., & Forbes, M. (2013). AphasiaBank: Data and methods. In N. Müller & M. Ball (Eds.), *Research methods in clinical linguistics and phonetics* (pp. 268–287). New York, NY: Wiley.
- Mann, W. C. & Thompson, S. A. (1985). Assertions from discourse structure. In *Proceedings of the 11th Annual Meeting of the Berkeley Linguistics Society* (pp. 245–258). Berkeley, CA. doi:10.3765/bls.v11i0.1886
- Mann, W. C. & Thompson, S. A. (1987). *Rhetorical structure theory: A theory of text organization* (tech. rep. No. ISI/RS-87-190). Los Angeles: Information Sciences Institute, University of Southern California. Retrieved from [http://www.sfu.ca/rst/05bibliographies/bibs/ISI\\_RS\\_87\\_190.pdf](http://www.sfu.ca/rst/05bibliographies/bibs/ISI_RS_87_190.pdf)

- Mann, W. C. & Thompson, S. A. (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text-Interdisciplinary Journal for the Study of Discourse*, 8(3), 243–281. doi:10.1515/text.1.1988.8.3.243
- Mann, W. C. & Thompson, S. A. (2001). Deux perspectives sur la théorie de la structure rhétorique (rst)[two views of rhetorical structure theory]. *Verbum*, (1), 9–29. Retrieved from <http://www-bcf.usc.edu/~billmann/WMLinguistic/2vsend.pdf>
- Manochiopinig, S., Sheard, C., & Reed, V. A. (1992). Pragmatic assessment in adult aphasia: A clinical review. *Aphasiology*, 6(6), 519–533. doi:10.1080/02687039208249489
- Marcu, D. (1997). *The rhetorical parsing, summarization, and generation of natural language texts* (Doctoral dissertation, University of Toronto, Toronto).
- Marcu, D. (1998). Improving summarization through rhetorical parsing tuning. In *Very large corpora (Proceedings of the COLING-ACL'98 workshop)* (pp. 206–215). Montreal. Retrieved from <http://www.aclweb.org/anthology/W98-1124>
- Marcu, D. (1999). Discourse trees are good indicators of importance in text. In I. Mani & M. T. Maybury (Eds.), *Advances in automatic text summarization* (pp. 123–136). Cambridge, MA: MIT Press.
- Marcu, D. (2000). The rhetorical parsing of unrestricted texts: A surface-based approach. *Computational linguistics*, 26(3), 395–448. doi:10.1162/089120100561755
- Marcu, D. (2003). Discourse Structures: Trees or graphs? <https://www.isi.edu/~marcu/discourse/DiscourseStructures.htm>.
- Marcu, D., Amorrortu, E., & Romera, M. (1999). Experiments in constructing a corpus of discourse trees. In *Towards standards and tools for discourse tagging. (Proceedings of the ACL'99 workshop)* (pp. 48–57). College Park, Maryland. Retrieved from <http://www.aclweb.org/anthology/W99-0307>
- 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. doi:10.1080/02687038.2011.584690
- Marini, A., Boewe, A., Caltagirone, C., & Carlomagno, S. (2005). Age-related differences in the production of textual descriptions. *Journal of psycholinguistic research*, 34(5), 439–463. doi:10.1007/s10936-005-6203-z
- Marini, A., Carlomagno, S., Caltagirone, C., & Nocentini, U. (2005). The role played by the right hemisphere in the organization of complex textual structures. *Brain and Language*, 93(1), 46–54. doi:10.1016/j.bandl.2004.08.002
- Marini, A. & Fabbro, F. (2007). Psycholinguistic models of speech production in bilingualism and multilingualism. In A. Ardila & E. Ramos (Eds.), *Speech and language disorders in bilinguals* (pp. 47–67). New York, NY: Nova Science Publishers.
- Marini, A., Spoletini, I., Rubino, I. A., Ciuffa, M., Bria, P., Martinotti, G., ... Siracusano, A., et al. (2008). The language of schizophrenia: An analysis of micro and macrolinguistic abilities and their neuropsychological correlates. *Schizophrenia research*, 105(1-3), 144–155. doi:10.1016/j.schres.2008.07.011
- Martin, J. R. (2001). Cohesion and texture. *The handbook of discourse analysis*, 35–53.
- McCullough, K. C., McCullough, G. H., Ruark, J. L., & Rainey, J. (2006). Pragmatic performance and functional communication in adults with aphasia. *The Journal of Speech and Language Pathology–Applied Behavior Analysis*, 1(2), 164–178. doi:10.1037/h0100193
- McNamara, D. S., Louwerse, M. M., McCarthy, P. M., & Graesser, A. C. (2010). Coh-matrix: capturing linguistic features of cohesion. *Discourse Processes*, 47(4), 292–330. doi:10.1080/01638530902959943
- McNeil, M. R., Doyle, P. J., Fossett, T. R., 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. doi:10.1080/02687040143000348
- Menn, L., Obler, L. K., & Miceli, G. (1990). *Agrammatic aphasia: A cross-language narrative sourcebook*. Amsterdam / Philadelphia: John Benjamins Publishing.
- Meuse, S. & Marquardt, T. P. (1985). Communicative effectiveness in broca's aphasia. *Journal of Communication Disorders*, 18(1), 21–34. doi:10.1016/0021-9924(85)90011-5
- Miceli, G., Silveri, M. C., Romani, C., & Caramazza, A. (1989). Variation in the pattern of omissions and substitutions of grammatical morphemes in the spontaneous speech of so-called agrammatic patients. *Brain and language*, 36(3), 447–492. doi:10.1016/0093-934X(89)90079-5
- Miltsakaki, E., Prasad, R., Joshi, A. K., & Webber, B. L. (2004). The Penn Discourse Treebank. In *Proceedings of LREC'04*. Lisbon. Retrieved from <https://www.seas.upenn.edu/~pdtb/papers/lrec04.pdf>

