

The importance of thematic informativeness in narrative discourse recovery in acute post-stroke aphasia

Amélie Brisebois, M.Sc.^{a,b}, Simona Maria Brambati, Ph.D.^{c,d}, Marianne Désilets-Barnabé B.Sc.^{a,b}, Johémie Boucher B.Sc.^{c,d}, Alberto Osa García, M.Sc.^{a,b}, Elizabeth Rochon, Ph.D.^{e,f,g,h}, Carol Leonard, Ph.D.ⁱ, Alex Desautels, M.D., Ph.D.^{a,j}, and Karine Marcotte, Ph.D.^{a,b}

^a *Centre de recherche du Centre intégré universitaire de santé et de services sociaux du Nord-de-l'Île-de-Montréal, Montréal, Québec, Canada.*

^b *École d'orthophonie et d'audiologie, Faculté de médecine, Université de Montréal, Montréal, Québec, Canada.*

^c *Centre de recherche de l'Institut Universitaire de Gériatrie de Montréal, Montréal, Québec, Canada.*

^d *Département de psychologie, Faculté des arts et des sciences, Université de Montréal, Montréal, Québec, Canada*

^e *Department of Speech-Language Pathology, Faculty of Medicine, University of Toronto, ^[L]^[L]Toronto, Canada*

^f *Toronto Rehabilitation Institute, Toronto, Canada*

^g *Heart and Stroke Foundation, Canadian Partnership for Stroke Recovery, Ontario, Canada*

^h *Rehabilitation Sciences Institute, University of Toronto*

ⁱ *School of Rehabilitation Sciences, University of Ottawa*

^j *Département des neurosciences, Faculté de médecine, Université de Montréal, Montréal, Québec, Canada.*

Corresponding author:

Karine Marcotte, Ph.D.

NARRATIVE DISCOURSE RECOVERY
Aphasiology (accepted)

Address: Centre de recherche du Centre intégré universitaire de santé et de services sociaux du
Nord-de-l'Île-de-Montréal, 5400 Gouin Ouest, Montréal, Québec, Canada, H4J 1C5

Phone number: 514-338-2222 extension 7710

Fax number: 514-340-2115

E-mail: karine.marcotte@umontreal.ca

The importance of thematic informativeness in narrative discourse recovery in acute post-stroke aphasia

Abstract

Background: Discourse analysis allows the examination of functional and ecological language impairment in post-stroke aphasia. Given its complexity, various methods of analysis have been developed to measure the multiple components of discourse. Clinical assessment usually includes discourse analysis, but how clinicians should assess recovery of discourse, particularly in acute care settings, is still a matter of debate.

Aims: This study aimed to measure improvements in discourse production in early post-stroke aphasia recovery.

Methods & Procedure: Twenty-three persons with aphasia following a first left middle cerebral artery stroke were recruited in the stroke unit of Hôpital du Sacré-Coeur de Montréal (May 2015-July 2018). Patients treated with thrombolysis (n=10) and untreated patients underwent two aphasia assessments (0 to 72 hours, 7 to 14 days post-onset). Discourse assessment consisted of the picture description task from the Western Aphasia Battery (Kertesz, 2006). Changes in microlinguistic and thematic informativeness measures between the two assessment periods were analyzed.

Outcomes & Results: In-depth microlinguistic analyses showed no significant difference between the initial and follow-up assessments. Conversely, some thematic informativeness variables improved significantly during the same time period. Patients treated with thrombolysis produced more thematic units than untreated patients at both assessments, but the change between the two groups was not significant.

Conclusions & Implications: This study suggests that thematic informativeness variables are sensitive to language improvement in early post-stroke aphasia recovery whereas no microlinguistic variables improved significantly in the same period. In contrast to previous evidence, the difference between patients treated with thrombolysis and untreated patients was not evident over time. The results suggest that thematic informativeness constitutes an interesting path to explore as

a routine clinical assessment in acute-care settings because it is time-efficient, simple to conduct and reliable in assessing early changes in the discourse production of individuals with aphasia in the acute post-stroke phase.

Keywords: aphasia, discourse, language recovery, acute stroke, informativeness

Introduction

Discourse assessment allows a detailed and relatively ecological assessment of language impairments in post-stroke aphasia. In fact, most clinicians who work with this population collect discourse samples. A minimum of 60% of speech-language pathologists working with people with aphasia (PWA) analyze discourse samples at least some of the time in their language assessment (Bryant, Spencer, & Ferguson, 2017). Discourse productions vary in important ways depending on the nature of the elicitation task (Pritchard, Hilari, Cocks, & Dipper, 2017). Topic-directed interview does not provide obligatory contexts for production and offers a high ecological validity (Bryant et al., 2017). Conversely, story retell and picture description generate more constrained productions, are more predictable and thus allow an easier comparison over time and between individuals (Bryant, Ferguson, & Spencer, 2016; Linnik, Bastiaanse, & Höhle, 2016). The constrained methods are thus more frequently used in both research and clinical practice. Widely used language batteries (e.g. Western Aphasia Battery (Kertesz, 2006)) offer qualitative grids to score picture description, but normative data are not available in Canadian-French. Only the Montreal-Toulouse Language Battery (Nespoulous et al., 1992) offers a content information grid to guide scoring the picture description task that has been validated in Canadian-French (Béland, Lecours, Giroux, & Bois, 1993). In current practices in Quebec, the analysis of a description task remains mostly based on clinical judgement (Monetta, 2014), providing speech-language pathologists with few guidelines to determine the success of the task

and even less to appreciate changes over time.

Although discourse assessment is an essential part of the clinical evaluation, little is known about how discourse impairment evolves over time (Bryant, Ferguson, & Spencer, 2016), particularly in the acute stage. Following natural recovery of language, discourse abilities recover spontaneously in the first weeks following a stroke, but this might be due to overall language recovery. Contemporary debates in aphasiology (Dietz & Boyle, 2018) confirm a clear need for researchers to develop a measure that captures meaningful outcomes, in order to target improvement seen in treatment and establish a better understanding of the recovery process. To date, it is still unclear if the acute assessment of discourse could have prognostic value for long-term aphasia outcomes.

Analysis of discourse samples can inform two main dimensions of language: the microlinguistic dimension which includes within-sentence linguistic measures, and the macrolinguistic dimension which refers to between-sentence measures (Marini et al., 2008). Microlinguistic abilities involve phonological, lexical, syntactic and semantic variables; whereas macrolinguistic abilities refer to discourse organization. Both of these levels of analysis inform how much information a speaker can convey (e.g.: Armstrong, 2000). A brief overview of discourse variables that have been studied in the field, with a special interest on acute aphasia discourse, will be presented.

Historically, microlinguistic analysis of speech samples has been able to inform how PWA's language abilities differ from those of healthy individuals and much important work has been done in this area (e.g. reviews such as Armstrong, 2000; Bryant, Ferguson, & Spencer, 2016; Pritchard, Hilari, Cocks, & Dipper, 2017). Some variables seem to have a higher diagnostic sensitivity for patients with language impairment, such as speech rate, percentage of semantic and phonological errors, Moving Average Token-Type Ratio, and number of verbs per utterance (Andreetta,

Cantagallo, & Marini, 2012; Andreetta & Marini, 2015; Covington & McFall, 2010; Saffran, Berndt, & Schwartz, 1989; Wright, Capilouto, Wagovich, Cranfill, & Davis, 2005). Research using these variables has indeed highlighted differences between healthy individuals and PWA. Refined linguistic analyses are critical to qualify the language performance of PWA; this is especially true when language impairment is mild (Dalton & Richardson, 2015; Fromm et al., 2017). Limited by lack of time, discourse analysis in clinical settings has been mostly based on a qualitative rather than a quantitative assessment. Many questions have arisen concerning the psychometric properties of these variables (Dietz & Boyle, 2018; Kintz & Wright, 2018): In the last decade, the desire to capture meaningful outcomes has led to an increasing number of metrics in discourse analysis that often lack reliability, validity and stability of data. With more than 500 measures identified in discourse analysis (Bryant et al., 2016), the gap between research and practice in discourse analysis has broadened. Refined phonological, syntactic and semantic analyses show evident impairments in early aphasia assessments, but these measures do not seem to evolve as much as expected in the first months of recovery, even when language therapy is provided (Carlomagno, Pandolfi, Labruna, Colombo, & Razzano, 2001; Larfeuil & Le Dorze, 1997; Marini, Caltagirone, Pasqualetti, & Carlomagno, 2007).

