

* Cet article a été accepté dans *Aphasiology* chez Taylor & Francis Online.

<http://dx.doi.org/10.1080/02687038.2021.1907295>

A longitudinal study of narrative discourse in post-stroke aphasia

Amélie Brisebois, M.Sc.^{a,b}, Simona Maria Brambati, Ph.D.^{c,d}, Johémie Boucher B.Sc.^{c,d},
Elizabeth Rochon, Ph.D.^{e,f,g,h}, Carol Leonard, Ph.D.ⁱ, Marianne Désilets-Barnabé B.Sc.^{a,b}
, 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, ^{[[[]]}SEP 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.

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

A longitudinal study of narrative discourse in post-stroke aphasia

Abstract

Background: Previous findings have demonstrated the importance of discourse analysis in post-stroke aphasia, as it allows for in-depth examination of language impairment and represents key components of functional communication.

However, little is known about the recovery of discourse over time.

Aims: The main aim of this study is to measure the longitudinal changes in descriptive discourse production from the acute to chronic stages of post-stroke aphasia recovery. The secondary aim is to explore the association between discourse measures and overall language impairment severity measures at different testing points.

Methods & Procedure: Seventeen French Canadian speakers with various types and severities of aphasia following a first left middle cerebral artery stroke participated in this study. They underwent three language assessments (acute: 0 to 72 hours; subacute: 7 to 14 days; chronic: 6 to 12 months post-onset). The picture description from the Western Aphasia Battery was analyzed at three time points. Changes in terms of thematic informativeness and microstructural variables were analyzed.

Outcomes & Results: Regarding the micro-structural variables, the mean length of utterances (MLU) and the number of words per minute showed significant positive changes between the acute and chronic phases. For the thematic informativeness measures, the number of thematic units (TUs), the number of thematic units per minute (TUs/min) and the number of thematic units per utterance (TUs/utt)

increased significantly between the acute and chronic phases. Positive correlations between TUs and MLU in the acute phase and a general language impairment severity measure in the acute and chronic phases suggest a relationship between these measures and global language performance suggesting the potential predictive value of these variables in the acute phase.

Conclusions & Implications: These findings support the use of thematic units in descriptive discourse analysis during an acute clinical examination of language as they require minimal additional time to score and track changes in post-stroke aphasia recovery. They capture long-term changes in discourse abilities and appear related to overall language measures in both the acute and chronic stages of recovery. The interpretation of the changes in MLU and the number of words per minute is less straightforward, as improvements in these measures carry different interpretations depending on the type of aphasia. Nonetheless, further studies are required to investigate test-retest reliability and the effect of therapy on the changes observed over time when using thematic units to document change in discourse.

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

Background

Discourse generally refers to language units larger than an utterance or a sentence (Kong, 2016) and is considered the most elaborated verbal production (Ska et al., 2004). Discourse tasks have received increasing interest over the last decade as they are now considered one of the most ecologically valid assessments of language impairments. According to the International Classification of Functioning, Disability, and Health (ICF) model, discourse is a key language component because its performance relates to social participation (Armstrong et al., 2012). It is no surprise that clinical impressions often rely on discourse assessment to document a diagnosis, establish goals, and measure outcomes (Boyle, 2020; Bryant et al., 2017). However, research on discourse abilities has received much less attention than the production of single words in acquired language disorders, largely because discourse analysis is time-consuming (Boles, 1998) and challenging considering the difficulty involved in identifying the key components that represent the complexity of post-stroke aphasia profiles (Linnik, Bastiaanse, & Höhle, 2016). The present study will focus on longitudinal changes in discourse production from the acute to chronic stages in French-speaking individuals with post-stroke aphasia. Also, this study aims to contribute to the ongoing questions on language predictors in stroke (Kiran & Thompson, 2019) by presenting an exploratory analysis of associations between a discourse assessment collected in the acute stage of recovery (i.e., baseline) and an overall language impairment severity measure collected both in acute and in chronic stages of recovery.

Discourse elicitation tasks in aphasia

The nature of discourse elicitation tasks allows for a wide range of spoken output (Pritchard, Hilari, Cocks, & Dipper, 2017), and the choice of discourse measures is related to the elicitation task (Stark, 2019). Among them, spontaneous and interactive conversational simulations or observation of functional activities (e.g., talking over the phone) provide valuable information about the impact of language impairments on everyday activities and social participation (Holland, Frattali, & Fromm, 1999). On the other hand, types of discourse tasks that are not interactive—such as picture description, story retell, or procedural discourse—generate more constrained productions, which allow for a more precise comparison over time and between individuals (Bryant et al., 2016; Linnik et al., 2016). Picture description tasks are widely used in the clinical assessment of aphasia. They consist of a detailed description of a pictorial stimulus representing a complex scene and provide a relatively constrained discourse sample with expected topics (Cherney, 1998). Thus, they allow for a standardized approach to studying language production in context and facilitate performance comparison over time and across different groups (Kong, 2016). Also, the number of words and the number of correct information units collected with a structured task such as picture descriptions are similar to that collected in unstructured speech-elicited tasks such as conversation (Doyle et al., 1995). Moreover, it seems that using a highly structured stimulus, such as a picture, provides a predictable language context because of a stable and shared semantic frame, allowing for more precise inter-individual comparisons (Bryant et al., 2016).

Discourse analysis

Spoken discourse analysis has the potential to provide important and meaningful outcome measures in aphasia (Wallace et al., 2018), but rigorous standardization of the analysis is needed to support its development (Stark et al., 2021). Spoken discourse can be analyzed by focusing on the microstructure of language (e.g., lexicon, syntax, word classes), by focusing on the macrostructure (e.g., story grammar, communicative effectiveness) or using measures that represent the interface of micro- and macro-structural structures (e.g., content units, main concepts) (Armstrong, 2000). These multiple aspects are important in supporting efficient spoken discourse and can be independently impaired. A growing body of evidence suggests that some discourse information measures lack adequate reliability and stability to reflect language changes related to natural recovery or therapy (Boyle, 2014; Pritchard et al., 2017; Stark et al., 2020), but there is a clear of data in less prevalent languages (i.e., French-Canadian). As such, the present study will focus on the changes over time of both microstructural variables and variables situated at the interface between micro and macrostructural ones, such as informativeness, in French-Canadian speaking individuals with post-stroke aphasia (Armstrong, 2000).

Microstructural measures

Microstructure refers to the language system itself, including phonological, lexical, semantic and grammatical processing (Andretta et al., 2012). Several micro-structural variables, such as overall verbal productivity, utterance length, and syntactic complexity, can be automatically extracted using the Computerized Language Analysis (CLAN)

(MacWhinney, 2000), a free software program which provides a detailed transcription method and speedy automatic computation of a wide range of language features. The program now has a database of language samples of individuals suffering from post-stroke aphasia and neurotypical individuals (AphasiaBank) (MacWhinney & Fromm, 2016). However, no patient data are yet available in French-Canadian and only 14 samples from neurotypical individuals are available in French.

Consistent findings have shown that microstructural measures of discourse in people with chronic aphasia differ from those of healthy individuals. Speech rate, percentage of semantical and phonological errors, moving average token-type ratio, and number of verbs per utterance are amongst the measures with the most diagnostic sensitivity (Andretta, Cantagallo, & Marini, 2012; Andretta & Marini, 2015; Covington & McFall, 2010; Saffran, Berndt, & Schwartz, 1989; Wright, Capilouto, Wagovich, Cranfill, & Davis, 2005). Despite the growing importance of spoken discourse analysis in post-stroke aphasia, little is known about the longitudinal changes on these microstructural measures. Previous findings suggest that they do not evolve as much as expected in the acute and sub-acute phases of recovery (Brisebois et al., 2020), or even when language therapy is provided (Carlomagno et al., 2001; Larfeuil & Le Dorze, 1997; Marini et al., 2007).

Discourse information measures

Despite methodological challenges, as part of a comprehensive discourse assessment, informativeness discourse measures have received much attention in the literature (see review by Pritchard et al., 2017), as aphasia affects the ability to convey information — a basic language skill related to functional communication (Armstrong et

al., 2012). Since post-stroke aphasia can have an impact on the quantity and quality of information communicated in discourse (Nicholas & Brookshire, 1993), both micro- (i.e., quantity) and macro- (i.e., quality) linguistic variables of informativeness have been investigated. Moreover, discourse variables appear on a continuum between micro- and macro- structural aspects (Armstrong, 2000). Therefore, informativeness is often considered ‘in between’ micro- and macro- structural aspects. On the one hand, the most micro-structural variables of informativeness allow for in-depth content and structure analyses, addressing the fundamental underpinnings of lexical processing. Micro-structural variables assessing informativeness, such as CIU and “lexical information units” have in common that they assess informativeness at the word level of connected speech and both have great diagnostic sensitivity. As opposed to their micro-structural informativeness counterparts, thematic informativeness assesses content beyond the word level, and are considered more macrostructural. Thematic informativeness is defined as expected semantic elements in the discourse (information, main events/ideas/concepts) that are relevant to the topic of a stimulus (Marini, Carlomagno, Caltagirone, & Nocentini, 2005; Marini et al., 2011). This type of analysis is particularly appropriate to assess discourse performance on a picture description task since the performance on such a task relies primarily on the patient’s abilities to interpret the stimulus and formulate a message representing that stimulus (Kong, 2016). This type of analysis has also been of particular interest because a determined set of information units allows for an easier comparison of the discourse performance between people with chronic aphasia and healthy speakers (Dalton & Richardson, 2015) and within groups of people with aphasia over time (Brisebois et al., 2020).