- Mirman, D., Strauss, T. J., Brecher, A., Walker, G. M., Sobel, P., Dell, G. S., & Schwartz, M. F. (2010). A large, searchable, web-based database of aphasic performance on picture naming and other tests of cognitive function. *Cognitive neuropsychology*, *27*(6), 495–504. doi:10.1016/j.sbspro.2010.08.066v
- Mitchell, M. W. (2011). Bias of the random forest out-of-bag (OOB) error for certain input parameters. *Open Journal of Statistics*, *1*(3), 205–211. doi:10.4236/ojs.2011.13024
- Mol, L., Krahmer, E., & van de Sandt-Koenderman, M. (2013). Gesturing by speakers with aphasia: How does it compare? *Journal of Speech, Language, and Hearing Research*, *56*(4), 1224–1236. doi:10.1044/1092-4388(2012/11-0159)
- Moore, J. D. & Pollack, M. E. (1992). A problem for RST: the need for multi-level discourse analysis. *Computational linguistics*, *18*(4), 537–544. Retrieved from <http://aclweb.org/anthology/J92-4007>
- Moser, M. & Moore, J. D. (1996). Toward a synthesis of two accounts of discourse structure. *Computational linguistics*, *22*(3), 409–419. doi:http://www.aclweb.org/anthology/J96-3006
- Muò, R., Cancialosi, P., Galimberti, L., Carrubba Cacciola, B., Gilardone, M., & Schindler, A. (2015). Validation of the Italian version of the American Speech-Language and Hearing Association – Functional Assessment of Communication Skills for adults (I-ASHA-FACS). *Aphasiology*, *29*(9), 1110–1130. doi:10.1080/02687038.2015.1010475
- Nakatani, C. H., Hirschberg, J., & Grosz, B. J. (1995). Discourse structure in spoken language: Studies on speech corpora. Paper presented at 95' AAAI Spring Symposium on Empirical Methods in Discourse Interpretation and Generation. Palo Alto. Retrieved from <http://nrs.harvard.edu/urn-3:HUL.InstRepos:2580299>
- 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, Language, and Hearing Research*, *36*(2), 338–350. doi:10.1044/jshr.3602.338
- 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, Language, and Hearing Research*, *38*(1), 145–156. doi:10.1044/jshr.3801.145
- Nicholas, M., Obler, L. K., Albert, M. L., & Helm-Estabrooks, N. (1985). Empty speech in Alzheimer's disease and fluent aphasia. *Journal of Speech, Language, and Hearing Research*, *28*(3), 405–410. doi:10.1044/jshr.2803.405
- O'Donnell, M. (2000a). RSTTool 2.4: a markup tool for Rhetorical Structure Theory. In *Proceedings of the first International Conference on Natural Language Generation* (Vol. 14, pp. 253–256). Association for Computational Linguistics. Mitzpe Ramon, Israel. Retrieved from <http://www.aclweb.org/anthology/W00-1434>
- O'Donnell, M. (2000b). RST-Tool: An RST analysis tool. In *Proceedings of the 6th European Workshop on Natural Language Generation*. Gerhard-Mercator University. Duisburg, Germany. Retrieved from <http://www.wagsoft.com/Papers/rsttool97.pdf>
- 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. doi:10.1044/jslhr.4203.636
- Olness, G. S. (2006). Genre, verb, and coherence in picture-elicited discourse of adults with aphasia. *Aphasiology*, *20*(02-04), 175–187. doi:10.1080/02687030500472710
- Olness, G. S. (2007). Narrative structure in aphasia: Impact of task and aphasia severity. In *Proceedings of the 37th Clinical Aphasiology Conference*. Scottsdale, AZ. Retrieved from <http://aphasiology.pitt.edu/1890/>
- Olness, G. S. & Gober, J. M. (2013). Mild conduction aphasia and narrative functionality: beyond the assessment of impairment (a case study). In *Proceedings of the 43rd Clinical Aphasiology Conference*. Tucson, AZ. Retrieved from <http://aphasiology.pitt.edu/2478/>
- Olness, G. S., Matteson, S. E., & Stewart, C. T. (2010). "Let me tell you the point": How speakers with aphasia assign prominence to information in narratives. *Aphasiology*, *24*(6-8), 697–708. doi:10.1080/02687030903438524
- Olness, G. S. & Ulatowska, H. K. (2011). Personal narratives in aphasia: Coherence in the context of use. *Aphasiology*, *25*(11), 1393–1413. doi:10.1080/02687038.2011.599365
- Olness, G. S., Ulatowska, H. K., Carpenter, C. M., Williams-Hubbard, L. J., & Dykes, J. C. (2005). Holistic assessment of narrative quality: A social validation study. *Aphasiology*, *19*(3-5), 251–262. doi:10.1080/02687030444000723
- 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. doi:10.1080/02687030244000095