Although the analysis of microlinguistic skills is important, informativeness is also crucial as it represents the ability of the patient to convey a message and thus captures a more complete picture of discursive abilities. The importance of informativeness variables is reflected by a large body of literature on the topic (as reviewed by Pritchard et al., 2017). Also, significant work has been done on informativeness at the lexical level. Among the measures developed to investigate the use of language, lexical informativeness has brought a functional scope to

microlinguistic analysis. This is a strong measure as it encompasses both the fundamental underpinnings of lexical processing and the functional aspect of language use, both involved in the communication process (Sherratt, 2007). In the early 90's, Nicholas and Brookshire (1995) introduced the "Correct Information Unit" (CIU) measure which can be applied to a variety of speech samples. CIU is a standardized rule-based scoring system that was developed to assess informativeness and efficiency of connected speech in PWA in comparison to healthy participants. Indeed, CIUs have been shown to have great diagnostic sensitivity and their validity has been supported by high inter-rater reliability (Craig et al., 1993). More recently, Andretta and Marini (2015) developed "Lexical Information Units", which are well formed words, grammatically and pragmatically accurate. Using this lexical informativeness variable, they showed that people with fluent aphasia produced less informative discourse than healthy controls, even after intensive therapy, showing once more the diagnostic sensitivity of informativeness and the critical importance of measuring informativeness in connected speech.

However, two main methodological barriers have been identified for the use of informativeness as standard analysis in clinical settings, at least in Quebec. Analyses are often based on subjective impressions due to the lack of standard measures for use with the Canadian-French population (Monetta, 2014). Second, informativeness assessment generally requires a detailed linguistic analysis, which is often impossible considering the lack of time (Bryant et al., 2017). The study of lexical information units is time consuming; thus, for ecological purposes, this has led researchers interested in language changes in acute aphasia to develop other instruments that use a broader scope to analyze discourse and explore informativeness at a macrolinguistic level. For instance, the Main Concept Analysis assesses presence, completeness and accuracy of the main

information of PWA's connected speech (Nicholas & Brookshire, 1995). These global measures can be classified as macrolinguistic variables as described by Armstrong (2000).

A growing body of literature focused on macrolinguistic abilities in the last few decades, has identified thematic (or global) informativeness as amongst the most promising measures, as it represents a fundamental aspect of successful communication. For the purpose of this paper, thematic informativeness will be defined as semantic elements (information, main events/ideas/concepts) conveyed that are relevant to the topic of a stimulus. There remains, however some controversy as to the promising value of thematic informativeness. Some studies do not report impairments of thematic informativeness in PWA (Albright & Purves, 2008; Doyle, Goda, & Spencer, 1995; Doyle et al., 2000); however many others do report differences in thematic informativeness between healthy individuals and PWA (Agis et al., 2016; Capilouto, Wright, & Wagovich, 2006; Nicholas & Brookshire, 1995; Ulatowska, Freedman-Stern, Doyel, Macaluso-Haynes, & North, 1983). Also, some report impairments with severe aphasia (Ulatowska et al., 1983); whereas no impairment has been identified in people with mild and moderate aphasia (Yorkston & Beukelman, 1980).

Yorkston and Beukelman (1980) were among the first to propose a measure of informativeness. "Content units" (CU) have been defined as relevant pieces of information mentioned by healthy speakers. They also can be described as general or thematic informativeness components. Recently, Agis et al. (2016) showed that the measure of CUs in the description of the Cookie Theft picture from the BDAE (Goodglass et al., 2001) provides valuable information about volume and lesion location in acute stroke. The use of CUs rather than the standard National Institute of Health Stroke Scale (NIHSS) scoring increased the NIHSS sensitivity without requiring much

more assessment time. More precisely, they (Agis et al., 2016) calculated total CU, syllables per CU, ratio of left:right CU, CU/minute and percent interpretive CU from the Cookie Theft discourse description obtained within 48 hours of stroke onset in patients with left and right hemisphere acute stroke. These informativeness measures were able to differentiate controls from brain-injured patients. Also, the left:right CU ratios (i.e., CU visually located on the left side of the Cookie Theft picture compared to those located on the right side of the picture) were able to differentiate patients with left-hemisphere damage from right-hemisphere damage. These variables increased NIHSS sensitivity to communication impairments following right-hemisphere lesions by capturing other cognitive functions such as neglect, attention and topic maintenance. Finally, they showed that lesion volume of the stroke was best explained by a model including CU measures for right and left hemisphere stroke. This study demonstrates that a brief picture description analysis, in other words variables including CUs, complement the original NIHSS score and supports prediction of brain damage volume and location. This close relationship between informativeness and neural damage reinforces the special interest accorded to thematic informativeness in acute stroke.

More recently, Furlanis et al. (2018) investigated the early recovery of the discourse abilities of 40 Italian speaking patients with ischemic stroke, 18 of whom were treated with thrombolysis. They used a rapid screening test, which included a discursive task. They investigated the impact of thrombolysis on language recovery in acute stroke. Briefly, intravenous thrombolysis is a treatment currently in use to facilitate reperfusion in patients with acute ischemic stroke. It has been demonstrated to significantly improve short and long term general outcome (e.g. Emberson et al., 2014; Rohde, Worrall, O. Halloran, Godecke, & Farrell, 2017; Röther et al., 2002; Seitz & Donnan, 2015; Seitz, Sukiennik, & Siebler, 2012; Wardlaw, Murray, Berge, & del

Zoppo, 2014). In their study, Furlanis et al. (2018) recruited forty Italian patients who underwent a language screening test at baseline, 24 hours and 72 hours post-stroke. They showed that “spontaneous speech” was the only variable that significantly improved, amongst other language variables, between baseline and 72 hours post-stroke and between 24 and 72-hours post-stroke. “Spontaneous speech” was calculated in terms of the number of semantic units named during the 1-minute description of the Cookie Theft picture of the Boston Diagnostic Aphasia Examination (Goodglass et al., 2001). These information units (n=16) were previously identified (Allibrio, Gori, Signorini, & Luzzatti, 2009) and participants were given 1 point for each named element. The scoring of the ‘spontaneous speech’ measure also included a maximum of 3 additional points for relevant elements that were named, but not listed among the 16 elements. As well, a maximum of 2 points was deducted each for syntactic and phonemic errors. The last two scores reflect a general impression of language impairment (see Appendix B). Because of this general scope, the ‘spontaneous speech’ measure of Furlanis et al. (2018) may be considered similar to the Content Units identified by Yorkston and Beukelman (1980). However, the “Spontaneous speech” score was mainly composed of thematic informativeness (i.e. macrolinguistic measure), though not solely as it includes a general appreciation of microlinguistic features. An overall impression of impairment of phonology and syntax was given by the assessor, which is different from microlinguistic analyses (Allibrio et al., 2009) and a maximum of two points for each domain could be subtracted from the thematic unit score. Interestingly, Furlanis et al’s (2018) findings provide new insights regarding informativeness since they demonstrated that ‘spontaneous speech’ was related to the first 72 hours of the recovery process in stroke.