Conflicting evidence has been reported in individuals with chronic post-stroke aphasia regarding information units. Indeed, this measure was reported to be sensitive to treatment response in PWA (Albright & Purves, 2008; Avent & Austermann, 2003; Coelho, Mchugh, & Boyle, 2000; Stark, 2010) and correlated with listener perceptions (Cupit et al., 2010; Ross, 1999). Many other studies have reported differences in information units between neurologically intact adults and PWA (Agis et al., 2016; Capilouto, Wright, & Maddy, 2016; Nicholas & Brookshire, 1995; Ulatowska et al., 1983). However, other studies have not identified impairments of information units in PWA compared to healthy speakers (Albright & Purves, 2008; Doyle et al., 1995, 2000). One possible explanation for the conflicting evidence is the severity of the language impairments of the patients included in each of these studies. For instance, impairments in informativeness have been reported in severe aphasia (e.g., Ulatowska et al., 1983), whereas other studies identified no such impairment in people with mild or moderate aphasia (e.g., Yorkston & Beukelman, 1980). However, recent research (e.g., Alyahya et al., 2020; Brisebois et al., 2020) has demonstrated that informativeness impairment affects PWA with various language impairment severities in acute, subacute and chronic stages of recovery.

Recently, a growing body of literature has investigated thematic informativeness in acute aphasia, as it allows for a comprehensive analysis of discourse performance and seems to capture changes in early language recovery. Thematic informative measure of Content Units (CUs) (Yorkston & Beukelman, 1980) in the description of the Cookie Theft picture from The Boston Diagnostic Aphasia Examination (BDAE) (Goodglass et al., 2001) provided valuable information about volume and lesion location within 48

hours of stroke onset in people with right- and left- hemisphere ischemic stroke with or without language impairment (Agis et al., 2016). The authors highlighted a close relationship between informativeness and neural damage, demonstrating that some variables, including CUs, improve prediction of brain damage volume and location. On a more practical note, CUs could be easily implemented in clinical settings as the scoring and the analysis require minimal additional time (Agis et al., 2016). Likewise, in a study investigating hyperacute language recovery, descriptive discourse, measured by the addition of information units and a global appreciation of phonological and syntactic performance, was the only variable amongst other language variables that significantly improved between 0 and 72 hours post-stroke in a group of PWA (Furlanis et al., 2018). Our team also recently reported on the early changes in informativeness in the picture description of the Western Aphasia Battery Revised (Kertesz, 2006) in a group of 23 PWA with different types of aphasia and severity (i.e., ranging from mild to severe) (Brisebois et al., 2020). We calculated informativeness in terms of macro-structural units which consisted of a specific list of thematic units previously identified based on the production of healthy speakers (as described in Brisebois et al., 2020). Hence, for each participant, a maximum of 16 thematic units could be produced, irrespective of the length/elaboration of the description. As a group, patients showed significant improvement in the number of thematic units produced between 0 to 72 hours and 7 to 14 days post-onset, whereas no micro-structural variables improved significantly during the same time. These results suggest that informativeness measures are of particular interest in acute post-stroke aphasia because they seem sensitive to change, they require < 3

minutes to score, and also contribute to the prediction of neural correlates related to recovery (Agis et al., 2016).

Longitudinal discourse studies in aphasia

Despite an increasing number of studies on discourse abilities in PWA (see review by Bryant, Ferguson, & Spencer, 2016) and the fact that aphasia recovery may take up to a year or even longer (e.g., Bernhardt et al., 2017), little is known about how discourse evolves from the acute to the chronic phase of aphasia recovery. Among the few studies conducted to date, cohesive abilities were documented with a narrative discourse task at 1-, 6-, and 12-months post-stroke in a group of patients who had not been diagnosed with aphasia following a left hemispheric stroke (Ellis et al., 2005). The analysis demonstrated that while the number of cohesive ties is stable in the first year following a stroke, the percentage of correct cohesive ties increased significantly during the same period. These results support the fact that the natural recovery of subtle language disruptions following a stroke is important. More recently, changes in the production of cohesive ties in narrative discourse have been investigated in a large group of participants with right (n=76) and left (n=145) hemisphere ischemic stroke at two time points -- <1 week and 6-12 months post-stroke (Stockbridge et al., 2021). This study demonstrated that deficits in cohesion occur whether the stroke is located in the right or left hemisphere, and that patterns of cohesive markers are different across the two groups of participants. However, the authors reported minimal changes over time in the production of cohesion markers, possibly due to very high interindividual variability. Indeed, discourse is unstable by nature, even in chronic PWA (Boyle, 2014, 2015). A recent scoping review showed that most longitudinal studies reporting on cognitive impairment have been conducted using

screening tools that focus on the impairment level, such as naming (Saa et al., 2019). Based on prevailing knowledge that improvements are expected in post-stroke aphasia during the acute and sub-acute phases of recovery (e.g., Bernhardt et al., 2017; Pedersen et al., 2004; Saur et al., 2006), quantitative description of changes in discourse over time should be reported more systematically to accurately describe the trajectory of discourse abilities after stroke.

Purpose

The main goal of the present study is, thus, to measure changes in spoken production of French-Canadian speaking individuals with post-stroke aphasia using the picture description task of the WAB-R (Kertesz, 2006). To do so, micro-structural variables that have been reported to be sensitive to language impairment (namely number of words per minute and mean length of utterance) (Andretta et al., 2012; Andretta & Marini, 2015; Fergadiotis & Wright, 2015; Fromm et al., 2016) were collected to document in-depth linguistic aspects of discourse production. Thematic informativeness variables (Brisebois et al., 2020) were also collected as they refer to broader discourse metrics that represent semantic content and linguistic form but are also easy to implement in clinical settings and can be scored rapidly. Based on previous longitudinal studies, (e.g., Furlanis et al., 2018), and also on the fact that language recovery continues for a long period after stroke (Bernhardt et al., 2017; Pedersen et al., 2004), we expect positive changes over the course of time. More specifically, we predict that thematic units will demonstrate an increase from the acute to chronic stages of testing. Also, considering prior results (Brisebois et al., 2020), micro-structural variables should improve later in recovery. Our secondary goal is to explore potential relationships between the discourse

measures in the acute period to a global language severity measure obtained in the acute and chronic stages. This secondary goal is twofold. First, we seek a relationship between the discourse scores and a global language severity score in the acute stage. Because it is believed that thematic units provide a general assessment of discourse ability (Brisebois et al., 2020), we expect a positive relationship between thematic units and the global language severity measure. Second, we will investigate whether discourse measures in the acute stage are associated with the long-term language severity measure in the chronic stage. Such an association would support the future investigation of the use of early discourse measures as a clinical predictor of long-term language outcomes in post-stroke aphasia.

Methods and Materials

Participants

Patients with aphasia were recruited in 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 2019 in the context of a larger project that sought to investigate the longitudinal recovery of white matter structural connections mediating early and late post-stroke aphasia recovery. No criteria concerning initial aphasia severity or lesion size were applied. Seventeen native French-Canadian speakers (eight women, mean age: 72.12 ± 12.67 years old; mean education: 12.59 ± 4.27 years) with various types of post-stroke aphasia following an ischemic stroke located in the left middle cerebral artery territory participated in the present study. Four were monolinguals (French-Canadian only), 10 were bilinguals (French-Canadian and another language) and

three spoke three languages. However, for all participants French-Canadian as their dominant language and the language used for testing. Individual clinical and sociodemographic data of all participants are presented in Table 1. The participants of the present study represent a subset of the participants of a previous study (n=17) that targeted early discourse recovery during the acute and subacute phases (Brisebois et al., 2020).

[Table 1 should be inserted here]

Table 1. Demographic and clinical variables of participants with post-stroke aphasia

Participant	Sex	Age	Educ.	Language status	Initial NIHSS score	Lesion vol (ml)	Days _{acute} (T1)	Days _{subacute} (T2)	Days _{chronic} (T3)	CS _{acute} (T1)	CS _{subacute} (T2)	CS _{chronic} (T3)
1	M	52	9	Monolingual	n/a	35	1	7	387	8.20	24.78	27.3
2	M	74	6	Monolingual	9	30	3	8	365	10.24	13.81	24.6
3	M	73	19	Bilingual	18	16	3	10	224	7.71	14.02	27.3
4	F	70	14	Trilingual	16	95	3	12	249	1.87	1.69	5.3
5	M	83	9	Bilingual	9	35	3	10	366	3.90	14.39	18.3
6	F	47	18	Trilingual	26	71	0	10	218	0.00	0.00	18.3
7	F	73	7	Trilingual	n/a	6	3	13	217	14.36	17.23	16.3
8	M	65	11	Bilingual	6	12	3	14	196	28.53	28.88	29.3
9	M	72	15	Bilingual	11	1	1	9	188	21.33	28.11	28.3

Participant	Sex	Age	Educ.	Language status	Initial NIHSS score	Lesion vol (ml)	Days _{acute} (T1)	Days _{subacute} (T2)	Days _{chronic} (T3)	CS _{acute} (T1)	CS _{subacute} (T2)	CS _{chronic} (T3)
10	M	73	11	Unilingual	n/a	16	1	8	231	12.76	14.79	24.5
11	M	64	15	Bilingual	n/a	2	1	11	277	27.46	28.90	28.5
12	F	95	6	Bilingual	1	13	2	9	251	16.27	22.86	23.5
13	F	60	12	Bilingual	7	0.29	3	13	232	23.60	21.73	21.5
14	M	91	19	Bilingual	7	0.17	3	15	383	20.09	25.08	25.5
15	F	85	16	Bilingual	n/a	10	2	8	227	26.79	27.70	27.5
16	F	81	15	Monol	17	14	2	11	255	12.30	22.39	26.5
17	F	68	12	Bilingual	4	14	3	12	219	29.50	28.78	28.5

Participant	Sex	Age	Educ.	Language status	Initial NIHSS score	Lesion vol (ml)	Days _{acute} (T1)	Days _{subacute} (T2)	Days _{chronic} (T3)	CS _{acute} (T1)	CS _{subacute} (T2)	CS _{chronic} (T3)
group Mean (SD)		72.12 (12.67)	12.59 (4.27)		11.09 (7.38)	20.99 (26.32)	2.18(1.01)	10.59(2.32)	263.82 (67.25)	15.58 (9.61)	21.77 (11.29)	25.11 (13.11)

NIHSS= National Institute of Health Stroke Scale; n/a= non-available in the medical chart; ^a CS_{acute} = (Naming_{acute} (/10) + Comprehension_{acute} (/10) + Repetition_{acute} (/10)), ^b CS_{subacute} = (Naming_{subacute} (/10) + Comprehension_{subacute} (/10) + Repetition_{subacute} (/10)), ^c CS_{chronic} = (Naming_{chronic} (/10) + Comprehension_{chronic} (/10) + Repetition_{chronic} (/10)); BDAE= Boston Diagnostic Aphasia Examination;

Exclusion criteria were a history of major psychiatric disorders, learning disabilities, or severe perceptual deficits, as identified by the on-call physician left-handedness, or additional neurological diagnoses. The study was approved by the ethics review board of the *Centre intégré universitaire de santé et de services sociaux du Nord-de-l'Ile-de Montréal* (Project #MP-32-2018-1478) and written informed consent was obtained from all participants.