- Packard, J. L. (2000). *The morphology of Chinese: A linguistic and cognitive approach*. Cambridge, UK: Cambridge University Press.
- Pedersen, P. M., Vinter, K., & Olsen, T. S. (2001). The communicative effectiveness index: psychometric properties of a Danish adaptation. *Aphasiology*, 15(8), 787–802. doi:10.1080/02687040143000195
- Penn, C. (1988). The profiling of syntax and pragmatics in aphasia. *Clinical linguistics & phonetics*, 2(3), 179–207. doi:10.1080/02699208808985255
- Penn, C., Frankel, T., & Wilkinson, R. (2015). Problems with the understandability of aphasic talk: Mentions of persons as a trouble source in interaction. *Aphasiology*, 29(3), 291–314. doi:10.1080/02687038.2014.986632
- Perkins, L. (1995). Applying conversation analysis to aphasia: Clinical implications and analytic issues. *European Journal of Disorders of Communication*, 30(3), 372–383. doi:10.3109/13682829509021449
- Perkins, L. (2003). Negotiating repair in aphasic conversation. In C. Goodwin (Ed.), *Conversation and brain damage* (pp. 147–162). New York, NY: Oxford University Press.
- Piehler, M. F. & Holland, A. L. (1984). Cohesion in aphasic language. In *Proceedings of the 14th Clinical Aphasiology Conference* (pp. 208–214). BRK Publishers. Seabrook Island, SC. Retrieved from <http://eprints-prod-05.library.pitt.edu/813/1/14-25.pdf>
- Poesio, M., Patel, A., & Di Eugenio, B. (2006). Discourse structure and anaphora in tutorial dialogues: An empirical analysis of two theories of the global focus. *Research on Language and Computation*, 4(2–3), 229–257. doi:10.1007/s11168-006-9005-z
- Polanyi, L. (1988). A formal model of the structure of discourse. *Journal of pragmatics*, 12(5–6), 601–638. doi:10.1016/0378-2166(88)90050-1
- Pomeroy, S. L., Tamayo, P., Gaasenbeek, M., Sturla, L. M., Angelo, M., McLaughlin, M. E., . . . Golub, T. R. (2002). Prediction of central nervous system embryonal tumour outcome based on gene expression. *Nature*, 415(6870), 436. doi:10.1038/415436a
- Porch, B. E. (1967). *Porch index of communicative ability: Theory and development*. Palo Alto: Consulting Psychologists Press.
- Prasad, R., Dinesh, N., Lee, A., Miltsakaki, E., Robaldo, L., Joshi, A., & Webber, B. L. (2008). The Penn Discourse Treebank 2.0. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC'08)*. European Language Resources Association (ELRA). Marrakech, Morocco. Retrieved from <https://www.seas.upenn.edu/~pdtb/papers/pdtb-lrec08.pdf>
- Prins, R. & Bastiaanse, R. (2004). Review. *Aphasiology*, 18(12), 1075–1091. doi:10.1080/02687030444000534
- 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. doi:10.1080/02687038.2014.993912
- Prutting, C. A. & Kirchner, D. M. (1983). Applied pragmatics. In T. M. Gallagher & C. A. Prutting (Eds.). San Diego, CA: Colledge Hill Press.
- Ramsberger, G. & Rende, B. (2002). Measuring transactional success in the conversation of people with aphasia. *Aphasiology*, 16(3), 337–353. doi:10.1080/02687040143000636
- Redeker, G. (1990). Ideational and pragmatic markers of discourse structure. *Journal of pragmatics*, 14(3), 367–381. doi:10.1016/0378-2166(90)90095-U
- Redeker, G. (2000). Coherence and structure in text and discourse. In H. Bunt & W. Black (Eds.), *Abduction, belief and context in dialogue* (pp. 233–263). Amsterdam / Philadelphia: John Benjamins.
- Reinhart, T. (1980). Conditions for text coherence. *Poetics today*, 1(4), 161–180. doi:10.2307/1771893
- Reitter, D. & Stede, M. (2003). Step by step: underspecified markup in incremental rhetorical analysis. In *Proceedings of 4th International Workshop on Linguistically Interpreted Corpora (LINC-03) at EACL 2003*. Budapest, Hungary. Retrieved from <http://www.aclweb.org/anthology/W03-2411>
- Renkema, J. (2009). *Discourse, of course: An overview of research in discourse studies*. Amsterdam / Philadelphia: John Benjamins Publishing.
- Ripich, D. N. & Terrell, B. Y. (1988). Patterns of discourse cohesion and coherence in Alzheimer's disease. *Journal of Speech and Hearing Disorders*, 53(1), 8–15. doi:10.1044/jshd.5301.08
- Roberts, C. (2012, December). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics*, 5(6), 1–69. doi:10.3765/sp.5.6
- Rogalski, Y., Altmann, L. J., Plummer-D'Amato, P., Behrman, A. L., & Marsiske, M. (2010). Discourse coherence and cognition after stroke: A dual task study. *Journal of communication disorders*, 43(3), 212–224. doi:10.1016/j.jcomdis.2010.02.001
- Rossi, E. & Bastiaanse, R. (2008). Spontaneous speech in Italian agrammatic aphasia: A focus on verb production. *Aphasiology*, 22(4), 347–362. doi:10.1080/02687030701407093