Despite an increasing number of studies on discourse abilities in PWA, little is known about how discourse evolves in the early phase (i.e. ten days) of post-stroke aphasia and less is known about patients treated with thrombolysis as compared to untreated patients. The aim of the present study is thus to measure discourse production changes in early post-stroke aphasia recovery by analyzing microlinguistic variables and thematic informativeness. Microlinguistic variables were chosen according to recent findings in aphasia, based on the criteria of sensitivity to language impairment (Andretta, Cantagallo, & Marini, 2012; Andretta & Marini, 2015; Fergadiotis & Wright, 2016; Fromm et al., 2016). It is hypothesized that thematic informativeness will provide a sensitive measure of discourse recovery in the early phase, based on the results of Furlanis et al. (2018). It is also hypothesized that patients treated with thrombolysis will have greater improvement than untreated patients.

Material and Methods

Participants

This study was part of a larger project directed by S.M.B. and K.M. which sought to investigate the longitudinal recovery of white matter structural connections that mediate early and late post-stroke aphasia recovery. Twenty-three native Canadian-French speakers with various forms of post-stroke aphasia, nine women; mean age: 70.2 ± 12.6 years; mean education: 11.7 ± 4.0 years, were recruited in the present study. Two were monolinguals (Canadian-French only), seventeen were bilinguals (Canadian-French and another language) and four spoke three languages or more. However, they all used Canadian-French as their first language. Among these patients, those who were eligible for thrombolysis received this treatment as recommended by the Canadian Stroke Guidelines (Boulanger et al., 2018). The decision to give this treatment was made by the

on-call neurologist and was made before the study was presented to the participant. Ten patients were treated with thrombolysis. Individual clinical and sociodemographic information of all participants is presented in Table 1. Participants were recruited from the stroke unit at Hôpital du Sacré-Coeur de Montréal (Centre intégré universitaire de santé et de services sociaux du Nord-de-l'île-de-Montréal, Québec) between May 2015 and July 2018. All participants were diagnosed with aphasia following an ischemic stroke by a neurologist and screened for eligibility by a member of the research team.

Table 1. Demographic and clinical variables of all participants.

Partici pant	Sex	Age	Educ.	T1 (number of days after onset)	T2 (number of days after onset)	Initial NIHSS score	Initial Aphasia type	Initial Severity (BDAE Scale¹)	CS at T1	CS at T2
Thrombolysis group (received rTPA)										
1	M	52	9	1	7	n/a	Transcortical mixed	Moderate to severe	16.57	8.20
2	M	74	6	3	8	9	Wernicke	Severe	3.56	10.24
12	M	65	11	3	14	6	Anomic	Mild	0.35	28.53
13	M	72	15	1	9	11	Transcortical mixed	Moderate to severe	6.78	21.33

¹ The BDAE Severity was applied to the CS score, as the BDAE language battery is not standardized in French.

NARRATIVE DISCOURSE RECOVERY
Aphasiology (accepted)

14	M	55	11	4	10	23	Global	Moderate to severe	9.90	10.63
15	M	73	11	1	8	n/a	Wernicke	Moderate to severe	14.79	14.79
16	M	64	15	1	11	n/a	Conduction	Mild	28.90	27.46
18	F	60	12	3	13	7	Anomic	Mild to moderate	21.73	23.60
22	F	81	15	2	11	17	Anomic	Mild	27.70	26.79
23	F	68	12	3	12	n/a	Anomic	Mild	28.78	26.74
r-TPA										
group		66.22	11.67	2.11						19.62
Mean		(9.6)	(3.04)	(1.17)	10.11(2.37)	12.17(6.59)			3.11 (5.29)	(8.22)
(SD)										
Non-treated group (no rTPA)										

NARRATIVE DISCOURSE RECOVERY
 Aphasiology (accepted)

3	M	78	10	3	10	n/a	Global	Severe	1.69	0.00
4	F	75	9	2	9	23	Global	Severe	0.00	0.00
5	M	61	10	3	10	6	Broca	Moderate to severe	4.02	11.32
6	M	49	9	2	11	6	Anomic	Mild to moderate	2.62	24.82
7	M	73	19	3	10	18	Wernicke	Severe	6.32	7.71
8	F	70	14	3	12	16	Global	Severe	1.69	1.87
9	M	83	9	3	10	9	Transcortical sensory	Moderate	10.49	3.90
10	F	47	18	0	10	26	Global	Severe	0.00	0.00
11	F	73	7	3	13	n/a	Transcortical sensory	Moderate	2.86	14.36

NARRATIVE DISCOURSE RECOVERY
Aphasiology (accepted)

17	F	95	6	2	9	1	Broca	Mild to moderate	6.59	16.27
19	M	91	19	3	15	7	Anomic	Mild to moderate	4.99	20.09
20	F	85	16	2	8	n/a	Transcortical mixed	Moderate	10.08	12.30
21	M	71	7	2	9	n/a	Transcortical motor	Moderate	3.24	18.59

No-										
rTPA		72.79	11.79							10.15
group		(13.91)	(4.59)	2.43(0.85)	10.57(1.87)	12.44(8.62)			3.94 (3.71)	(8.44)
Mean										
(SD)										

NIHSS= National Institute of Health Stroke Scale, rTPA= Recombinant tissue plasmin activator/thrombolysis; n/a= non-available in the medical chart; BDAE= Boston Diagnostic Aphasia Examination; CS= Composite score of language

Participants presented with language impairment due to a first stroke affecting the territory of the left middle cerebral artery. No criteria concerning aphasia severity in the acute phase or lesion size were applied. Only patients who had been fully fluent speakers of Canadian-French before the stroke were included. Patients with a history of major psychiatric disorders, learning disabilities or severe perceptual deficits were excluded. Left-handed patients and patients with additional neurological diagnoses were also excluded from the study. None of the patients presented with aphasia due to sub-thalamic stroke or a pronounced subcortical arteriosclerosis. The study was approved by the ethics review board of the Centre intégré universitaire de santé et de services sociaux du Nord-de-l'Île-de Montréal (Project #MP-32-2018-1478) and written informed consent was obtained from all participants.

Assessments

Initial assessment (T1) occurred within the first 72 hours, mean = 2.3 days; SD= 1.0, and the follow-up assessment (T2) took place approximately ten days, mean= 10.3 days, SD= 2.0, post-stroke. Specific timing for each assessment is specified in Table 1. As part of the larger project, all participants underwent a complete language assessment. Therefore, language impairments were carefully documented at T1, T2 and at a six-month follow up. Based on Lazar et al. (2010) a composite score was developed and adapted for the Canadian-French speaking population. Although the WAB is one of the most widely used batteries, previous studies have shown that aphasia classification by the WAB matched the clinical impression of only half of the patients in acute assessment (Swindell, Holland, & Fromm, 1984), which puts into question the sensitivity of the tasks included in this composite score. Consequently, sections of

protocols that are widely used with Canadian-French speaking patients were used. This Composite Score (CS) consisted of three subscores: Comprehension, Repetition and Naming. For the Comprehension subscore, it combined the Word-Sentence Comprehension Task (max = 47 points) of the Montreal-Toulouse (Nespoulous, Lecours, & Lafond, 1986) and the revised (short) version of the Token Test (De Renzi & Faglioni, 1978) (max = 36 points), which includes oral comprehension of words, sentences and sequential commands. Repetition was assessed with the repetition task of the Montreal-Toulouse (Nespoulous et al., 1986): (2 points for each word/nonword (n=30) and 5 points for each sentence (n=3), max = 75 points). Finally, Naming subscore was composed of the DO-80 (Test de dénomination orale d'images: Picture naming test) (Deloche & Hannequin, 1997) (max = 80 points) and the semantic fluency task (max = 25 points) of the Protocole Montréal de la Communication (Joanette, Ska, & Côté, 2004). Each of the three modalities was computed to a possible score of 10, with a maximum CS = 30. Initial impairment ($CS_{initial}$) and sub-acute impairment ($CS_{10\text{days}}$) are reported in Table 1. A one-way ANOVA showed no significant difference between the group who received thrombolysis and the untreated group concerning $CS_{initial}$, $F(1,21) = 0.197$, $p=.662$, whereas a significant difference was observed for $CS_{10\text{days}}$ $F(1,21) = 7.288$, $p=.013$, with the participants who received thrombolysis showing better language performance. Severity scoring was based on results obtained on these tasks and rated on the BDAE severity scale (Goodglass et al., 2001).

