Article publié dans American Journal of Speech-Language Pathology et qui peut être retracé à l'adresse suivante : https://doi.org/10.1044/2022_AJSLP-22-00084

Validation of videoconference administration of picture description from the Western Aphasia Battery- Revised in neurotypicals Canadian French speakers

Karine Marcotte[1,2], Arianne Lachance [1], Amélie Brisebois[1,2], Patrizia Mazzocca[1], Marianne Désilets-Barnabé[3], Noémie Desjardins[2], and Simona Maria Brambati[3,4,5]

[1] École d'orthophonie et d'audiologie, Faculté de médecine, Université de Montréal, Montréal, Québec, Canada.

[2] Centre de recherche du Centre intégré universitaire de santé et de services sociaux du Nord-de-l'Île-de-Montréal, Montréal, Québec, Canada.

[3] Faculté de médecine, Université de Montréal, Montréal, Québec, Canada.

[4] Centre de recherche de l'Institut Universitaire de Gériatrie de Montréal, Montréal, Québec, Canada.

[5] Département de psychologie, Faculté des arts et des sciences, Université de Montréal, Montréal, Québec, Canada

***Correspondence:**

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

Abstract

Purpose: During the COVID-19 pandemic, clinicians and researchers have increasingly used remote online assessments to pursue their activities, but mostly with tests not validated for videoconference administration. The current study aims to validate the remote online administration of picture description in Canadian French neurotypical speakers and to explore the TUs checklist recently developed.

Method: Spoken discourse elicited through the picture description task of the Western Aphasia Battery-Revised (WAB-R) was collected from Canadian French neurotypical speakers from Quebec aged between 50 and 79 years old. Forty-seven (47) participants completed the task in person, and 49 participants completed the task by videoconference. Videos of each discourse sample were transcribed using CHAT conventions. Microstructural variables were extracted using the Computerized Language Analysis (CLAN) program whereas thematic informativeness was scored for each sample using thematic units (TUs). Chi-square tests were conducted to compare both groups on each TUs. T-tests were also performed on the total score of TUs and microstructural variables.

Results: Groups were matched on sex, age, and education variables. The t-tests revealed no intergroup difference for the total TU score and for the microstructural variables (e.g., mean length of utterances, number of words per minute). Chi-square tests showed no significant intergroup difference for all 16 TUs.

Conclusions: These findings support remote online assessment of the picnic scene of the WAB-R picture description in Canadian French neurotypical speakers. These results also validate the 16 TUs most consistently produced. The use of videoconference could promote and improve the recruitment of participants who are usually less accessible, such as people using assistive mobility technologies.

INTRODUCTION

During the COVID-19 pandemic, researchers worldwide had to stop their activities conducted in person with human subjects. Thus, to continue their research activities, many teams turned to remote online assessments (e.g., Bratan et al., 2021; Cassarino et al., 2022). Telecommunications technology used to deliver communication assessments and interventions, also known as telehealth in clinical settings, existed even before the COVID-19 pandemic (e.g., Mashima & Doarn, 2008; Wales et al., 2017). A recent systematic review revealed that patients in rural settings, representing nearly half of the worldwide population, appreciated telehealth services, including speech-language therapy (Harkey et al., 2020). Telerehabilitation fosters a more natural and ecological therapeutic context and could potentially facilitate assessments integrated into the patients' daily life activities (McCue et al., 2010; Pitt et al., 2019).