Assessments

The participants underwent three language assessments over time. The first assessment (T1; acute phase) occurred within the first 72 hours post-onset (mean=2.2 days; SD= 1.0), the second assessment (T2; subacute phase) took place approximately 10 days post-onset (range=7 to 15 days, mean=10.6 days, SD=2.3), and the third assessment (T3; chronic phase) took place approximately eight months (range=188 to 387 days, mean=263.8 days; SD=67.3) post-onset. The specific timing for each assessment is indicated for each PWA in Table 1. By the last assessment, all participants were medically stable and had received speech-language therapy, as recommended by the Canadian Stroke Guidelines (Boulangier et al., 2018). The amount, intensity, and the nature of language therapy varied across participants. As the study took place in Canada, which has a public health care system, all participants received a certain amount of language therapy, which ranged from a just a few sessions to a several months. Taken as a group, the therapy provided differed regarding the timing, intensity, duration and type, in response to each individual's language impairments (type and severity) and personal

goals. At the time of the third assessment, no participants were actively involved in speech-language therapy.

All PWA completed a comprehensive language assessment. As in Brisebois et al. (2020), language impairments were thoroughly documented and evaluated at T1, T2, and T3. Based on Lazar et al. (2010), where language recovery with English speakers was investigated, we developed a composite score (CS), adapted for the French-speaking population. Sections of protocols that are widely used with French-Canadian speaking patients were used. This Composite Score (CS) consisted of three subscores:

Comprehension, Repetition and Naming (Osa García et al., 2020). The Comprehension subscore, combined the Word-Sentence Comprehension Task (max = 47 points) of the Montreal-Toulouse test (Nespoulous et al., 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 test (Nespoulous et al., 1986): (2 points for each word/nonword (n=30) and 5 points for each sentence (n=3), max = 75 points). Finally, the Naming sub-score 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 et al., 2004). Each of the three domains had a possible maximum score of 10, resulting in a maximum composite score (CS) of 30. Initial (CS_{acute}), subacute (CS_{subacute}), and chronic (CS_{chronic}) individual and mean scores are reported in Table 1. Severity scoring and aphasia type were based on the results obtained on these tasks, clinical judgement and overall rating on the BDAE severity scale (Goodglass et al., 2001). All

participants were asked to produce an oral description of the Picnic picture of the Western Aphasia Battery –Revised (Kertesz, 2006). They were encouraged to talk in sentences and pay attention to all aspects of the picture. Picture description samples were filmed using a Sony HDR-PJ540 camera (9.2 mega pixels).

Procedure

The procedures, including the thematic informativeness coding and analysis as well as the micro-structural analysis, have been previously reported (Brisebois et al., 2020). 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).

Thematic informativeness coding and analysis

Thematic Units (TUs) were selected based on the WAB picture descriptions of 45 healthy French-Canadian speakers (as in Brisebois et al. (2020)). Sixteen TUs identified by at least 75% of these participants were included in the analysis grid and given a score of one point (see Brisebois et al., 2020). A similar selection procedure has also been employed in previous studies (e.g., Marini et al., 2011). This measure includes the total number of specific units that the participants produced, where a maximum of 16 thematic units could be obtained. A list of Thematic Units is provided in Appendix 1. Efficiency was also considered by calculating the number of TUs per minute and the number of TUs per utterance. Each variable was computed in acute, subacute and chronic phases, yielding to these variables: TU_{Sacute} , TU_{Subacute} , TU_{Schronic} , $TUs/\text{min}_{\text{acute}}$, $TUs/\text{min}_{\text{subacute}}$,

TUs/min_{chronic}, TUs/utt_{acute}, TUs/utt_{subacute} and TUs/utt_{chronic}.

Micro-structural analysis

Transcriptions underwent detailed linguistic and textual analysis focusing on micro-structural measures known to be affected in aphasia (e.g., Andreetta, Cantagallo, & Marini, 2012). All micro-structural variables were extracted using the program EVAL of the Computerized Language ANalysis (CLAN) program (MacWhinney, 2000) for each sample, and a separate script on CLAN for the Moving Average Token-Type Ratio (MATTR) (Covington, 2007). MATTR allows for comparisons of varying sample lengths; it also 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 (MacWhinney, 2000), was between 0 and 142 seconds (T1) for the first assessment, from 0 to 132 seconds for the second assessment (T2), and from 10 to 242 seconds for the third assessment. Utterance segmentation, transcription and scoring for utterances, and identification of lexical errors were conducted as per the Codes for the Humans Analysis of Transcripts (CHAT) manualized protocol (MacWhinney, 2000), incorporating additional guidance for French users of this program (Colin & Le Meur, 2016). Extracted productivity measures were defined as the total number of completed words and the number of words per minute. Grammatical/syntactic complexity was measured by the mean length of utterance (MLU), the number of verbs/utterance and propositional density (Brown et al., 2008). Also, percentages of phonological and semantic errors, as a function of total number of content words, were computed.

Phonological errors included phonemic and phonetic paraphasias, false starts, *conduites*

d'approche, and neologisms. Semantic errors included verbal paraphasias (i.e., replacement of a target word by a semantically related or unrelated word). Finally, the percentage of adequate utterances was calculated, based on the number of utterances produced without any errors. Each variable was computed in acute, subacute and chronic phases.

Inter-rater reliability

Inter-rater reliability (IRR) was conducted on the transcriptions of six participants (33% of the transcripts) at all three time points (n=18 transcriptions) by a second rater (M.D.-B.). Two-way random effects intraclass correlation coefficients (ICC) were calculated on micro-structural variables extracted from CLAN (i.e., the total number of words, the duration of the sample, and the number of utterances) and TUs to determine consistency between raters (as in Brisebois et al., 2020; Marcotte et al., 2017). The results of these analyses are reported in Table 2. All variables met the threshold of high reliability, $ICC > .80$ (Streiner & Norman, 2008).

[Table 2 should be inserted here]

Table 2. Inter-rater reliability express as two-way random effect intraclass correlations (ICC)

	T1 Acute assessment Cronbach's alphas (α)	T2 Subacute assessment Cronbach's alpha (α)	T3 Chronic assessment Cronbach's alphas (α)
Total number of words	.986	.954	1.0
Duration of the sample	.971	.869	.958
Number of utterances	.904	.974	.969
Thematic Units	.985	.968	1.0

TUs were independently manually scored by two of the authors (A.B. and M.D.-B.) for all discourse samples.

Data analysis

All statistical analyses were done using SPSS® v25.0, with the significance level set at $p < .05$ after Bonferroni correction for multiple comparisons.

Time effect

We first aim to measure discourse production changes in early and late post-stroke aphasia by analyzing micro-structural variables and thematic informativeness. Thus, separate analyses were conducted on the dependent variables of thematic informativeness, the micro-structural variables noted earlier (e.g., verbs per utterance), and the language composite score (CS), with time as a repeated measure. The variables of thematic informativeness (TUS, TUs per minute and TUs per utterance), moving average token-type ratio (MATTR), verbs per utterance, density, percentage of semantic and

phonological paraphasias, percentage of adequate utterances, and the CS showed a non-normal distribution (a Shapiro Wilk normality test being $p < .05$ for these variables). As such, a non-parametric Friedman test with Bonferroni post-hoc for paired comparisons was conducted on these variables. The variables of number of total words, number of words per minute, and mean length of utterances showed a normal distribution and so a repeated measure mixed ANOVA, using Bonferroni correction for the post-hoc comparisons, was conducted for these variables.

Exploratory Correlations

The secondary aim of this study is to explore associations between the discourse measures in the acute period and a global language severity measure obtained in both the acute and chronic stages. Hence, Pearson bivariate correlations were computed to assess the possible association between initial measures of discourse that significantly changed over time (i.e., TU_{acute} , $TU_{/min}_{acute}$, $TU_{/utt}_{acute}$, MLU_{acute} , and number of words per minute_{acute}) and the general composite score of aphasia in both acute and chronic stages (CS_{acute} and $CS_{chronic}$).

Results

Time effect

Thematic informativeness variables

Group data

A non-parametric Friedman test demonstrated a significant effect of time on Thematic Units (TUs) $\chi^2(2) = 17.100$; $p < .001$; Kendall's $W = .503$. Bonferroni post-hoc for matched comparisons showed a significant mean improvement for TU_{chronic} compared to TU_{acute} of 4.44 TUs ($p < .001$). This test also showed a significant effect of time on TUs per minute $\chi^2(2) = 6.918$; $p < .05$; Kendall's $W = .216$. Post-hoc comparisons indicated a significant mean improvement in $TUs/\text{min}_{\text{chronic}}$ compared to $TUs/\text{min}_{\text{acute}}$ of 5.86 TUs/min ($p = .031$). Similar results were obtained with TUs/utt. Results revealed a significant effect of time $\chi^2(2) = 12.689$; $p < .01$; Kendall's $W = .397$. Bonferroni post-hoc for paired comparisons demonstrated a significant improvement of $TUs/\text{utt}_{\text{chronic}}$ compared to $TUs/\text{utt}_{\text{acute}}$ of 0.24 TU/utt ($p = .002$). The results of these analyses are reported in Table 3.