- Ruigendijk, E., Vasić, N., & Avrutin, S. (2006). Reference assignment: Using language breakdown to choose between theoretical approaches. *Brain and Language*, *96*(3), 302–317. doi:10.1016/j.bandl.2005.06.005
- Ruiter, M. B., Kolk, H. H., Rietveld, T. C., Dijkstra, N., & Lotgering, E. (2011). Towards a quantitative measure of verbal effectiveness and efficiency in the Amsterdam-Nijmegen Everyday Language Test (ANELT). *Aphasiology*, *25*(8), 961–975. doi:10.1080/02687038.2011.569892
- Sacks, H. (1972). An initial investigation of the usability of conversational data for doing sociology. In D. N. Sudnow (Ed.), *Studies in social interaction* (pp. 31–73). New York, NY: The Free Press.
- Sacks, H., Schegloff, E. A., & Jefferson, G. (1978). A simplest systematics for the organization of turn taking for conversation. In J. Schenkein (Ed.), *Studies in the organization of conversational interaction* (pp. 7–55). Elsevier. doi:10.1016/B978-0-12-623550-0.50008-2
- Saffran, E. M., Berndt, R. S., & Schwartz, M. F. (1989). The quantitative analysis of agrammatic production: Procedure and data. *Brain and language*, *37*(3), 440–479. doi:10.1016/0093-934X(89)90030-8
- Saldert, C., Johansson, C., & Wilkinson, R. (2015). An interaction-focused intervention approach to training everyday communication partners: A single case study. *Aphasiology*, *29*(3), 378–399. doi:10.1080/02687038.2014.948803
- Sanders, T. (1993). On the cognitive and text-analytic status of coherence relations. In *Intentionality and structure in discourse relations: Proceedings of a workshop sponsored by the Special Interest Group on Generation of the Association for Computational Linguistics*. Columbus, OH. Retrieved from <http://www.aclweb.org/anthology/W93-0230>
- Sanders, T. (1997). Semantic and pragmatic sources of coherence: On the categorization of coherence relations in context. *Discourse processes*, *24*(1), 119–147. doi:10.1080/01638539709545009
- Sanders, T. & Pander Maat, H. (2006). Cohesion and coherence: Linguistic approaches. In K. Brown (Ed.), *The encyclopaedia of language and linguistics* (Vol. 2, pp. 591–595). Amsterdam: Elsevier. doi:10.1016/B0-08-044854-2/00497-1
- Sanders, T. & Spooren, W. (1999). Communicative intentions and coherence relations. In W. Bublitz, U. Lenk, & E. Ventola (Eds.), *Coherence in spoken and written discourse* (pp. 235–250). Amsterdam / Philadelphia: John Benjamins Publishing.
- Sanders, T. & Spooren, W. (2001). Text representation as an interface between language and its users. In T. Sanders, J. Schilperoord, & W. Spooren (Eds.), *Text representation: Linguistic and psycholinguistic aspects* (pp. 1–26). Amsterdam / Philadelphia: John Benjamins Publishing.
- Sanders, T., Spooren, W. P., & Noordman, L. G. (1992). Toward a taxonomy of coherence relations. *Discourse processes*, *15*(1), 1–35. doi:10.1080/01638539209544800
- Scharp, V. L., Tompkins, C. A., & Iverson, J. M. (2007). Gesture and aphasia: Helping hands? *Aphasiology*, *21*(6-8), 717–725. doi:10.1080/02687030701192273
- Schiffrin, D. (1994). *Approaches to discourse*. Blackwell textbooks in linguistics. Oxford, UK: Blackwell Publishers.
- Segal, M. R. (2004). *Machine learning benchmarks and random forest regression*. Center for Bioinformatics and Molecular Biostatistics, University of California, San Francisco. Los Angeles, CA. Retrieved from <https://escholarship.org/uc/item/35x3v9t4>
- Shapiro, L. R. & Hudson, J. A. (2014). Coherence and cohesion in children’s stories. In J. Costermans & M. Fayol (Eds.), *Processing interclausal relationships* (pp. 33–58). New York, NY: Psychology Press.
- Sherratt, S. (2007). Multi-level discourse analysis: A feasible approach. *Aphasiology*, *21*(3-4), 375–393. doi:10.1080/02687030600911435
- Sherratt, S. M. & Penn, C. (1990). Discourse in a right-hemisphere brain-damaged subject. *Aphasiology*, *4*(6), 539–560. doi:10.1080/02687039008248506
- Sherratt, S. & Bryan, K. (2012). Discourse production after right brain damage: Gaining a comprehensive picture using a multi-level processing model. *Journal of Neurolinguistics*, *25*(4), 213–239. doi:10.1016/j.jneuroling.2012.01.001
- Stark, J. A. (2010). Content analysis of the fairy tale Cinderella – A longitudinal single-case study of narrative production: “From rags to riches”. *Aphasiology*, *24*(6-8), 709–724. doi:10.1080/02687030903524729
- Stede, M. (2004). The Potsdam Commentary Corpus. In *Proceedings of the 2004 ACL Workshop on Discourse Annotation* (pp. 96–102). Association for Computational Linguistics. Barcelona, Spain. Retrieved from <https://www.aclweb.org/anthology/W/W04/W04-0213.pdf>
- Stede, M. (2008). Disambiguating rhetorical structure. *Research on Language and Computation*, *6*(3-4), 311–332. doi:10.1007/s11168-008-9053-7