The existing literature in speech-language pathology suggests high agreement between remote online assessments and in-person assessments and interventions in both adults and children (Edwards et al., 2012). Only three tests developed for language assessment of adults with chronic acquired aphasia have been validated for videoconference administration: the Boston Naming Test–Second Edition (Hill et al., 2009; Short form; Kaplan et al., 2001), the Boston Diagnostic Aphasia Examination–Third Edition (Short form; Goodglass et al., 2001; Hill et al., 2009; Palsbo, 2007; Theodoros et al., 2008) and more recently the Western Aphasia Battery-Revised (WAB-R, Kertesz, 2006; Dekhtyar et al., 2020). Participants with a variety of aphasia severities completed videoconference assessments, and the results were similar to in-person assessments. Further research is needed to support the validity of remote assessments with adults, including both adults with communication impairments and neurotypical adults, in order to establish the feasibility of the videoconference administration, develop standardized guidelines (Dekhtyar et al., 2020) and standardized normative data. Challenges have been previously identified with videoconference assessment using pictorial stimuli when participants were connected using a phone or when participants had poor internet speed when using tasks involving a timing component (Wood et al., 2021). Considering the increasing use of virtual assessment, a recent technical report proposed detailed guidelines to conduct videoconference assessments of language, which included spoken discourse tasks, in both neurotypical individuals and people with aphasia (Doub et al., 2021). A procedure sheet, an ideal environment checklist, examples of troubleshooting as well as material to ensure comprehension of study procedure for people with aphasia were proposed. More studies using standardized guidelines are required to develop normative data adapted to videoconference for both neurotypical individuals and people with aphasia.

Discourse-oriented studies are of great importance, considering that gaps in this level can significantly impede individuals' communication skills and impact their social skills (e.g., Galski et al., 1998; Steel & Togher, 2019; Worrall et al., 2011). In the present study, we refer to the functional definition of spoken discourse, where spoken discourse refers to language beyond the sentence level for a specific purpose (Armstrong, 2000). Up until now, clinicians mainly assessed picture description using transcription-less analysis, such as the qualitative scales for the BDAE-3 (Goodglass et al., 2001) and the WAB-R (Kertesz, 2006). These scales have shown high interrater reliability for experienced raters in their previous versions, but the reliability is below expectations for community speech-language pathologists (Gordon, 1998). Although limited within clinical speech language pathology, linguistic discourse analysis can allow an increased detection of many language deficits. A growing number of studies applied linguistic discourse analysis over the past decades. A recent review identified over 500 different discursive measures in 165 studies (Bryant et al., 2016). These measures refer to either macrostructure or microstructure features. Namely, macrostructural variables include discourse-level processing such as informativeness, coherence and cohesion, while microstructural variables refer to discourse's lexical and grammatical components (Stark et al., 2021). Additionally, several tasks can be used to elicit spoken discourse. These include structured or semi-structured narratives and story-retell of well-known stories, procedural elicitation, and picture description. The length and content of productions vary significantly depending on the nature of the eliciting task (Bryant et al., 2016). Among these various genres, picture descriptions provide a relatively constrained spoken discourse sample with expected content (Chenery & Murdoch, 1994). This allows the use of a more standardized approach and facilitates performance comparison over time and across different groups. Picture description tasks are widely used in clinical settings to elicit connected speech samples and are part of most language assessment batteries.

Very few studies investigated either macrostructure or microstructure features in spoken discourse during videoconference assessment and in-person assessments. Also, most of them were conducted with patients presenting communication impairments and as part of a language assessment battery. To the best of our knowledge, only two studies specifically compared discourse assessment in videoconference and in-person administration. The first study (Brennan et al., 2004) was conducted in a group of 40 subjects who suffered from either a traumatic brain injury or a stroke using the Story Retelling Procedure (SRP; Doyles et al., 1998). No significant difference was found across all subjects between the videoconference and in-person administrations. However, the sub-group of patients who suffered from a traumatic brain injury performed more poorly in the videoconference administration. Moreover, the difference between the two administrations was not influenced by sociodemographic variables (i.e. age, gender and education). SRP is a well-described method measuring discourse production and comprehension of spoken narratives in persons with aphasia, but this method essentially differs from commonly employed discourse tasks in clinical settings (McNeil et al., 2007). The second study was conducted on 20 individuals who suffered from a traumatic brain injury (Turkstra et al., 2012). Various spoken discourse tasks were administered both in a videoconference and in-person and analyzed using several microstructural measures. Interestingly, the results showed no significant differences between videoconference and in-person assessments in various discourse tasks regarding several microstructural measures (i.e., total number of words, total number of c-units, type-token ratio, and number of mazes (revisions and repetitions). One of the major caveats of the previous studies is that the videoconference environment was simulated as participants were in an adjacent room and not at home, where there is no assistant to set up the connection and potential distraction in the room (Cherney & van Vuuren, 2012). Recently, videoconference administration of the complete WAB-R, which includes spoken discourse production, has been compared to in-person administration in a group of 20 adults with chronic aphasia (Dekhtyar et al., 2020). However, the comparison between in-person and videoconference administration has been made on the total WAB-R scores (i.e., aphasia quotient, language quotient and cortical quotient), but not on each subscore. Therefore, this study does not provide information regarding the reliability of linguistic spoken discourse analysis between in-person and videoconference administration.