[Table 3 should be inserted here]

Table 3. Mean scores (SD) of discourse variables at each assessment timepoint and the effect of time.

				Repeated measure mixed ANOVA		
				*Non parametrical Friedman test		
				Time effect		
Variables	T1 (acute)	T2 (subacute)	T3 (chronic)	T1-T2	T2-T3	T1-T3
	MEAN (SD)	MEAN (SD)	MEAN (SD)			
Thematic units*	5.62 (5.40)	8.29 (5.72)	10.06 (4.75)	$\chi^2 (2) = 17.100, p < .001, \text{Kendall's } W = .503$		
				p=.062	p=.368	p< .001
Thematic units per minute*	6.07 (5.53)	11.38 (10.16)	11.93 (7.35)	$\chi^2 (2) = 8.400, p < .02, \text{Kendall's } W = .247$		
				p=.368	p=.595	p<.02
Thematic units per utterance*	0.32 (0.37)	0.60 (0.49)	0.56 (0.25)	$\chi^2 (2) = 14.431, p < .001, \text{Kendall's } W = .424$		
				p=.077	p=.435	p<.001

				Repeated measure mixed ANOVA		
				*Non parametrical Friedman test		
				Time effect		
Variables	T1 (acute)	T2 (subacute)	T3 (chronic)	T1-T2	T2-T3	T1-T3
	MEAN (SD)	MEAN (SD)	MEAN (SD)			
Total words	114.67 (88.916)	132.13 (116.30)	178.29 (139.77)	F (2, 13) =2.507, p=.120 partial η^2 =.278		
Words per minute	117.67 (38.83)	119.68 (41.27)	147.21 (45.74)	F (2, 13) =5.941, p<.015 partial η^2 =.478		
				p=1.0	p=.080	p<.02
	5.36 (2.84)	6.65 (4.16)	7.49 (3.28)	F (2, 13) = 3.831, p<.05 partial η^2 =.371		

				Repeated measure mixed ANOVA		
				*Non parametrical Friedman test		
				Time effect		
Variables	T1 (acute)	T2 (subacute)	T3 (chronic)	T1-T2	T2-T3	T1-T3
	MEAN (SD)	MEAN (SD)	MEAN (SD)			
Mean length of utterances (words)				p=.435	p=.648	p<.05
Moving average Token Type ratio (MATTR)*	0.82 (0.19)	0.90 (0.06)	0.92 (0.04)	$\chi^2 (2) = 1.385, p=.500$		
Verb per utterance*	0.26 (0.27)	0.32 (0.40)	0.32 (0.31)	$\chi^2 (2) = 0.552, p=.759$		
Density*	0.23 (0.11)	0.25 (0.13)	0.30 (0.08)	$\chi^2 (2) = 3.3733, p=.155$		

				Repeated measure mixed ANOVA		
				*Non parametrical Friedman test		
				Time effect		
Variables	T1 (acute)	T2 (subacute)	T3 (chronic)	T1-T2	T2-T3	T1-T3
	MEAN (SD)	MEAN (SD)	MEAN (SD)			
% semantic paraphasia*	1.27 (2.02)	0.48 (0.59)	0.64 (0.85)	$\chi^2 (2) = 2.462, p=.292$		
% phonological paraphasia*	1.83 (2.93)	0.42 (0.34)	1.04 (1.37)	$\chi^2 (2) = 2.087, p=.352$		
% adequate utterance*	65.49 (40.05)	71.28 (31.47)	82.62 (28.44)	$\chi^2 (2) = 6.582, p<.05, Kendall's W = .219$		
				p=.1.0	p=.249	p=.053

Individual data

In terms of individual results, TUs scores showed a positive change from the acute testing phase to the chronic testing phase for 15 participants; one patient's score dropped by one point and another clearly showed a ceiling effect. If we consider results for TUs/min, the individual data revealed improvement (i.e., a positive change) from the acute stage to the chronic stage for 13 patients; the other patients showed a decrease of TUs/min. Also, for TUs/utt, 15 participants demonstrated a positive change during this period; another showed stability over time and another a decrease in TUs/utt. Individual results are presented in Figure 1.

[Figure 1 should be inserted here]

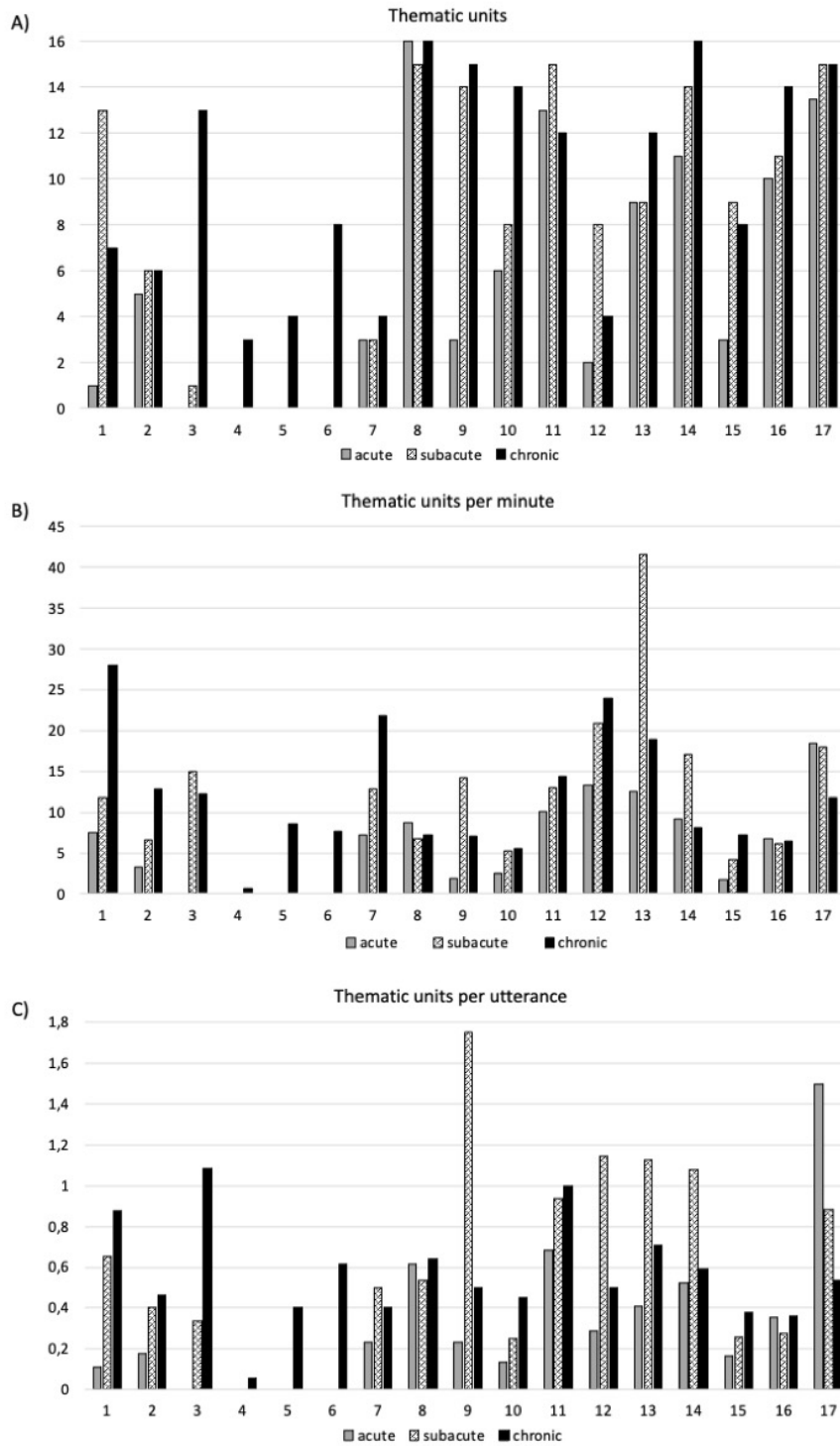


Figure 1. Individual results in acute (48-72h), subacute (7-14 days) and chronic (>180 days) stages post stroke for the macro-linguistic measures that significantly changed over time. The upper panel (A) represents the Thematic Units, the middle panel (B) represents the Thematic Units per minute, and the lower panel (C) the Thematic Units per utterance.

Micro-structural variables

Group data

Repeated measures mixed ANOVAs demonstrated a significant effect of time on the number of words per minute $F(2, 13) = 5.941; p < .02$; partial $\eta^2 = .478$. The results of these analyses are reported in Table 3. A Bonferroni post-hoc test for matched comparisons showed a significant mean improvement in number of words per minute_{chronic} compared to number of words per minute_{acute} of 29.54 words per minute ($p < .009$). A significant effect of time on mean length of utterance (MLU) $F(2, 13) = 3.831; p < .049$; partial $\eta^2 = .371$ was also found, and post-hoc comparisons demonstrated a significant mean improvement in MLU_{chronic} as compared to MLU_{acute} of 2.13 MLU (words) ($p < .039$). A non-parametric Friedman test demonstrated a significant effect of time on the percentage of adequate utterances produced, $\chi^2(2) = 6.582; p < .037$; Kendall's $W = .219$. However, the Bonferroni post-hoc test for matched comparisons revealed no significant differences between the acute and chronic phases. All the other micro-structural variables did not show any significant effect of time.