- Stede, M. (2011). *Discourse processing* (G. Hirst, Ed.). Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers. doi:10.2200/S00354ED1V01Y201111HHLT015
- Stent, A. (2000). Rhetorical structure in dialog. In *Proceedings of the first International Conference on Natural Language Generation* (pp. 247–252). Association for Computational Linguistics. Mitzpe Ramon, Israel. Retrieved from <http://www.aclweb.org/anthology/W00-1433>
- Swindell, C. S., Holland, A. L., & Fromm, D. (1984). Classification of aphasia: WAB type versus clinical impression. In *Proceedings of the 14th Clinical Aphasiology conference* (pp. 48–54). Seabrook Island, SC: BRK Publishers. Retrieved from <http://eprints-prod-05.library.pitt.edu/795/1/14-07.pdf>
- Taboada, M. (2001). *Collaborating through talk: The interactive construction of task-oriented dialogue in English and in Spanish* (Doctoral dissertation, Universidad Complutense de Madrid).
- Taboada, M. (2004). *Building coherence and cohesion: Task-oriented dialogue in English and Spanish*. Pragmatics & Beyond New Series. Amsterdam / Philadelphia: John Benjamins Publishing.
- Taboada, M. (2006). Discourse markers as signals (or not) of rhetorical relations. *Journal of pragmatics*, 38(4), 567–592. doi:10.1016/j.pragma.2005.09.010
- Taboada, M. & Mann, W. C. (2006a). Applications of rhetorical structure theory. *Discourse studies*, 8(4), 567–588. doi:10.1177/1461445606064836
- Taboada, M. & Mann, W. C. (2006b). Rhetorical structure theory: Looking back and moving ahead. *Discourse studies*, 8(3), 423–459. doi:10.1177/1461445606061881
- Tanskanen, S.-K. (2006). *Collaborating towards coherence: Lexical cohesion in English discourse*. Pragmatics & Beyond New Series. Amsterdam / Philadelphia: John Benjamins Publishing.
- Tesak, J. & Code, C. (2008). *Milestones in the history of aphasia: Theories and protagonists*. London: Psychology Press.
- Thompson, C. K., Meltzer-Asscher, A., Cho, S., Lee, J., Wieneke, C., Weintraub, S., Mesulam, M., et al. (2013). Syntactic and morphosyntactic processing in stroke-induced and primary progressive aphasia. *Behavioural neurology*, 26(1-2), 35–54. doi:10.3233/BEN-2012-110220
- Thompson, S. (1994). Aspects of cohesion in monologue. *Applied linguistics*, 15(1), 58–75. doi:10.1093/applin/15.1.58
- Ulatowska, H. K., Allard, L., & Chapman, S. B. (1990). Narrative and procedural discourse in aphasia. In Y. Joannette & H. H. Brownell (Eds.), (pp. 180–198). New York, NY: Springer. doi:10.1007/978-1-4612-3262-9\_8
- Ulatowska, H. K. & Chapman, S. B. (1994). Discourse macrostructure in aphasia. In R. L. Bloom, L. K. Obler, S. De Santi, & J. S. Ehrlich (Eds.), *Discourse analysis and applications: Studies in adult clinical populations* (pp. 29–46). Hillsdale, NJ: Lawrence A. Erlbaum.
- Ulatowska, H. K., Chapman, S. B., Johnson, J., & Branch, C. (1999). Macrostructure and inferential processing in discourse of aphasic patients. *Psychology of Language and Communication*, 3(2), 3–20.
- Ulatowska, H. K., Doyel, A. W., Stern, R. F., Haynes, S. M., & North, A. J. (1983). Production of procedural discourse in aphasia. *Brain and Language*, 18(2), 315–341. doi:10.1016/0093-934X(83)90023-8
- Ulatowska, H. K., Hayashi, M. M., Cannito, M. P., & Fleming, S. G. (1986). Disruption of reference in aging. *Brain and language*, 28(1), 24–41. doi:10.1016/0093-934X(86)90088-X
- Ulatowska, H. K., North, A. J., & Macaluso-Haynes, S. (1981). Production of narrative and procedural discourse in aphasia. *Brain and language*, 13(2), 345–371. doi:10.1016/0093-934X(81)90100-0
- Ulatowska, H. K. & Olness, G. S. (1997). Some observations on narratives by aphasics and their contributions to narrative theory. *Journal of Narrative and Life History*, 7(1), 259–264. doi:10.1075/jnlh.7.32som
- Ulatowska, H. K. & Olness, G. S. (2000). Discourse revisited: Contributions of lexico-syntactic devices. *Brain and Language*, 71(1), 249–251. doi:10.1006/brln.1999.2261
- Ulatowska, H. K. & Olness, G. S. (2007). Pragmatics in discourse performance: Insights from aphasiology. 28(02), 148–158. doi:10.1055/s-2007-970572
- Ulatowska, H. K., Olness, G. S., Samson, A., Keebler, M., & Goins, K. (2004). On the nature of personal narratives of high quality. *Advances in Speech Language Pathology*, 6(1), 3–14. doi:10.1080/14417040410001669453
- Ulatowska, H. K., Olness, G. S., Wertz, R., Samson, A., Keebler, M., & Goins, K. (2003). Relationship between discourse and Western Aphasia Battery performance in African Americans with aphasia. *Aphasiology*, 17(5), 511–521. doi:10.1080/0268703034400102
- Ulatowska, H. K., Olness, G. S., & Williams, L. J. (2004). Coherence of narratives in aphasia. *Brain and Language*, 1(91), 42–43. doi:10.1016/j.bandl.2004.06.024