Lately, our research team documented in-person performance of thematic informativeness in individuals with aphasia speaking Canadian French (Brisebois et al., 2020, 2021) using the picture description task (i.e. picnic scene) from the Western Aphasia Battery-Revised (Kertesz, 2006). Picture description of the picnic scene is a widely used task because it has a mean duration of 90 seconds and can be completed at the bedside. Informativeness is a macrostructural measure that assesses the ability of an individual to convey relevant information about a given stimulus (Armstrong, 2000). A variety of measures have been used to examine informativeness, such as correct information unit (CIU; Nicholas and Brookshire 1993), main content analysis (MCA; Nicholas and Brookshire 1995) and more recently, thematic units (TUs; Marini et al., 2011a). TUs quickly quantifies thematic informativeness in discourse production. The TU checklist is based on a finite set of semantic themes, which should increase its reliability (Brookshire & Nicholas, 1994).

Briefly, we demonstrated that the TUs score in picture description is sensitive to language impairment and appears related to overall language measure (i.e. a composite score composed of lexico-semantic, comprehension and repetition tasks; García et al., 2020) in both acute and chronic stages of recovery in post-stroke aphasia (Brisebois et al., 2020, 2021). Conversely, very few microstructural variables showed the same qualities. Other groups have reported that microstructural variables were also sensitive to language impairment (Andreetta et al., 2012; Andreetta & Marini, 2015; Fergadiotis & Wright, 2016; Fromm et al., 2016). The mean length of utterances (MLU), the duration, the number of words per minute, the total number of utterances and the moving average token-type ratio (MATTR) are examples of measures with the most sensitivity. To our knowledge, no previous studies reported on the reliability of linguistic spoken discourse analysis between in-person and real-life videoconference administration.

Given that linguistic discourse analysis provides additional sensitivity to language impairments (Bryant et al., 2016), it is important to validate its use in remote videoconference assessment of spoken discourse with adults, including neurotypicals. Validating remote videoconference administration in neurotypical adults is crucial as many studies include a control group of neurotypical adults to differentiate from communication impairments by establishing standardized reference data. Also, better knowledge of expected typical spoken discourse abilities would help researchers develop more accurate models of spoken discourse (Sherratt & Bryan, 2019). Indeed, there is a clear need to provide an additional perspective on how non-linguistic variables (e.g., executive functioning) interacts with the linguistics variables, which are not yet included in current models of discourse processing. A better comprehension of the relationships between the various variables (i.e., cognitive, linguistic, demographic and social cognition) involved in discourse processing will allow a better differentiation between normal and impaired discourse abilities. However, this requires large samples of neurotypical adults which can be easier using videoconference administration. Before recruiting larger groups of neurotypical adults using videoconference, we need to validate its use and establish detailed protocols. Thus, the purpose of the current study is two-fold. First, it aimed to validate the remote videoconference assessment of spoken discourse using the WAB-R picture description task in neurotypical Canadian French speakers using microstructural measures and macrostructural (i.e. thematic informativeness), using the best practices guidelines recently proposed (Stark et al., 2022). We expected similar results across the two modalities based on previous findings (Georgeadis et al., 2004; Turkstra et al., 2012). Second, it aimed to continue the exploration of the TUs checklist that we have recently developed (Brisebois et al., 2020, 2021). We expected that each TUs would be produced by a similar proportion of participants in both the *in-person* group and *videoconference* group.