Individual data

Looking at individual data from the acute to chronic phases, 14 participants showed an increase in the number of words per minute; whereas, the others demonstrated a decline ($n=3$). With respect to mean length of utterance, 13 participants showed an improvement; whereas the other patients demonstrated either stability (i.e., no change between the acute and chronic stages, $n=2$) or a decline in MLU ($n=2$). These data are illustrated in Figure

2.

[Figure 2 should be inserted here]

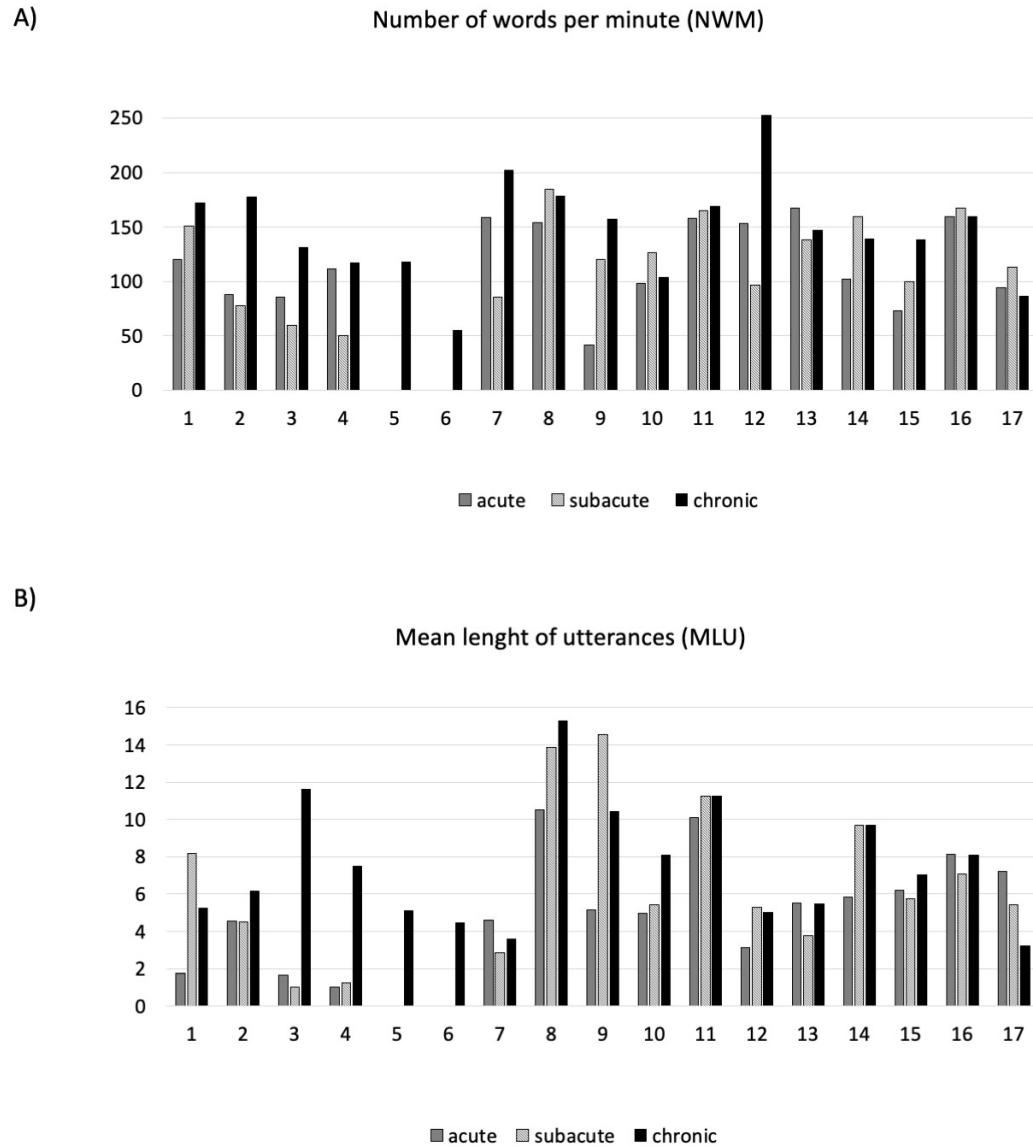


Figure 2. Individual results in acute (48-72h), subacute (7-14 days) and chronic (>180 days) stages post stroke for the micro-linguistic measures that significantly changed over time. The upper panel (A) represents the number of words per minute and the lower panel (B) the mean length of utterances.

Language composite scores

A non-parametric Friedman test demonstrated a significant effect of time on language Composite Score (CS) $\chi^2(2) = 15.134; p < .001$; Kendall's $W = .445$. The Bonferroni post-hoc test for matched comparisons showed a significant mean improvement for $CS_{Chronic}$ as compared to CS_{acute} of 9.93 points ($p < .001$). Considering individual data from the acute to chronic phases, CS improved for 15 patients; whereas the 2 others demonstrated a slight decrease in the CS. Figure 3 represents individual results for the language composite score.

[Figure 3 should be inserted here]

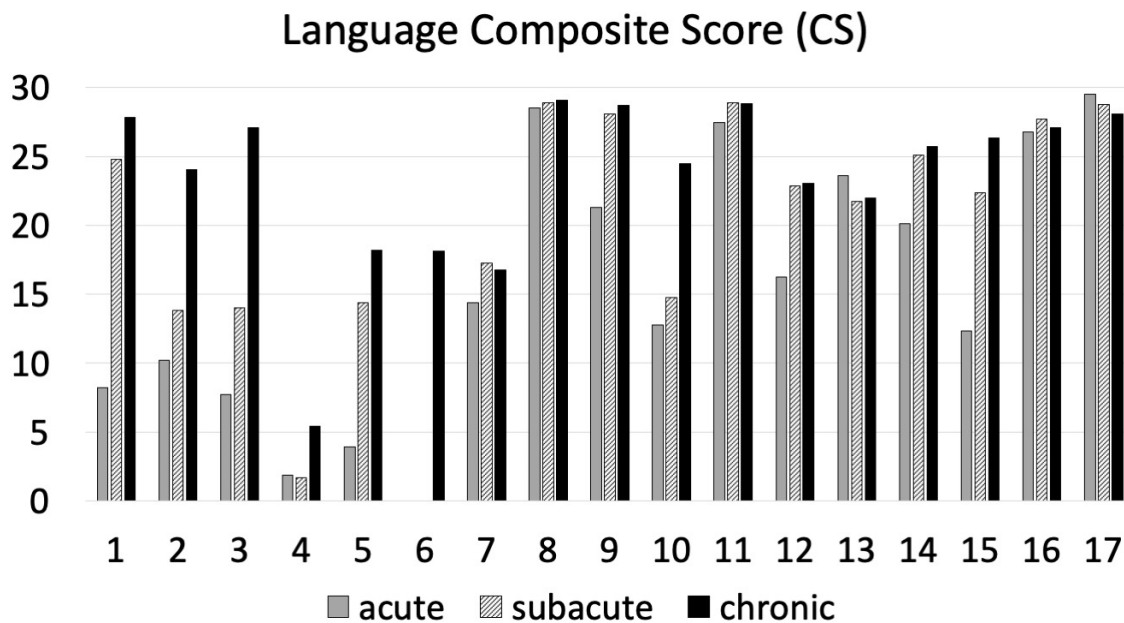


Figure 3. Individual results for the language Composite Score in acute (48-72h), subacute (7-14 days) and chronic (>180 days) stages post stroke.

Correlations

Pearson bivariate correlations were conducted to assess the possible association between initial scores of discourse production that significantly changed over time (i.e., Thematic Units (TUs), TUs per minute, TUs per utterance, mean length of utterance (MLU), and number of words per minute) and a general composite score of language impairment (CS) in the acute and chronic stages. Significant correlations were found and are reported in Table 4. Namely, CS_{chronic} was found to be positively correlated with $TU_{\text{Sacute}} r(17) = .523, p=.031$, $MLU_{\text{acute}} r(17) = .530, p=.042$, and $CS_{\text{acute}} r(17) = .631, p=.007$. These results indicate a positive association between the acute and chronic scores for the named variables. Also, CS_{acute} was found to be positively associated with $TU_{\text{Sacute}} r(17) = .888, p< .001$, $TUs/min_{\text{acute}} r(17) = .745, p=.001$, $TUs/utt_{\text{acute}} r(17) = .790, p< .001$, and $MLU_{\text{acute}} r(17) = .869, p< .001$. These results demonstrate a positive association between the initial language score (CS_{acute}) and the named discourse variables in the acute stage of recovery (please see Table 4). Figure 4 illustrates the significant correlations between TU_{Sacute} and CS_{acute} and CS_{chronic} .