- 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. doi:10.1310/tsr2001-36
- Ulatowska, H. K. & Sadowska, M. (1992). Some observations on aphasic texts. In S. J. J. Hwang & W. R. Merrifield (Eds.), *Language in context: essays for robert e. longacre* (Vol. 107, pp. 51–66). SIL International Publications in Linguistics. The Summer Institute of Linguistics and the University of Texas at Arlington. Retrieved from <http://www.mexico.sil.org/resources/archives/8907>
- Van Dijk, T. A. (1976). Narrative macro-structures. *PTL: A journal for descriptive poetics and theory of literature, 1*, 547–568. Retrieved from <http://www.discourses.org/OldArticles/Narrative%20macrostructures.pdf>
- Van Dijk, T. A. (1977). *Text and context: Explorations in the semantics and pragmatics of discourse*. Longman linguistics Library. London: Longman.
- Van Dijk, T. A. (1980). *Macrostructures: An interdisciplinary study of global structures in discourse, interaction, and cognition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Van Dijk, T. A. (1997). The study of discourse. In T. A. Van Dijk (Ed.), *Discourse as structure and process* (Vol. 1, pp. 1–34). Sage Publications.
- Van Dijk, T. A. & Kintsch, W. (1983). *Strategies of discourse comprehension*. New York, NY: Academic Press.
- Van Leer, E. & Turkstra, L. (1999). The effect of elicitation task on discourse coherence and cohesion in adolescents with brain injury. *Journal of Communication Disorders, 32*(5), 327–349. doi:10.1016/S0021-9924(99)00008-8
- Vermeulen, J., Bastiaanse, R., & Van Wageningen, B. (1989). Spontaneous speech in aphasia: A correlational study. *Brain and language, 36*(2), 252–274. doi:10.1016/0093-934X(89)90064-3
- von Heusinger, Klaus. (1999). *Intonation and information structure. The representation of focus in phonology and semantics* (Habilitation thesis, Universität Konstanz, Konstanz, Germany).
- Wagenaar, E., Snow, C., & Prins, R. (1975). Spontaneous speech of aphasic patients: A psycholinguistic analysis. *Brain and language, 2*, 281–303. doi:10.1016/S0093-934X(75)80071-X
- Walker, M. A. (1996). Limited attention and discourse structure. *Computational Linguistics, 22*(2), 255–264. Retrieved from <https://users.soe.ucsc.edu/~maw/papers/walker-clsquib96.ps>
- Walker, M. A. (1998). Centering, anaphora resolution, and discourse structure. In M. A. Walker, A. K. Joshi, & E. F. Prince (Eds.), *Centering in discourse* (pp. 401–435). Oxford: Clarendon Press. Retrieved from <https://users.soe.ucsc.edu/~maw/papers/cent-alg13.pdf>
- Wang, W. S. & Sun, C. (2015). *The Oxford handbook of Chinese linguistics*. New York: Oxford University Press.
- 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. doi:10.1177/154596802401105199
- Wertz, R. T., Collins, M. J., Weiss, D., Kurtzke, J. F., Friden, T., Brookshire, R. H., . . . Porch, B. E., et al. (1981). Veterans Administration cooperative study on aphasia: A comparison of individual and group treatment. *Journal of Speech, Language, and Hearing Research, 24*(4), 580–594. doi:10.1044/jshr.2404.580
- Wertz, R. T., Deal, J. L., & Robinson, A. J. (1984). Classifying the aphasias: a comparison of the Boston Diagnostic Aphasia Examination and the Western Aphasia Battery. In *Proceedings of the 14th Clinical Aphasiology Conference* (pp. 40–47). BRK Publishers. Seabrook Island, SC. Retrieved from <http://aphasiology.pitt.edu/archive/00000794/>
- Westerhout, E. & Monachesi, P. (2005). Corpus of Dutch Aphasic Speech. In *Proceedings of the 16th Meeting of Computational Linguistics in the Netherlands* (pp. 83–98). University of Amsterdam. Amsterdam. Retrieved from <https://staff.fnwi.uva.nl/k.simaan/D-Papers/clinproc.pdf>
- Westerhout, E. & Monachesi, P. (2006). A pilot study for a Corpus of Dutch Aphasic Speech (CoDAS). In *Proceedings of the 5th International Conference on Language Resources and Evaluation* (pp. 1648–1653). European Language Resources Association (ELRA). Retrieved from [http://www.lrec-conf.org/proceedings/lrec2006/pdf/672\\_pdf.pdf](http://www.lrec-conf.org/proceedings/lrec2006/pdf/672_pdf.pdf)
- Wilkinson, R. (Ed.). (2015). Aphasiology. *Conversation and aphasia: Advances in analysis and intervention, 29.3*. doi:10.1080/02687038.2014.974138
- Wilkinson, R., Beeke, S., & Maxim, J. (2010). Formulating actions and events with limited linguistic resources: Enactment and iconicity in agrammatic aphasic talk. *Research on language and social interaction, 43*(1), 57–84. doi:10.1080/08351810903471506