Among the language tasks, PWA were asked to describe what was happening in the Picnic picture of the Western Aphasia Battery (Kertesz, 2006) and to produce sentences. They were encouraged to pay attention to all aspects of the picture. Picture description samples were filmed using Sony HDR-PJ540 camera (9.2 mega pixels). The picnic picture from the WAB (Kertesz, 2006) was part of the protocol in the hospital

where data were collected; Furlanis et al. (2018) used the Cookie Theft picture (Goodglass et al., 2001).

Microlinguistic discourse analysis

Videos of each discourse sample were imported and transcribed in ELAN (Sloetjes & Wittenburg, 2008) by an experienced speech-language pathologist (A.B.) and a student in speech-language pathology (M.D.-B.) using CHAT conventions (MacWhinney, 2000). Transcriptions underwent detailed linguistic and textual analysis focusing on microlinguistic measures known to be affected in aphasia (e.g. Andreetta, Cantagallo, & Marini, 2012). All microlinguistic variables were extracted using EVAL of CLAN (MacWhinney, 2000) for each sample, and a separate script on CLAN for the Moving Average Token-Type Ratio (MATTR) (Covington, 2007). MATTR allows comparisons of varying sample lengths and measures lexical diversity by calculating a token-type ratio for non-overlapping segments of a sample. Sample duration range, as calculated by the program EVAL of CLAN (Mac Whinney, 2000), was between 0 and 142 seconds (T1) for the first assessment and from 0 to 132 seconds for the second assessment (T2). Utterances segmentation, transcription and scoring for utterances and lexical errors were conducted following the CHAT manual (MacWhinney, 2000) with additional guidance for French users of this program (Colin & Le Meur, 2016). Productivity measures extracted were defined as the total number of completed words and number of words per minute. Grammatical/ syntactic complexity was measured by the number of verbs per utterance and density (Brown, Snodgrass, Kemper, Herman, & Covington, 2008). Also, percentages of phonological and semantic errors were computed. Phonological errors included phonemic and phonetic paraphasias, false starts, *conduites d'approche* and neologisms. Finally, the percentage of adequate utterances was calculated, based on the number of utterances produced without any errors. As in Andreetta and Marini

(2015), semantic errors included semantic paraphasias (i.e., replacement of a target word by a semantically related or unrelated word) and verbal paraphasias. Percentages were obtained by dividing the total number of errors by total number of content words and this obtained value was multiplied by 100.

Thematic information units (TUs)

TUs were selected based on the WAB picture description of forty-five healthy Canadian-French speakers that were recruited in other (as yet, unpublished) studies conducted by our research team. This group is composed of 15 males and 30 females, mean age: 70.96 ± 9.62 years, and mean education: 15.73 ± 4.68 years. Similar to Marini et al. (2011), sixteen TUs identified by at least 75% of these participants were included in the analysis grid and given a score of one point (see *Appendix A* for TU selection). This variable includes the total number of specific units produced by the participants. In other words, for each participant a maximum of 16 thematic units could be obtained, irrespective of the length/elaboration of the description. Efficiency was also considered by calculating the number of TUs per minute and the number of TUs per utterance. This measure of TUs differs from the “spontaneous speech” measure used by Furlanis et al. (2018), as it does not include extra points for other relevant named elements that were named, yet not listed among the 16 elements, nor does it include an overall impression of syntactic and phonological errors.

Global informativeness measure (GIM)

A measure similar to the ‘spontaneous speech’ measure used by Furlanis et al. (2018) was developed and called the General Informativeness Measure (GIM). One point was given for each of the 16 selected TUs and .5 extra points for each thematic unit produced that was relevant to the picture (no ceiling score was applied). A maximum of

2 points was subtracted for each of phonemic and syntactic errors. Efficiency calculations were also performed (GIM per minute and GIM per utterance). Appendix B describes the scoring rules used for the GIM in the present study.

TUs and GIMs are closely related and one might argue that TU performance could potentially drive the GIM performance. However, TU has a ceiling score of 16 and represents only the naming of the 16 predetermined themes. In contrast, the GIM does not have a ceiling, and also takes into account syntactic and phonological impairments. Therefore, the analysis of both variables is clinically relevant.

Inter-rater reliability

All transcriptions were made by the first author (A.B.). Inter-rater reliability (IRR) was conducted on the transcriptions of 10 randomly selected participants at both time points (n=20 transcriptions) by a second rater (M.D.-B.). For all measures (microlinguistic, TU and GIM), discrepancies between raters were resolved through a discussion between the two raters until a consensus was reached. Two-way random effects intraclass correlation coefficients (ICC) were calculated on all microlinguistic variables to determine consistency between raters (as in Marcotte et al. 2017).

The results of this analysis are reported in Table 2. Most variables met the threshold of high reliability $ICC > .80$ (Streiner & Norman, 2008). In particular, inter-rater reliability was high for the number of utterances, total words, MLU, MATTR, verbs/utterance, density and % adequate utterances at both times of assessment. Inter-rater reliability was also high for words/minute and % phonological errors (including phonemic and phonetic paraphasias, false starts, *conduites d'approche* and neologisms) at the first assessment, but not at the second assessment. Unexpectedly, inter-rater reliability for semantic paraphasias was poor at both timepoints.

[Table 2 should be inserted here]

The variables TU and GIM were independently scored by two of the authors (A.B. and M.D.-B.) for 100% of the discourse samples (at both time points). Inter-rater reliability was high for both variables (TU: ICC = .997; GIM: ICC = .993) (Streiner & Norman, 2008).

For all measures (microlinguistic, TU and GIM), discrepancies between raters were resolved through a discussion between the two raters until a consensus was reached.

Data analysis

All statistical analyses were done using SPSS® v25.0. Two-factor mixed-design ANOVAs were conducted with group (treated with thrombolysis and untreated) as the between-subjects factor and time (T1 and T2) as the within-subjects factor for each microlinguistic variable, TUs and GIM. The significance level was set at $p < .05$ after Bonferroni correction for multiple comparisons. Test-retest reliability was assessed using Pearson correlations between TU and GIM results at both times.

Results

Participants

Table 1 reports the mean scores of participant characteristics. A one-way ANOVA demonstrated no significant differences between the treated and the untreated groups concerning sex, $F(1,21) = 1.576, p=.199$, age, $F(1,21) = 1.527, p=.230$, education, $F(1,21) = 0.005, p=0.946$, timing of testing (T1 and T2), $F(1,21) = 0.571, p=.458$ and $F(1,21) = 0.270, p=.609$, initial NIHSS score, $F(1,21) = 0.004, p=.948$, and CS at T1

$F(1,21) = 0.197, p=.662$. A significant difference in language score between the treated group vs the untreated group was found at T2 only $F(1,21) = 7.288, p=.013$. As a group, patients treated with thrombolysis had higher language scores (i.e. milder aphasia) 7-10 days post-stroke than the untreated group.