[Table 4 should be inserted here]

Table 4. Correlations between discourse and language variables

Variables		TU _{sacute}	TU _{s/minacute}	TU _{s/uttacute}	Words/min _{acute}	MLU _{acute}	CS _{acute}	CS _{chronic}
Variables	Statistics	I	II	III	IV	V	VI	VII
I	r	1	.672	.804	.367	.894	.888	.523
	p		0.003	0.000	0.178	0.000	0.000	0.031
	n	17	17	17	15	15	17	17
II	r		1	.831	.508	.412	.745	.373
	p			0.000	0.053	0.127	0.001	0.140
	n		17	17	15	15	17	17
III	r			1	.143	.616	.790	.442
	p				.612	0.014	0.000	0.075
	n			17	15	15	17	17
IV	r				1	.315	.350	-.145
	p					0.253	0.201	0.606

Variables		TU _{sacute}	TU _{s/minacute}	TU _{s/uttacute}	Words/min _{acute}	MLU _{acute}	CS _{acute}	CS _{chronic}
Variables	Statistics	I	II	III	IV	V	VI	VII
	n				15	15	15	15
V	r					1	.869	.530
	p						0,000	0.042
	n					15	15	15
VI	r						1	.631
	p							0,007
	n						17	17

[Figure 4 should be inserted here]

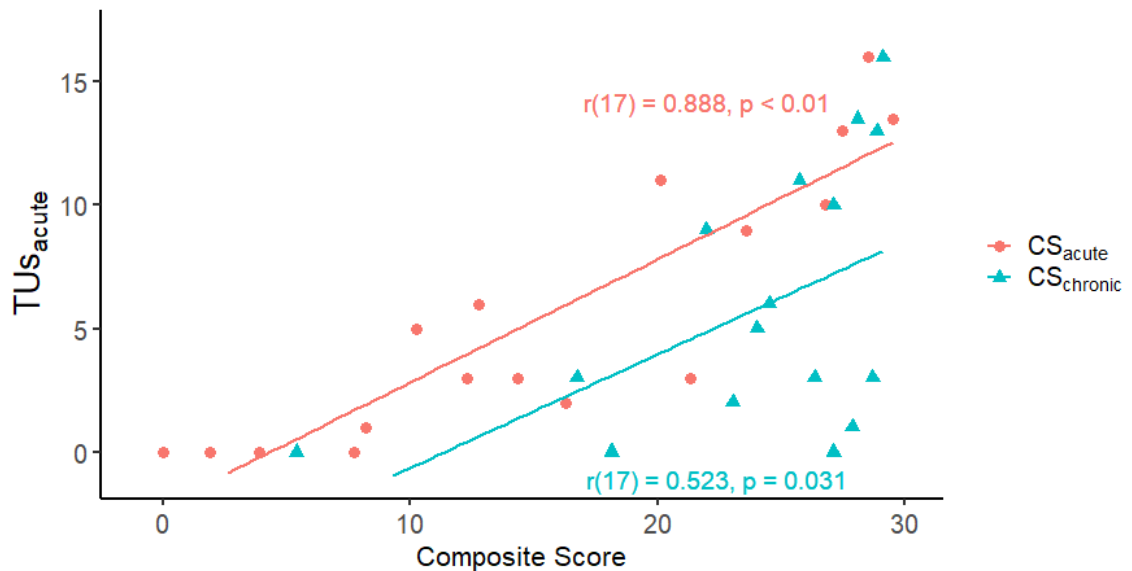


Figure 4. Pearson correlations' scatterplots between TUS_{acute} and CS_{acute} (circles in red) and TUS_{acute} and CS_{chronic} (triangles in blue).

Discussion

Our findings suggest that the production of thematic units in descriptive discourse during an acute clinical examination of language is a promising measure because it seems to capture long-term changes in discourse abilities and seems to be related to an overall language measure in both the acute and chronic stages of recovery. The interpretation of the changes in mean length of utterances (MLU) and maybe also in number of words per minute is less straightforward, as improvements in these measures carry different meanings depending on the type of aphasia (Audrey Holland et al., 2017). Granted the test-retest reliability of this measure requires further investigation with larger sample

sizes, we propose that thematic informativeness may be considered in the routine clinical assessment of individuals with aphasia to track longitudinal changes in post-stroke aphasia.

The results showed significant positive changes between the acute and chronic phases for the micro-structural variables—more specifically, for the mean length of utterances and the number of words per minute. For the thematic informativeness measures, the number of thematic units (TUs), the number of thematic units per minute (TUs/min) and the number of thematic units per utterance (TUs/utt) increased significantly between the acute and chronic phases, but not between the acute and subacute phases. Significant positive moderate correlations between TUs and MLU in the acute phase and a general language measure in the acute and chronic phases suggest a relationship between discourse measures and an overall language measure and, potentially, the predictive value of these variables in the acute phase.

The present study aimed to measure descriptive discourse production changes in early and late post-stroke aphasia recovery by analyzing micro-structural variables and thematic informativeness. As we hypothesized, based on our previous study (Brisebois et al., 2020), some micro-structural variables showed an increase between the acute and chronic stages of recovery. Namely, the number of words per minute and MLU showed significant improvements between the acute and chronic stages, while no other micro-structural variable changed significantly in the same period. Interestingly, among the variables previously identified as having the most diagnostic sensitivity (Andreetta, Cantagallo, & Marini, 2012; Andreetta & Marini, 2015; Covington & McFall, 2010; Saffran, Berndt, & Schwartz, 1989; Wright, Capilouto, Wagovich, Cranfill, & Davis,

2005), only speech rate, measured by the number of words per minute, increased significantly from the acute to the chronic stage of recovery in our study. As highlighted by Holland et al. (Audrey Holland et al., 2017), increases in some discourse measures—namely, MLU—do not always represent improvement. More precisely, these authors mention that a higher MLU in a person with Wernicke’s aphasia might not reflect an improvement in language function. We believe that the same logic applies to the number of words per minute. Therefore, it is possible that the positive changes in MLU and number of words per minute reflect changes, but not necessarily improvement in terms of language function.

Our results extend previous findings that documented improvement in thematic informativeness in the early stages of recovery (Brisebois et al., 2020; Furlanis et al., 2018). Surprisingly, the present cohort did not demonstrate significant differences between the acute and subacute stages in the thematic informativeness measures, which, rather, resulted in statistical trends. The discrepancy between the present results and those previously obtained by our team (Brisebois et al., 2020) could be explained by a smaller sample size due to the longitudinal aspect of the present study. In our previous study, we reported on 23 PWA, but only 17 PWA participated in the third assessment, therefore representing 26% of group attrition. However, the combination of findings from this previous study and the present one support further investigation and a growing interest in considering informativeness measures in the investigation of the longitudinal recovery of aphasia, including their use in clinical settings.

Interestingly, this group of participants well represents the cosmopolitan and linguistic diversity of Montreal. While French is the official language of the city,

according to the latest census profile (Statistics Canada, 2017), nearly 60% of the population has a knowledge of both Canadian official languages (i.e., English and French), and more than 21% of Montrealers are trilingual. Our sample was composed of bilinguals (N=10), trilinguals (N=3) and monolinguals (N=4) (a minority population in Montreal). Nevertheless, they all had French-Canadian as their first language, and all were tested in this language. We believe that by including not only monolinguals, we have a more representative sample of the population in our city and even our province. When looking at individual data, no specific trend seemed to distinguish the linguistic groups from one another. However, the present group's diversity might influence obtained results and one must be careful when comparing with other studies that are mostly based on monolingual populations, particularly when looking at micro-structural results, which are highly language dependent. For instance, a tendency of French-Canadian speakers to accept English equivalents or a French pronunciation of an English word has been observed in a study that examined name agreement for the 60 stimuli of the Boston Naming Test (Kaplan et al., 1983) in Quebec (Roberts & Doucet, 2011).

Conversely, it could be argued that a measure such as thematic informativeness transcends language barriers because it represents the ability to express semantic content related to a specific stimulus (Brisebois et al., 2020) but not a specific lexical unit. In fact, during a picture description task, the richness of the verbal output is determined by the ability to interpret the stimulus (i.e., the content of the picture) and the capacity to formulate the ideas that the stimulus depicts (Kong, 2016). Therefore, TU improvements represent both the recovery of linguistic skills and the ability to interpret the stimulus. Indeed, similar results in Italian (Furlanis et al., 2018) support the idea that thematic

informativeness reflects the prelinguistic conceptual stage of discourse production as illustrated in Sherrat's model (2007).

Discourse assessment is a very common practice in clinical settings (Bryant et al., 2017), but an important gap between research and practice in discourse analysis has been noted (Dietz & Boyle, 2018), including in the French-Canadian community, where normative data is very scarce. The results of this study may lead to the development of a clinical tool incorporating a list of thematic units, that is responsive to change in post-stroke aphasia. The present results support future investigations to establish test-retest reliability in a larger sample size and also to evaluate the clinical validity of this measure.

Our second aim was to explore the association between discourse variables in the acute stage and overall language scores in the acute and chronic stages. Our results indicated that, amongst variables that significantly improved over time, Thematic Units (TU_{acute}), the mean length of utterances (MLU_{acute}), and the language composite score (CS_{acute}) in the acute stage were associated with the language composite score in the chronic stage (CS_{chronic}). Not surprisingly, the strongest correlation was between CS_{acute} and CS_{chronic} , which is clearly in line with the idea that initial severity of the language impairment is associated with long-term language outcomes (Lazar et al., 2010). Positive moderate correlations were also found between discourse variables (i.e., TU_{acute} and MLU_{acute}) and the general language measure (CS) in the chronic phase, which suggests a relationship between these initial discourse measures and language scores obtained in the chronic stage of language recovery. Therefore, TU_{acute} should be taken into account to improve predictions about the long-term language performance of individuals with acute aphasia. Also, strong correlations were found between TU_{acute} , TU_{minacute} , TU_{uttacute} ,

MLU_{acute}, and CS_{acute}, which suggests that discourse scores are at least partially connected to language scores in acute post-stroke aphasia. If we add to these interpretations the fact that thematic informativeness seems to transcend languages (i.e., similar results have been obtained in Italian (Furlanis et al., 2018) and English (Agis et al., 2016)), our hypothesis that thematic units can be a general language assessment tool is strengthened (Brisebois et al., 2020). More research is needed to integrate this knowledge into clinical practice. However, our data support the importance of discourse assessment, as it seems to be a good indicator of overall language performance, including both expressive and receptive components. On a more practical note, we believe that thematic informativeness, similar to Content Units (CUs) (Agis et al., 2016), constitutes an interesting path to explore as a routine clinical evaluation because it is time-efficient, simple to conduct, and reliable in assessing initial impairment and changes in the discourse production of individuals with aphasia throughout the recovery continuum.