- Wolf, F. & Gibson, E. (2005). Representing discourse coherence: A corpus-based study. *Computational linguistics*, 31(2), 249–287. doi:10.1162/0891201054223977
- Wolf, F., Gibson, E., Fisher, A., & Knight, M. (2005). The Discourse GraphBank (LDC2005T08). Web Download. Linguistic Data Consortium. Philadelphia, PA.
- Wright, H. H. (2011). Discourse in aphasia: An introduction to current research and future directions. *Aphasiology*, 25(11), 1283–1285. doi:10.1080/02687038.2011.613452
- Wright, H. H. & Capilouto, G. J. (2009). Manipulating task instructions to change narrative discourse performance. *Aphasiology*, 23(10), 1295–1308. doi:10.1080/02687030902826844
- 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., Capilouto, G. J., & Koutsoftas, A. (2013). Evaluating measures of global coherence ability in stories in adults. *International journal of language & communication disorders*, 48(3), 249–256. doi:10.1111/1460-6984.12000
- Wright, H. H., Capilouto, G. J., Wagovich, S., Cranfill, T., & Davis, J. (2005). Development and reliability of a quantitative measure of adults' narratives. *Aphasiology*, 19(3-5), 263–273. doi:10.1080/02687030444000732
- Wright, H. H., Koutsoftas, A. D., Capilouto, G. J., & Fergadiotis, G. (2014). Global coherence in younger and older adults: Influence of cognitive processes and discourse type. *Aging, Neuropsychology, and Cognition*, 21(2), 174–196. doi:10.1080/13825585.2013.794894
- Wright, H. H., Koutsoftas, A., Fergadiotis, G., & Capilouto, G. J. (2010). Coherence in stories told by adults with aphasia. *Procedia - Social and Behavioral Sciences*, 6, 111–112. doi:10.1016/j.sbspro.2010.08.056
- Wulfeck, B., Bates, E., Juarez, L., Opie, M., Friederici, A., Macwhinney, B., & Zurif, E. (1989). Pragmatics in aphasia: Crosslinguistic evidence. *Language and Speech*, 32(4), 315–336. doi:10.1177/002383098903200402
- Yiu, E. M.-L. (1992). Linguistic assessment of Chinese-speaking aphasics: Development of a Cantonese aphasia battery. *Journal of Neurolinguistics*, 7(4), 379–424. doi:10.1016/0911-6044(92)90025-R
- Yorkston, K. M. & Beukelman, D. R. (1980). An analysis of connected speech samples of aphasic and normal speakers. *Journal of speech and hearing disorders*, 45(1), 27–36. doi:10.1044/jshd.4501.27



# Appendices

## Appendix A

## Examples of RST relations based on the classification suggested by Carlson and Marcu (2001) (Chapter 3)

Class	Relation	Example		
		Satellite	Nucleus	Satellite
1. Attribution			[嗷][Mum told him]	[把遮啦][to bring an umbrella]
2. Background	a. background	[有一日朝早][One day in the morning]	[咪叫小朋友遮] [mum asked the child to bring an umbrella]	
	b. circumstance	[行到半路][Walking halfway]	[突然落起大雨][it started to rain suddenly]	
3. Condition		[如果你想好食][If you want it to taste better]	[最好搽牛油][better put on butter]	
4. Contrast			[叫遮] [唔][Mum told him to bring an umbrella] [he refused to bring]	
5. Elaboration	a. set-member		[火腿食法][There are two ways to prepare the ham]	[一切][一成放落去包][to julienne it] [or to put the whole piece on the bread]
	b. process-step		[整蛋三文治][To make an egg and ham sandwich]	[首先要煎熟蛋同火腿] [跟著面包][first pan-fry an egg and ham][then put (them) on the bread]
6. Evaluation	a. interpretation		[小朋友我唔遮][The child said I won't bring an umbrella]	[始都唔遮][At last he didn't bring the umbrella]
	b. comment		[整火腿蛋治][Making an egg and ham sandwich]	[哩反而熟悉][in fact this is more familiar (to me)]
7. Explanation	a. evidence		[天好][The weather is nice]	[唔落雨][it won't rain]
	b. reason		[我唔遮啦][I won't bring an umbrella]	[天好][the weather is nice]
8. Joint	a. joint		[小朋友翻把遮] [返][The child took the umbrella] [(and) continued to walk to school]	
	b. list		[第一材料一蛋] [仲要一火腿][The first ingredient is an egg] [and (we) also need a piece of ham]	
9. Manner-means	a. manner	[好嬲][Mum very angry] (Angry is a stative verb)	[小朋友][scolded the child]	