Test-retest stability

Test-retest stability of GIM was calculated using Pearson correlations, as in Capilouto et al. (2006), on a subset of thirteen healthy participants who participated in the larger project. It was composed of 7 males and 6 females, mean age: 62.77 ± 6.43 years, and mean education: 14.83 ± 2.17 years. The picture description task of the Western Aphasia Battery (A. Kertesz, 2006) was used and followed the same method and procedure as that described in the present study. The mean time between the two testing points was 272.23 ± 71.22 days (as compared to 8.08 ± 1.83 days for PWA) A significant correlation between T1 and T2 measurements was found for GIM, $r = .775, p < .001$, indicating that GIM is stable. Test-retest stability for TUs was not calculated since this subset of participants was included in the initial selection of thematic units.

Microlinguistic variables

Table 3 reports the mean scores, collapsed across group, for each microlinguistic variable obtained at both T1 and T2, as well as the F value. None of the microlinguistic variables changed significantly.

[Table 3 should be inserted here]

Informativeness

Thematic Units

Table 4 shows the mean values, collapsed across group, of TUs, TUs per minute and TUs per utterance both at T1 and T2 as well as the F values. The results revealed a main effect of time, with significant improvements in thematic informativeness at T2 in terms of TUs, $F(1,21) = 7.731$; $p = 0.011$, showing an average increase of 2.04 TUs, and also in terms of TUs per minute, $F(1, 21) = 4.787$; $p = .040$, showing an average increase of 2.94 TUs/min. Overall, most patients, regardless of group, were able to express more thematic units at T2 compared to T1. However, change in terms of TUs per utterance, $F(1, 21) = 1.995$; $p = .173$ was not significant.

[Table 4 should be inserted here]

The results also revealed a significant main effect of group, $F(1,21) = 8.048$, $p = .010$, but there was no interaction between group and time. Namely, PWA treated with thrombolysis produced more TUs both at T1 and T2 than untreated PWA, and both groups showed a similar improvement between T1 and T2. Figure 1a) represents the number of TUs produced by each individual at both T1 and T2. When looking at the individual data, TU improved for 12 patients; other patients showed either stability (i.e. no change in the number of TUs between T1 and T2, $n=11$) or a slight deterioration (i.e. a decreased number of TUs produced at T2; $n=2$). GIM improved for 11 patients; other patients showed either stability (i.e. no change in GIM score between T1 and T2; $n=7$) or a slight deterioration (i.e. a decreased of GIM score at T2; $n=5$). GIM/utt improved for 11 patients; other patients showed either stability (i.e. no change in GIM/utt between

NARRATIVE DISCOURSE RECOVERY

Aphasiology (accepted)

T1 and T2; n=7) or a slight deterioration (i.e. a lower GIM/utt at T2; n=5). When all these results are considered together, 14/23 individuals showed an improvement as reflected in an increased number of TUs, a higher GIM score or better GIM/utt at T2.

[Figure 1 should be inserted here]

Figure 1 also shows that patients who received thrombolysis demonstrated overall better performance. Of all PWA treated with thrombolysis (10 patients), seven patients improved, two remained stable and one patient produced fewer TUs at T2. Among the untreated PWA, six out of 13 improved, five showed remained stable and one deteriorated.

General Informativeness Measure

Table 4 displays the mean values, collapsed across groups of GIM, GIM per minute and GIM per utterance both at 48-72 hours and 7-14 days post stroke. Improvement in thematic informativeness over time is represented by positive changes for all three variables at T2, but differences are only significant for GIM and GIM per minute. More precisely, an average increase of 2.83 GIM, $F(1, 21) = 6.393$; $p = 0.020$, was observed between T1 and T2. For GIM per minute, an average increase of 3.88 TUs per minute, $F(1,21) = 7.972$; $p = 0.010$, was observed between T1 and T2. Results for GIM per utterance did not reach statistical significance, $F(1,21) = 3.290$; $p = .084$. Figure 1b) represents the GIM for each individual at both T1 and T2. Regardless of the treatment group, individually, 12 PWA showed an improvement in GIM at T2, with the remaining patients demonstrating stability (n=6) or a slight deterioration (n=5).

A significant main effect of group was also found with the average GIM significantly higher in patients treated with thrombolysis than in the untreated patients, $F(1,21) = 8.502, p=.008$ (see Table 4). For the efficiency measures (GIM/minute and GIM/utterance), there were no significant differences between treated and untreated PWA. There was also no interaction between group and time. Figure 1c) represents the GIM per utterance for each individual at both T1 and T2 and indicates that patients treated with thrombolysis produced a higher value of GIM/utterance than untreated PWA.

Discussion

The present study reports improvements on thematic informativeness during the picnic description of the WAB in 15 of 23 individuals with acute post-stroke aphasia between 2 days and 10 days post-onset. Thematic informativeness improvements were observed by measuring the TUs, and also by using the GIM which combines the number of TUs and a global appreciation of phonological and syntactic impairments. Moreover, an improvement in the efficiency of thematic informativeness was found, as measured by the number of TUs and GIM per minute. No significant improvements were found on the microlinguistic measures known to reflect language impairment: total words, words/minute, verb/utterance, and density (e.g. Andreetta & Marini, 2015). Although informativeness has been shown to improve following specific therapies in case studies (Marini, Caltagirone, Pasqualetti, & Carlomagno, 2007) very few studies have revealed discourse changes in a group of patients with aphasia in the early phase of recovery. Consistent with that of others (Agis et al., 2016; Furlanis et al., 2018), the present findings support the critical importance of measuring thematic informativeness in connected speech of PWA in the acute stage of stroke.

Most in-depth studies on discourse analysis have examined improvements in discourse production in individuals with chronic post-stroke aphasia (Andreetta, Cantagallo, & Marini, 2012; Capilouto et al., 2006; Doyle et al., 1995; Marini et al., 2007; Nicholas & Brookshire, 1993). However, it is widely acknowledged that the greatest degree of language recovery takes place within the first two weeks after stroke (Kertesz & McCabe, 1977; Laska, Hellblom, Murray, Kahan, & Von Arbin, 2001; Pedersen et al., 1995; Wade, Hower, David, & Enderby, 1986). The present study shows that thematic informativeness reveals early discourse recovery better than the microlinguistic measures used -- supporting the need for further investigation of this measure in clinical settings. Others have clearly demonstrated the clinical sensitivity of microlinguistic variables (e.g. Andreetta et al., 2012; Nicholas & Brookshire, 1993) and discourse changes following therapy (Larfeuil & Le Dorze, 1997; Marini et al., 2007). However, despite high inter-rater reliability for most measures, in the present investigation no significant changes on microlinguistic variables in the acute phase when no therapy was provided were observed. One reason for this difference in findings might be explained by the composition of the sample. Our study included patients with a wider variety of aphasia severities and types than previous studies on the subject. Although this diversity can be considered a strength of the study because it represents a clinical reality, error types and error coding may have been substantially different than in previous studies.

All things considered, thematic informativeness, as measured here by TUs and GIM, constitutes an interesting outcome measure of discourse production with PWA in the early phase of recovery. This study included twenty-three PWA with varying types of language impairment and severity, reflecting well the clinical variations seen in

aphasia. The results suggest, therefore that these two measures may be suitable for the assessment of aphasia in the early stages.