Nonetheless, our results must be interpreted with caution. Indeed, the extent of our results is limited by the small number of patients who were able to maintain their participation throughout the year of data collection. Current knowledge suggests that administering more than one discourse task is mandatory to establish test-retest reliability (Boyle, 2014) and also to address a comprehensive language picture of PWA (Stark, 2019). In fact, the present study includes only one sample at each time point, which limits database size and therefore a sound stability analysis. However, the lexical informativeness measure (i.e., correct information units and the number of correct information units per minute) and the number of words per minute have demonstrated sufficient test-retest correlations (<.85) (Boyle, 2014) to support their suitability for

group research studies (Fitzpatrick et al., 1998) at least in the chronic stage of recovery. Besides, many studies have used a single-picture description task to investigate content variables in groups of PWA (e.g., Agis et al., 2016) and a wide variety of metrics have been studied using an oral picture description task (Bryant et al., 2016). Future studies should include many discourse genres to incorporate factors such as listener familiarity, topic, and emotionality. Neuropsychological approaches tend to focus on the impairment, whereas functional language assessment examines how people with aphasia successfully convey meaning. Hence, clinical assessments should also include different discourse genres to gain a full picture of the person's everyday communication skills (Papathanasiou & Coppens, 2012). Granted, the present results were obtained with only one type of discourse and elicitation technique; importantly, however, this study is the first to describe discourse recovery in French-Canadian speakers from the acute to the chronic stage of recovery. Reference or normative data on discourse measures are scarce in French-Canadian, which limits diagnostic use (Monetta et al., 2014). Therefore, there is a clear need to provide normative data to target spontaneous and therapy-induced improvements and, hopefully, establish a better understanding of discourse informativeness measures and document their predictive value.

Another drawback of this study is the limited knowledge concerning the effect of the different variables related to therapy, namely timing, type, duration and intensity, provided between the subacute and the chronic data collection points. The number of patients included in the present study did not allow us to include the multiple dimensions of therapy in our statistical analysis. Nevertheless, the present study took place in Canada which has a public health care system. This means that all patients received speech and

language therapy based on their individual needs and their personal objectives in therapy, which is different than in other countries. Participants with the mildest impairments received only a few sessions, whereas participants with severe impairments received extended and intensive therapy. Also, most longitudinal studies that did not specifically investigated the effect of therapy have not investigated the effect of treatment on longitudinal changes (e.g., Hillis et al., 2018; Stockbridge et al., 2019). However, there is still a clear need to further investigate the effect of the multiple dimensions of therapy on longitudinal changes in post-stroke aphasia, not only for changes observed in discourse tasks as in the present study, but on all language abilities. Furthermore, discourse performance relies on a complex interaction of linguistic and cognitive processes (Sherratt & Bryan, 2012) which should be considered when one is seeking to draw a picture of discourse performance in post-stroke patients.

Conclusion

This study aimed to measure changes in descriptive discourse production from the acute to chronic stages of post-stroke aphasia recovery in a group of 17 PWA following a first left middle cerebral artery stroke. Analyses demonstrated a significant effect of time (i.e., improvement) for the number of total thematic units and of thematic units per minute (TUs/min) as well as mean length of utterance (MLU) and number of words per minute from the acute stage to the chronic stage of recovery. Correlation analyses demonstrated that amongst the discourse variables that significantly improved over time were TUs and MLU in the acute phase. These preliminary correlation results suggest that thematic

informativeness obtained from a simple descriptive discourse task in the acute phase encourage research to demonstrate whether they could potentially be a useful measure for predicting aphasia outcome in the chronic phase. These findings also support the use of TUs in assessing discourse performance in PWA. Further longitudinal studies, which include more participants and also more discourse tasks should be conducted to document the longitudinal course of discourse recovery in post-stroke aphasia.

Acknowledgments

We are very grateful to all the participants and their families for their contribution to this study. Also, we wish to thank Alberto Osa García and Anne-Marie Chouinard for their help in data processing and preparation of this manuscript.

Funding

This project was funded by a grant-in-aid from the 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". A.B. holds a scholarship from the "Fonds de Recherche du Québec – Santé" and J.B. from the Canadian Institute on Health Research (CIHR).

References

- Agis, D., Goggins, M. B., Oishi, K., Oishi, K., Davis, C., Wright, A., Kim, E. H., Sebastian, R., Tippett, D. C., Faria, A., & Hillis, A. E. (2016). Picturing the Size and Site of Stroke With an Expanded National Institutes of Health Stroke Scale. *Stroke*, 47(6), 1459–1465. <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>
- Alyahya, R. S. W., Halai, A. D., Conroy, P., & Lambon Ralph, M. A. (2020). A unified model of post-stroke language deficits including discourse production and their neural correlates. *Brain*, 143(5), 1541–1554. <https://doi.org/10.1093/brain/awaa074>
- 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, Sara, Cantagallo, A., & Marini, A. (2012). Narrative discourse in anomic aphasia. *Neuropsychologia*, 50(8), 1787–1793.
- 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.
- Armstrong, E., Bryant, L., Ferguson, A., & Simmons-Mackie, N. (2012). Approaches to Assessment and Treatment of Everyday Talk in Aphasia. In Ilias Papathanasiou & P. Coppens (Eds.), *Aphasia and related neurogenic communication disorders*. Papathanasiou, I., Coppens, P., Potagas, C. (2nd ed., pp. 269–285). Jones & Barlett learning.
- Avent, J. R., & Austerlitz, S. (2003). Reciprocal scaffolding: A context for communication treatment in aphasia. *Aphasiology*, 17(4), 397–404. <https://doi.org/10.1080/02687030244000743>
- Bernhardt, J., Hayward, K. S., Kwakkel, G., Ward, N. S., Wolf, S. L., Borschmann, K., Krakauer, J. W., Boyd, L. A., Carmichael, S. T., Corbett, D., & Cramer, S. C. (2017). Agreed definitions and a shared vision for new standards in stroke recovery research: The Stroke Recovery and Rehabilitation Roundtable taskforce. *International Journal of Stroke*, 12(5), 444–450. <https://doi.org/10.1177/1747493017711816>
- 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. [https://doi.org/10.1016/S0021-9924\(98\)00005-7](https://doi.org/10.1016/S0021-9924(98)00005-7)
- Boulanger, J., Lindsay, M., Gubitza, G., Smith, E., Stotts, G., Foley, N., Bhogal, S., Boyle, K., Braun, L., Goddard, T., Heran, M., Kanya-Forster, N., Lang, E., Lavoie, P., McClelland, M., O’Kelly, C., Pageau, P., Pettersen, J., Purvis, H., ... 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>
- Boyle, M. (2014). Test-retest stability of word retrieval in aphasic discourse. *Journal of*

- Speech, Language, and Hearing Research*, 57(3), 966–978.
https://doi.org/10.1044/2014_JSLHR-L-13-0171
- Boyle, M. (2015). Stability of word-retrieval errors with the Aphasiabank stimuli. *American Journal of Speech-Language Pathology*, 24(4), 5953–5960.
https://doi.org/10.1044/2015_AJSLP-14-0152
- Boyle, M. (2020). Choosing Discourse Outcome Measures to Assess Clinical Change. *Seminars in Speech and Language*, 41(1), 1–9. <https://doi.org/10.1055/s-0039-3401029>
- Brisebois, A., Brambati, S. M., Désilets-Barnabé, M., Boucher, J., García, A. O., Rochon, E., Leonard, C., Desautels, A., & Marcotte, K. (2020). The importance of thematic informativeness in narrative discourse recovery in acute post-stroke aphasia. *Aphasiology*, 34(4), 472–491. <https://doi.org/10.1080/02687038.2019.1705661>
- 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.
<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.
- 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., & Maddy, K. M. (2016). Microlinguistic processes that contribute to the ability to relay main events: influence of age. *Aging, Neuropsychology, and Cognition*, 23(4), 445–463.
<https://doi.org/10.1080/13825585.2015.1118006>
- 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. *Arch Phys Med Rehabil*, 82(8), 1073–1080.
<https://doi.org/10.1053/apmr.2001.25155>
- Cherney, L. R. (1998). *Analyzing discourse in communicatively impaired adults* (C. A. Coelho & B. B. (Barbara B. Shadden (eds.)). Gaithersburg, Md. : Aspen Publishers.
- Coelho, C. A., Mchugh, R. E., & Boyle, M. (2000). Semantic feature analysis as a treatment for aphasic dysnomia: A replication. *Aphasiology*, 14(2), 133–142.
<https://doi.org/10.1080/026870300401513>
- Colin, C., & Le Meur, C. (2016). *Adaptation du projet AphasiaBanks 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)*.
- 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>
- Cupit, J., Rochon, E., Leonard, C., & Laird, L. (2010). Social validation as a measure of improvement after aphasia treatment: Its usefulness and influencing factors. *Aphasiology*, 24(11), 1486–1500. <https://doi.org/10.1080/02687031003615235>
- 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. <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 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., Carroll, B., Lustig, A., & 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>
- Ellis, C., Rosenbek, J. C., Rittman, M. R., & Boylstein, C. A. (2005). Recovery of cohesion in narrative discourse after left-hemisphere stroke. *Journal of Rehabilitation Research and Development*, 42(6), 737–746. <http://www.ncbi.nlm.nih.gov/pubmed/16680611>
- Fergadiotis, G., & Wright, H. H. (2015). Modelling confrontation naming and discourse performance in aphasia. *Aphasiology*, 30(4), 364–380. <https://doi.org/10.1080/02687038.2015.1067288>
- Fitzpatrick, R., Davey, C., Buxton, M. J., & Jones, D. R. (1998). Evaluating patient-based outcome measures for use in clinical trials. *Health Technology Assessment*, 2(14). <https://doi.org/10.3310/hta2140>
- Fromm, D., Greenhouse, J., Hou, K., Russell, G. A., Cai, X., Forbes, M., Holland, A., & 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-l-15-0401
- Furlanis, G., Ridolfi, M., Polverino, P., Menichelli, A., Caruso, P., Naccarato, M., Sartori, A., Torelli, L., Pesavento, V., & Manganotti, P. (2018). Early Recovery of Aphasia through Thrombolysis: The Significance of Spontaneous Speech. *J Stroke Cerebrovasc Dis*, 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., Goodglass, H., Kaplan, E., & Kaplan, E. (2001). *The Boston Diagnostic Aphasia Examination : BDAE-3 long form kit*.
- Hillis, A. E., Beh, Y. Y., Sebastian, R., Breining, B., Tippett, D. C., Wright, A., Saxena, S., Rorden, C., Bonilha, L., Basilakos, A., Yourganov, G., & Fridriksson, J. (2018). Predicting recovery in acute poststroke aphasia. *Ann Neurol*, 83(3), 612–622. <https://doi.org/10.1002/ana.25184>
- Holland, A., Frattali, C., & Fromm, D. (1999). *CADL-2 : Communication Activities of Daily Living*. Pro-Ed.