	b. mean	[用筷子][Use chopsticks]	[打蛋][(to) whisk the egg]	
10. Relations of cause	a. consequence		[真落雨][It really rained]	[小朋友晒][the child was drenched]
	b. purpose		[走翻返去][Went back]	[翻把遮][(to) take an umbrella]
11. Summary	a. summary			(作腿蛋治的步) [哩就腿蛋治](Procedures in making sandwich) [This is an egg and ham sandwich]
	b. restatement		[大都得][Bigger is ok]	[大面包都可以][it is fine to be bigger than a piece of bread]
12. Temporal	a. after	[煎完之後][After pan-frying it]	[就放落包度][put it on the bread]	
	b. same-time		[一路行] [一路真落雨][walk and walk] [it really rained at the same time]	
	c. sequence		[首先煎蛋] [跟著煎火] [最後放包][First pan-fry the egg] [then pan-fry the ham] [lastly put them on the bread]	
13. Reformulation	a. correction		[翻件雨][Take the raincoat]	[翻把雨][Take the umbrella]
	b. false-start	[男仔...][The boy is...]	[就落雨啦][It rains]	
	c. retracing		[咪到] [有少少嬲][Mum saw (it)] [(and) was a bit angry]	[到晒][saw him drenched]
14. Others	a. same-unit		[一女人] [] [嗌][A woman [(who) should be his mother] (embedded unit) [told him]	
	b. rhetorical-question		[唔我啦][(He) did not listen to what I said]	[解唔遮][Why didn't (he) bring an umbrella?]

## Appendix B

### Essential and elaborative components in the discourse samples (Chapter 3)

Discourse component	Information	% of occurrence in the transcripts of the aphasic group
Story-telling of “refused umbrella”		
Orientation	a. Mum gave the kid an umbrella	93.00%
	b. The child refused	93.00%
Complicating action	c. The child went out	60.00%
	d. It started to rain	100.00%
	e. The child went home	33.00%
	f. The child got drenched	53.00%
[Evaluation]	g. Mum was angry	13.00%
	h. Mum gave the child the umbrella	20.00%
Resolution	i. The child brought the umbrella	100.00%
	j. The child went out	33.00%
[Coda]	k. That’s the end of the story	0.00%
Procedural discourse of “Egg and ham sandwich”		
Procedural steps	[a. Beat an egg]	40.00%
	[b. Whisk the egg]	27.00%
	c. Pan-fry the egg	93.00%
	d. Pan-fry the ham	93.00%
	e. Put on a piece of bread	100.00%

*Note: Elaborative components (based on a 70% cut-off criterion using the NBDs’ transcripts) are added in square parentheses.*

## Appendix C

### Individual characteristics of aphasic participants (Chapter 4)

Pp	Sex	Age	Education	Aphasia type	Severity	Time PO	Etiology	Lesion location	Hemiplegia
1	F	51	secondary	dynamic, complex motor	severe	5.5 y	CVA	l MCA	right-sided pyramidal insufficiency
2	F	47	unfinished	dynamic, complex motor	moderate	2 y / 3 m	CVA	lMCA	right-sided hemiparesis, mostly in the arm
3	M	56	secondary	dynamic, complex motor	moderate	1 y / 2 m	CVA	lMCA	signs of right-sided pyramidal insufficiency
4	M	73	higher	Sensory	severe	1 y / 4 m	CVA	lMCA	signs of right-sided pyramidal insufficiency
5	M	50	secondary	complex motor	mild	7 y / 9 m	CVA	lMCA	spastic hemiplegia
6	M	70	unknown	sensory, acoustic-mnestic	moderate	4-5 m	CVA	unknown	unknown
7	M	52	higher	complex motor	moderate	3 m	CVA	lMCA	right arm and leg hemiparesis
8	F	40	secondary, vocational	sensory, acoustic-mnestic	moderate	3 m	CVA	lMCA	right-sided pyramidal insufficiency
9	F	67	higher	sensory, acoustic-mnestic	moderate	2.5 y	CVA	lMCA	signs of right-sided pyramidal insufficiency
10	M	58	secondary	sensory	moderate	1 y / 9 m	iCVA	lMCA	right-sided pyramidal insufficiency

*Note: CVA == cerebrovascular accident, iCVA == ischaemic CVA, lMCA == left middle cerebral artery*



## **Eidesstattliche Erklärung**

Hiermit erkläre ich an Eides statt, dass ich die vorgelegte Dissertation selbstständig und ohne unzulässige fremde Hilfe angefertigt und verfasst habe, dass alle Hilfsmittel und sonstigen Hilfen angegeben und dass alle Stellen, die ich wörtlich oder dem Sinne nach aus anderen Veröffentlichungen entnommen habe, kenntlich gemacht worden sind. Ich versichere weiterhin, dass die Dissertation in der vorgelegten oder einer ähnlichen Fassung bei keiner anderen Fakultät eingereicht worden ist.

Berlin, 1. September 2015