Thematic informativeness was chosen for three main reasons. First, there have been only two studies that have used measures of thematic informativeness in the acute phase. Furlanis et al. (2018) reported significant changes between the different timepoints using a measure similar to GIM. Also, Agis et al. (2016) reported that the inclusion of the measure of total CUs with NIHSS strengthened the prediction of lesion location and volume compared to NIHSS alone. In the present study, two measures of informativeness, the General Informativeness Measure (GIM) and Thematic Units (TUs), demonstrated that both measures seem to be sensitive to improvement in early post-stroke aphasia recovery. TUs alone are much simpler to score than GIM, which requires additional scoring. However, GIM includes both macro- and micro-linguistic measures of discourses (Andreetta et al., 2012; Marini et al., 2011), providing a multidimensional measure able to capture early improvements in descriptive discourse. Therefore, bedside scoring of TUs and GIM should be explored and tested to be used eventually by clinicians. These methods would be simpler than any detailed method already in use, which would respond to the latest concerns raised concerning discourse studies (Dietz & Boyle, 2018). These measures can be easily adapted to other instruments, such as the picture description task of the BDAE (Kaplan, Goodglass, & Weintraub, 1983) and the MT-86 (Nespoulous et al., 1992) -- instruments that are also widely used in clinical settings with Canadian-French-speaking PWA. The results from Furlanis et al. (2018) were based on discourse samples obtained from Italian speakers and the participants in Agis et al. (2016) were English speakers. Our data are based on discourse samples obtained from Canadian-French speakers. These findings, therefore, suggest that the notion of informativeness might be a feature that transcends languages.

All together, these results suggest that thematic informativeness might be more related to global communication rather than to specific language features. In a future study, the comparison of thematic informativeness with an overall language measure in the acute phase, as has been suggested by Wilson et al. (2019), could prove the effectiveness of TUs and GIM as a general language assessment tool.

The second reason for investigating thematic informativeness was that, as a measure, it is easier to implement in acute aphasia assessments. Lexical informativeness assessment and phonological analysis, in other words microlinguistic analyses, rely on transcriptions which are often too time consuming to be conducted in acute clinical settings (Bryant et al., 2017). TU scoring is based on a finite set of themes, which increases the reliability of its use (Brookshire & Nicholas, 1994). Early assessments of PWA are generally time constrained as many patients cannot tolerate lengthy protocols (Marshall & Wright, 2007) and, as a result, bedside assessments must be quick to administer. Picture description only requires a few minutes with the patient and is easy to carry out, which reconciles quantifiable measures with practical clinical requirements.

Third, thematic informativeness seems to be more reliable than other informativeness measures. In the present study, IRR was excellent for both GIM and TUs. Although variables similar to thematic informativeness measures have already obtained very good inter-rater reliability (IRR) scores (e.g. content units in Agis et al., 2016), the results of the present study serve as a basis to build an effective instrument with very good inter-rater reliability to assess discourse production in acute aphasia.

It is generally accepted that thrombolysis has a positive effect on language recovery (Emberson et al., 2014; Rohde, Worrall, O. Halloran, Godecke, & Farrell, 2017; Röther et al., 2002; Seitz & Donnan, 2015; Seitz, Sukiennik, & Siebler, 2012;

Wardlaw, Murray, Berge, & del Zoppo, 2014), but little is known about the specific language features influenced by this treatment. The present results extend previous findings and suggest that reperfusion therapies increase very early recovery on thematic informativeness, but the effects of thrombolysis attenuate in the following days. The above results demonstrated that patients treated with thrombolysis produced more TUs than untreated patients. However, in contrast to the results of a previous study that used a similar variable (Furlanis et al., 2018), there was no difference over time between PWA treated with thrombolysis and those who were not treated. One possible explanation could be the difference in the timings of the assessments. In Furlanis et al. (2018) testing first occurred in the very first hours following the stroke and the second assessment was 2 days post-onset. In contrast, in the present study, testing was first conducted two days post-onset. In other words, the last assessment in Furlanis et al. was conducted at the time of our first assessment. The present study did not capture the very early changes that Furlanis et al. (2018) were able to identify, when several neurological processes occur.

This study, showing early improvement of informativeness in PWA, is part of a larger on-going longitudinal study. Further analyses will be completed using the long-term assessment data to determine whether thematic informativeness in the acute phase also predicts long-term (i.e. 6 months post-stroke) recovery. Although there was no difference between the two groups in the acute recovery phase, our results support previous evidence that GIM and TUs should be tested in clinical practice in the very early phase because they capture the longitudinal pattern of aphasia recovery.

Given the multiple statistical analyses that were done, we acknowledge that these results must be considered preliminary and that the analysis of discourse from a larger sample of individuals with aphasia will be important to further support these

findings. Another potential limitation is that segmentation agreement was not calculated. However, inter-rater reliability was high for the variables that support the main purpose of this paper, including the number of thematic units and utterances, suggesting that segmentation agreement would also be high. Also, test-retest stability remains to be established for the variable TU, even though stability for GIM was good. In future studies, it would be interesting to analyze the discourse of PWA before thrombolysis administration to allow for the direct assessment of thrombolysis on discourse recovery. As well, considering that thematic units and not lexical units were studied, the present results are easily translatable to other French-speaking populations and likely other languages; however, the microlinguistic variables may not be as generalizable. Finally, while in the present study thematic informativeness was scored based on transcriptions, it will be interesting to investigate whether our measure could be scored online by clinicians such as has been demonstrated by Agis et al. (2016).

Conclusion

This study aimed at describing early improvements in the discourse production of PWA. Thematic informativeness improved significantly in a group of 23 acute post-stroke patients with various aphasia types and varying levels of severity during the first week post-onset, whereas the microlinguistic features studied did not show significant changes. These preliminary results suggest that the analysis of thematic informativeness may be a useful measure that reflects early discourse recovery. Consequently, a measure of thematic informativeness could be considered as a clinical instrument in acute care settings.

Declaration of interest statement

The authors report no conflicts of interest.

Funding

This project was funded by a grant-in-aid from Heart and Stroke Foundation (grant number G-16-00014039) to K.M., E.R., A.D. and S.M.B..

K.M. and S.M.B. hold a Career Award from the “Fonds de Recherche du Québec – Santé”.

J.B. holds a scholarship from the “Fonds de Recherche du Québec – Santé”.

References

- Agis, D., Goggins, M. B., Oishi, K., Oishi, K., Davis, C., Wright, A., ... Hillis, A. E. (2016). Picturing the Size and Site of Stroke With an Expanded National Institutes of Health Stroke Scale. *Stroke; a Journal of Cerebral Circulation*.
<https://doi.org/10.1161/STROKEAHA.115.012324>
- Albright, E., & Purves, B. (2008). Exploring SentenceShaper™: Treatment and augmentative possibilities. *Aphasiology*, 22(7–8), 741–752.
<https://doi.org/10.1080/02687030701803770>
- Allibrio, G., Gori, M., Signorini, G., & Luzzatti, C. (2009). Un esame del linguaggio per la diagnosi dei deficit afasici al letto del malato. *Giornale Di Psicologia*.
- Andreetta, S., Cantagallo, A., & Marini, A. (2012). Narrative discourse in anomic aphasia. *Neuropsychologia*, 50(8), 1787–1793.
<https://doi.org/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.
<https://doi.org/10.1080/02687038.2014.979394>
- Armstrong, E. (2000). Aphasic discourse analysis: The story so far. *Aphasiology*, 14(9), 875–892. <https://doi.org/10.1080/02687030050127685>
- Béland, R., Lecours, A. R., Giroux, F., & Bois, M. (1993). The MT-86 β aphasia battery: A subset of normative data in relation to age and level of school education (Part II). *Aphasiology*, 7(4), 359–382. <https://doi.org/10.1080/02687039308249516>
- Boulanger, J., Lindsay, M., Gubitz, G., Smith, E., Stotts, G., Foley, N., ... Butcher, K. (2018). Canadian Stroke Best Practice Recommendations for Acute Stroke Management: *Prehospital, Emergency Department, and Acute Inpatient Stroke*

Care, 6th Edition, Update 2018. International Journal of Stroke, 13(9), 949–984.