- Holland, A., Fromm, D., Forbes, M., & MacWhinney, B. (2017). Long-term Recovery in Stroke Accompanied by Aphasia: A Reconsideration. *Aphasiology*, *31*(2), 152–165. <https://doi.org/10.1080/02687038.2016.1184221>
- Joanette, Y., Ska, B., & Côté, H. (2004). *Protocole Montréal d'Évaluation de la Communication*. Ortho Édition.
- Kaplan, E., Goodglass, H., & Weintraub, S. (1983). *Boston Naming Test*. Lea & Febiger.
- Kertesz, A. (2006). *Western Aphasia Battery- Revised*. Pearson.
- Kiran, S., & Thompson, C. K. (2019). Neuroplasticity of Language Networks in Aphasia: Advances, Updates, and Future Challenges. *Frontiers in Neurology*, *10*, 295. <https://doi.org/10.3389/fneur.2019.00295>
- Kong, A. P. H. (2016). *Analysis of Neurogenic Disordered Discourse Production*. Routledge. <https://doi.org/10.4324/9781315639376>
- 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*, *11*(8), 783–811. <https://doi.org/10.1080/02687039708250456>
- 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*. Lawrence Erlbaum Associates.
- MacWhinney, Brian, & Fromm, D. (2016). AphasiaBank as BigData. *Seminars in Speech and Language*, *37*(1), 10–22.
- 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), 2217–2227. https://doi.org/10.1044/2017_JSLHR-L-17-0012
- 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., Andreetta, S., del Tin, S., Carlomagno, S., Tin, S. del, & Carlomagno, S. (2011). A multi-level approach to the analysis of narrative language in aphasia. *Aphasiology*, *25*(11), 1372–1392. <https://doi.org/10.1080/02687038.2011.584690>
- Marini, Andrea, Caltagirone, C., Pasqualetti, P., & Carlomagno, S. (2007). Patterns of language improvement in adults with non-chronic non-fluent aphasia after specific therapies. *Aphasiology*, *21*(2), 164–186. <https://doi.org/10.1080/02687030600633799>
- Marini, Andrea, 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. <https://doi.org/10.1016/j.bandl.2004.08.002>
- Monetta, L., & Tous les membres de l'axe 4 du REPAR. (2014). *Fiche descriptive des outils validés et/ou normés en franco-québécois pour l'évaluation du langage et de*

la parole, 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*.
- 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. <http://www.ncbi.nlm.nih.gov/pubmed/7731205>
- 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. <https://jslhr.pubs.asha.org/article.aspx?articleid=1779418>
- Osa García, A., Brambati, S. M., Brisebois, A., Désilets-Barnabé, M., Bedetti, C., Rochon, E., Leonard, C., Desautels, A., & Marcotte, K. (2020). Predicting early post-stroke aphasia outcome from initial aphasia severity. *Frontiers in Neurology, 11*, 120. <https://doi.org/10.3389/FNEUR.2020.00120>
- Papathanasiou, I., & Coppens, P. (2012). Aphasia and related neurogenic communication disorders: basic concepts and operational definitions. In *Aphasia and related neurogenic communication disorders*. Papathanasiou, I., Coppens, P., Potagas, C. Jones & Barlett learning.
- Pedersen, P. M., Vinter, K., & Olsen, T. S. (2004). Aphasia after stroke: type, severity and prognosis. The Copenhagen aphasia study. *Cerebrovascular Diseases (Basel, Switzerland), 17*(1), 35–43. <https://doi.org/10.1159/000073896>
- 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>
- Roberts, & Doucet. (2011). Performance of French-speaking Quebec adults on the Boston Naming Test. *Canadian Journal of Speech-Language Pathology and Audiology, 35*, 254–267.
- Ross, K. B. (1999). Comparison of impairment and disability measures for assessing severity of, and improvement in, aphasia. *Aphasiology, 13*(2), 113–124. <https://doi.org/10.1080/026870399402235>
- Saa, J. P., Tse, T., Baum, C., Cumming, T., Josman, N., Rose, M., & Carey, L. (2019). Longitudinal evaluation of cognition after stroke – A systematic scoping review. *PLOS ONE, 14*(8), e0221735. <https://doi.org/10.1371/journal.pone.0221735>
- 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. <http://www.ncbi.nlm.nih.gov/pubmed/2804622>
- Saur, D., Lange, R., Baumgaertner, A., Schraknepper, V., Willmes, K., Rijntjes, M., & Weiller, C. (2006). Dynamics of language reorganization after stroke. *Brain, 129*(Pt 6), 1371–1384. <https://doi.org/10.1093/brain/awl090>
- 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. <https://doi.org/10.1016/J.JNEUROLING.2012.01.001>
- Ska, B., Duong, A., & Joanne, Y. (2004). Discourse impairments. In R.D. Kent (Ed.),

- The MIT encyclopedia of communication disorders* (pp. 302–304). The MIT press.
- Sloetjes, H., & Wittenburg, P. (2008). *Annotation by category-ELAN and ISO DCR*.
http://www.lrec-conf.org/proceedings/lrec2008/pdf/208_paper.pdf
- Stark, B. C. (2019). A Comparison of Three Discourse Elicitation Methods in Aphasia and Age-Matched Adults: Implications for Language Assessment and Outcome. *American Journal of Speech-Language Pathology*, 28(3), 1067–1083.
https://doi.org/10.1044/2019_AJSLP-18-0265
- Stark, B. C., Dutta, M., Murray, L. L., Bryant, L., Fromm, D., MacWhinney, B., Ramage, A. E., Roberts, A., den Ouden, D. B., Brock, K., McKinney-Bock, K., Paek, E. J., Harmon, T. G., Yoon, S. O., Themistocleous, C., Yoo, H., Aveni, K., Gutierrez, S., & Sharma, S. (2021). Standardizing Assessment of Spoken Discourse in Aphasia: A Working Group With Deliverables. *American Journal of Speech-Language Pathology*, 30(1S), 491–502. https://doi.org/10.1044/2020_AJSLP-19-00093
- 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. <https://doi.org/10.1080/02687030903524729>
- Statistics Canada. (2017). Montréal, V [Census subdivision], Quebec and Montréal, TÉ [Census division], Quebec(table). In *Census Statistics Canada Catalogue no. 98-316-X2016001*. <https://www12.statcan.gc.ca/census-recensement/2016/dp-pd/prof/index.cfm?Lang=E>
- Stockbridge, M. D., Berube, S., Goldberg, E., Suarez, A., Mace, R., Ubellacker, D., & Hillis, A. E. (2019). Differences in linguistic cohesion within the first year following right- and left-hemisphere lesions. *Aphasiology*, 1–15.
<https://doi.org/10.1080/02687038.2019.1693026>
- Stockbridge, M. D., Berube, S., Goldberg, E., Suarez, A., Mace, R., Ubellacker, D., & Hillis, A. E. (2021). Differences in linguistic cohesion within the first year following right- and left-hemisphere lesions. *Aphasiology*, 35(3), 357–371.
<https://doi.org/10.1080/02687038.2019.1693026>
- 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>
- 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. <http://www.ncbi.nlm.nih.gov/pubmed/6192869>
- Wallace, S. J., Worrall, L., Rose, T., Le Dorze, G., Cruice, M., Isaksen, J., Kong, A. P. H., Simmons-Mackie, N., Scarinci, N., & Gauvreau, C. A. (2018). Which outcomes are most important to people with aphasia and their families? an international nominal group technique study framed within the ICF. *Disability and Rehabilitation*, 39(14), 1364–1379. <https://doi.org/10.1080/09638288.2016.1194899>
- 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>
- 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. <http://www.ncbi.nlm.nih.gov/pubmed/7354627>

Appendix A

Thematic Units (TUs) produced by at least 75% of 45 healthy speakers (Brisebois et al., 2020).

Words	% of identification
<i>Bateau</i> (boat)	91.1
<i>Cerf-volant</i> (kite)	93.3
<i>Château de sable/sable</i> (sand castle/sand)	93.3
<i>Chien</i> (dog)	91.1
<i>Femme/madame/maman</i> (woman/Mrs./mom)	93.3
<i>Fille</i> (girl)	91.1
<i>Garçon</i> (boy)	97.8
<i>Homme/monsieur/papa</i> (man, Mr., dad)	95.6
<i>Lac/rivière/mer/eau</i> (lake, river, sea, water)	75.6
<i>Lire</i> (to read)	88.9
<i>Maison/chalet</i> (house, country house)	86.7
<i>Pêcher/pêcheur</i> (to fish/fisherman)	100
<i>Pique-nique</i> (picnic)	91.1
<i>Radio/écouter de la musique</i> (radio/to listen to music)	75.6
<i>Boire/breuvage</i> (to drink, beverage)	77.8
<i>Voiture/auto</i> (car)	82.2