<https://doi.org/10.1177/1747493018786616>

Brookshire, R. H., & Nicholas, L. E. (1994). Speech sample size and test-retest stability of connected speech measures for adults with aphasia. *Journal of Speech and Hearing Research, 37(2)*, 399–407. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8028321>

Brown, C., Snodgrass, T., Kemper, S. J., Herman, R., & Covington, M. A. (2008). Automatic measurement of propositional idea density from part-of-speech tagging. *Behavior Research Methods, 40(2)*, 540–545. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/18522065>

Bryant, L., Ferguson, A., & Spencer, E. (2016). Linguistic analysis of discourse in aphasia: A review of the literature. *Clin Linguist Phon, 30(7)*, 489–518. <https://doi.org/10.3109/02699206.2016.1145740>

Bryant, L., Spencer, E., & Ferguson, A. (2017). Clinical use of linguistic discourse analysis for the assessment of language in aphasia. *Aphasiology, 31(10)*, 1105–1126. <https://doi.org/10.1080/02687038.2016.1239013>

Capilouto, G. J., Wright, H. H., & Wagovich, S. A. (2006). Reliability of main event measurement in the discourse of individuals with aphasia. *Aphasiology, 20(2–4)*, 205–216. <https://doi.org/10.1080/02687030500473122>

Carlomagno, S., Pandolfi, M., Labruna, L., Colombo, A., & Razzano, C. (2001). Recovery from moderate aphasia in the first year poststroke: Effect of type of therapy. *Archives of Physical Medicine and Rehabilitation*. <https://doi.org/10.1053/apmr.2001.25155>

Colin, Capucine; Le Meur, C. (2016). *ADAPTATION DU PROJET APHASIABANK À*

LA LANGUE FRANÇAISE: Contribution pour une évaluation informatisée du discours oral de patients aphasiques. Université Paul Sabatier, Toulouse III.

Covington, M. A. (2007). *MATTR user manual (CASPR Research Report 2007–05)*. Athens, GA.

Covington, M. A., & McFall, J. D. (2010). Cutting the Gordian Knot: The Moving-Average Type–Token Ratio (MATTR). *Journal of Quantitative Linguistics*, 17(2), 94–100. <https://doi.org/10.1080/09296171003643098>

Craig, H. K., Winkelseth, M., Carry, L., Walley, J., Bardach, L., Higman, B., ...

Sheimo, D. (1993). Quantifying connected speech samples of adults with chronic aphasia. *Aphasiology*. <https://doi.org/10.1080/02687039308249503>

Dalton, S. G., & Richardson, J. D. (2015). Core-Lexicon and Main-Concept Production During Picture-Sequence Description in Adults Without Brain Damage and Adults With Aphasia. *American Journal of Speech-Language Pathology*, 24(4), S923–S938. https://doi.org/10.1044/2015_AJSLP-14-0161

De Renzi, E., & Faglioni, P. (1978). Normative data and screening power of a shortened version of the Token Test. *Cortex*, 14(1), 41–49. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/16295108>

Deloche, G., & Hannequin, D. (1997). *Test de dénomination orale d'images: DO-80*. Éditions du centre de psychologie appliquée.

Dietz, A., & Boyle, M. (2018). Discourse measurement in aphasia research: have we reached the tipping point? *Aphasiology*, 32(4), 459–464. <https://doi.org/10.1080/02687038.2017.1398803>

Doyle, P. J., Goda, A. J., & Spencer, K. A. (1995). The Communicative Informativeness and Efficiency of Connected Discourse by Adults With Aphasia Under Structured

NARRATIVE DISCOURSE RECOVERY

Aphasiology (accepted)

and Conversational Sampling Conditions. *American Journal of Speech-Language Pathology*, 4(4), 130. <https://doi.org/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.

<https://doi.org/10.1080/026870300401306>

Emberson, J., Lees, K. R., Lyden, P., Blackwell, L., Albers, G., Bluhmki, E., ... Stroke

Thrombolysis Trialists' Collaborative Group. (2014). Effect of treatment delay, age, and stroke severity on the effects of intravenous thrombolysis with alteplase for acute ischaemic stroke: a meta-analysis of individual patient data from randomised trials. *The Lancet*, 384(9958), 1929–1935.

[https://doi.org/10.1016/S0140-6736\(14\)60584-5](https://doi.org/10.1016/S0140-6736(14)60584-5)

Fergadiotis, G., & Wright, H. H. (2016). Modelling confrontation naming and discourse performance in aphasia. *Aphasiology*, 30(4), 364–380.

<https://doi.org/10.1080/02687038.2015.1067288>

Fromm, D., Forbes, M., Holland, A., Dalton, S. G., Richardson, J., & MacWhinney, B.

(2017). Discourse Characteristics in Aphasia Beyond the Western Aphasia Battery Cutoff. *Am J Speech Lang Pathol*, 26(3), 762–768.

https://doi.org/10.1044/2016_AJSLP-16-0071

Fromm, D., Greenhouse, J., Hou, K., Russell, G. A., Cai, X., Forbes, M., ...

MacWhinney, B. (2016). Automated Proposition Density Analysis for Discourse in Aphasia. *J Speech Lang Hear Res*, 59(5), 1123–1132.

https://doi.org/10.1044/2016_jslhr-1-15-0401

Furlanis, G., Ridolfi, M., Polverino, P., Menichelli, A., Caruso, P., Naccarato, M., ...

- Manganotti, P. (2018). Early Recovery of Aphasia through Thrombolysis: The Significance of Spontaneous Speech. *Journal of Stroke and Cerebrovascular Diseases*, 27(7), 1937–1948.
<https://doi.org/10.1016/j.jstrokecerebrovasdis.2018.02.043>
- Goodglass, H., Kaplan, E., Barresi, B., Goodglass, H., Goodglass, H., Goodglass, H., ... Kaplan, E. (2001). The Boston Diagnostic Aphasia Examination : BDAE-3 long form kit.
- Joanette, Y., Ska, B., & Côté, H. (2004). *Protocole Montréal d'Évaluation de la Communication*. Isbergues, France: Ortho Édition.
- Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test*. Philadelphia: Lea & Febiger.
- Kertesz, A. (2006). *Western Aphasia Battery- Revised*. Pearson.
- Kertesz, A., & McCabe, P. (1977). Recovery patterns and prognosis in aphasia. *Brain*.
<https://doi.org/10.1093/brain/100.1.1>
- Kintz, S., & Wright, H. H. (2018). Discourse measurement in aphasia research. *Aphasiology*, 32(4), 472–474. <https://doi.org/10.1080/02687038.2017.1398807>
- Larfeuil, C., & Le Dorze, G. (1997). An analysis of the word-finding difficulties and of the content of the discourse of recent and chronic aphasic speakers. *Aphasiology*.
<https://doi.org/10.1080/02687039708250456>
- Laska, A. C., Hellblom, A., Murray, V., Kahan, T., & Von Arbin, M. (2001). Aphasia in acute stroke and relation to outcome. *Journal of Internal Medicine*.
<https://doi.org/10.1046/j.1365-2796.2001.00812.x>
- Lazar, R. M., Minzer, B., Antonello, D., Festa, J. R., Krakauer, J. W., & Marshall, R. S.

- (2010). Improvement in aphasia scores after stroke is well predicted by initial severity. *Stroke*, *41*(7), 1485–1488.
<https://doi.org/10.1161/STROKEAHA.109.577338>
- Linnik, A., Bastiaanse, R., & Höhle, B. (2016). Discourse production in aphasia: a current review of theoretical and methodological challenges. *Aphasiology*, *30*(7), 765–800. <https://doi.org/10.1080/02687038.2015.1113489>
- MacWhinney, B. (2000). *The CHILDES Project: Tolls for Analyzing Talk* (Vol. 3rd Editio). Mahwah, NJ: Lawrence Erlbaum Associates.
- Marcotte, K., McSween, M.-P., Pouliot, M., Martineau, S., Pauzé, A.-M., Wiseman-Hakes, C., & Macdonald, S. (2017). Normative study of the functional assessment of verbal reasoning and executive strategies (FAVRES) test in the French-Canadian population. *Journal of Speech, Language, and Hearing Research*, *60*(8). https://doi.org/10.1044/2017_JSLHR-L-17-0012
- Marini, A., Andretta, S., del Tin, S., & Carlomagno, S. (2011). A multi-level approach to the analysis of narrative language in aphasia. *Aphasiology*. <https://doi.org/10.1080/02687038.2011.584690>
- Marini, A., Caltagirone, C., Pasqualetti, P., & Carlomagno, S. (2007). Patterns of language improvement in adults with non-chronic non-fluent aphasia after specific therapies. *Aphasiology*. <https://doi.org/10.1080/02687030600633799>
- Marini, A., Galetto, V., Zampieri, E., Vorano, L., Zettin, M., & Carlomagno, S. (2011). Narrative language in traumatic brain injury. *Neuropsychologia*, *49*(10), 2904–2910. <https://doi.org/10.1016/j.neuropsychologia.2011.06.017>
- Marini, A., Spoletini, I., Rubino, I. A., Ciuffa, M., Bria, P., Martinotti, G., ... Spalletta, G. (2008). The language of schizophrenia: An analysis of micro and

- macrolinguistic abilities and their neuropsychological correlates. *Schizophrenia Research*. <https://doi.org/10.1016/j.schres.2008.07.011>
- Marshall, R. C., & Wright, H. H. (2007). Developing a Clinician - Friendly Aphasia Test, *16*(November), 295–315. [https://doi.org/10.1044/1058-0360\(2007/035\)](https://doi.org/10.1044/1058-0360(2007/035))
- Monetta, L. et tous les membres du R. regroupement stratégique #4. (2014). *Outils d'évaluation orthophoniques franco- québécois validés ou normés, de 1980 à 2014*.
- Nespoulous, J. L., Lecours, A. R., & Lafond, D. (1986). MT-86-Protocole Montréal-Toulouse d'examen linguistique de l'aphasie. *Ortho-Edition: Isbergues*.
- Nespoulous, J. L., Lecours, A. R., Lafond, D., Lemay, M. A., Puel, M., Joannette, Y., ... Rascol, A. (1992). *Protocole Montréal-Toulouse d'examen linguistique de l'aphasie: MT-86 module standard initial, M1b(2e édition révisée par Renée Béland et Francine Giroux)*. Isbergues, France: Ortho Edition.
- Nicholas, L. E., & Brookshire, R. H. (1993). A system for quantifying the informativeness and efficiency of the connected speech of adults with aphasia. *J Speech Hear Res*, *36*(2), 338–350.
- Nicholas, L. E., & Brookshire, R. H. (1995). Presence, completeness, and accuracy of main concepts in the connected speech of non-brain-damaged adults and adults with aphasia. *Journal of Speech and Hearing Research*, *38*(1), 145–156. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/7731205>
- Pedersen, P. M., Stig Jørgensen, H., Nakayama, H., Raaschou, H. O., Olsen, T. S., Jørgensen, H. S., ... Olsen, T. S. (1995). Aphasia in acute stroke: Incidence, determinants, and recovery. *Annals of Neurology*, *38*(4), 659–666. <https://doi.org/10.1002/ana.410380416>

- Pritchard, M., Hilari, K., Cocks, N., & Dipper, L. (2017). Reviewing the quality of discourse information measures in aphasia. *International Journal of Language & Communication Disorders*, 52(6), 689–732. <https://doi.org/10.1111/1460-6984.12318>
- Rohde, A., Worrall, L., O. Halloran, R., Godecke, E., & Farrell, A. (2017). 162: BRISBANE EVIDENCE-BASED LANGUAGE TEST, BRISBANE EBLT: A NEW DIAGNOSTIC TEST FOR THE IDENTIFICATION OF ACUTE POST-STROKE LANGUAGE DISORDERS. *BMJ Open*, 7(Suppl 1). Retrieved from http://bmjopen.bmj.com/content/7/Suppl_1/bmjopen-2016-015415.162.abstract
- Röther, J., Schellinger, P. D., Gass, A., Siebler, M., Villringer, A., Fiebach, J. B., ... Kompetenznetzwerk Schlaganfall Study Group. (2002). Effect of intravenous thrombolysis on MRI parameters and functional outcome in acute stroke. *Stroke*, 33(10), 2438–2445. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12364735>
- 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. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/2804622>
- Seitz, R. J., & Donnan, G. A. (2015). Recovery Potential After Acute Stroke. *Frontiers in Neurology*, 6, 238. <https://doi.org/10.3389/fneur.2015.00238>
- Seitz, R. J., Sukiennik, J., & Siebler, M. (2012). Outcome after systemic thrombolysis is predicted by age and stroke severity: an open label experience with recombinant tissue plasminogen activator and tirofiban. *Neurology International*, 4(2), 9. <https://doi.org/10.4081/ni.2012.e9>
- Sherratt, S. (2007). Multi-level discourse analysis: A feasible approach. *Aphasiology*,

21(3–4), 375–393. <https://doi.org/10.1080/02687030600911435>

Sloetjes, H., & Wittenburg, P. (2008). *Annotation by category-ELAN and ISO DCR*.

Retrieved from http://www.lrec-conf.org/proceedings/lrec2008/pdf/208_paper.pdf

Streiner, D. L., & Norman, G. R. (2008). Reliability. In *Health Measurement Scales* (pp. 167–210). Oxford University Press.

<https://doi.org/10.1093/acprof:oso/9780199231881.003.0008>

Swindell, C. S. C., Holland, A. A. L., & Fromm, D. (1984). Classification of aphasia: WAB type versus clinical impression. In *Clinical Aphasiology*.

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

Wade, D. T., Hewer, R. L., David, R. M., & Enderby, P. (1986). Aphasia after stroke: Natural history and associated deficits. *Journal of Neurology Neurosurgery and Psychiatry*. <https://doi.org/10.1136/jnnp.49.1.11>

Wardlaw, J. M., Murray, V., Berge, E., & del Zoppo, G. J. (2014). Thrombolysis for acute ischaemic stroke. *Cochrane Database of Systematic Reviews*, (7), CD000213. <https://doi.org/10.1002/14651858.CD000213.pub3>

Wilson, S. M., Eriksson, D. K., Brandt, T. H., Schneck, S. M., Lucanie, J. M., Burchfield, A. S., ... Kidwell, C. S. (2019). Patterns of Recovery From Aphasia in the First 2 Weeks After Stroke. *Journal of Speech, Language, and Hearing Research*, *62*(3), 723–732. https://doi.org/10.1044/2018_JSLHR-L-18-0254

Wright, H. H., Capilouto, G., Wagovich, S., Cranfill, T., & Davis, J. (2005).

Development and reliability of a quantitative measure of adults' narratives.

Aphasiology, *19*(3–5), 263–273. <https://doi.org/10.1080/02687030444000732>

NARRATIVE DISCOURSE RECOVERY

Aphasiology (accepted)

Yorkston, K. M., & Beukelman, D. R. (1980). An analysis of connected speech samples

of aphasic and normal speakers. *The Journal of Speech and Hearing Disorders*,

45(1), 27–36.