MATERIALS AND METHODS

Participants

Ninety-six neurotypical participants aged between 50 and 79 years old were recruited. Forty-seven participants completed the task in person, and 49 participants completed the task in a videoconference assessment. Inclusion criteria were: being at least 50 years old and having Canadian (Quebec) French as a language of use. All participants were recruited in the Montreal (Quebec) area. Exclusion criteria included severe mental illness, acquired or developmental language impairments, neurological impairments (including neurocognitive disorders), traumatic brain injury, and self-reported uncorrected visual or auditive deficits. A cognitive screening using the Montreal Cognitive Assessment (MOCA; Nasreddine et al., 2005) with the instruction for remote administration (https://www.mocatest.org/remote-moca-testing/) was completed only for the individuals from the *videoconference group*. For the *in-person group*, no cognitive screening was used, but all participants lived independently, and none reported functional impairments in their daily life activities

Participants from the *in-person group* were recruited before the COVID-19 pandemic through the participants' bank of the *Centre de recherche de l'Institut Universitaire de Gériatrie de Montréal* (http://www.criugm.qc.ca/en/participate.html) and through social media ads. Twenty-seven participants were recruited for a project which sought to establish normative data for picture description tasks and had been approved by the ethical committee from the *Centre intégré universitaire de santé et de services sociaux du Centre-Sud-de-l'Île-de-Montréal (Comité d'éthique de la recherche — Vieillissement et neuroimagerie,* # CER VN 17-18-12). Twenty other participants were recruited for a project which sought to investigate longitudinal changes in post-stroke aphasia approved by the ethical committee at *Centre de recherche du Centre intégré universitaire de santé et de services sociaux du Nord-de-l'Île-de-Montréal* (CIUSSS-NIM; # MP-32-2018-1478). Because of the pandemic of COVID-19, we modified our research protocols and continued our recruitment and testing using videoconference. All participants from the *videoconference group* were recruited during the COVID-19 pandemic in a project which sought to investigate longitudinal spoken discourse changes following a stroke. This project was approved by the ethical committee at CIUSSS-NIM (# 2020-1900). Written informed consent was obtained from all participants.

The *in-person group* (31 females, 16 males; 67.4 ± 7.7 years; mean education: 15.2 ± 2.8 years) and *videoconference group* (30 females, 19 males; 64.53 ± 6.79 years; mean education: 16.24 ± 2.80 years) were matched on sex (t = .70, p = .49), age (t = 1.91, p = .06) and education variables (t = -1.92, p = .06).

Procedure

For the *in-person group*, testing took place in our research facilities, during which the participant was seated and alone with the examiner in a private office. Spoken discourse audio samples were recorded using a Sony IC recorder icd-px312 for 27 participants and a Sony HDR-PJ540 camera (9.2

megapixels) for the remaining 20. The picture stimulus was presented on the desk in front of the participant. For the *videoconference group*, the detailed study procedure is available in Supplementary Material 1. Briefly, all participants were tested and recorded during a videoconference using the Zoom platform (https://zoom.us). Participants were instructed to use a computer or an ipad rather than a phone to be able to see the pictorial stimuli. Most participants were able to connect to the Zoom virtual meeting independently. Four participants were assisted with the connection to the Zoom virtual meeting and instructed by phone by the research assistant. When they were connected, the assessor asked the participant to adjust the video camera and audio if necessary, which all participants have been able to do by themselves. The assessor (either a trained registered speech-language pathologist or a trained student in speech-language pathology) also asked the participant to ensure a private and quiet testing environment. The picture stimulus had been scanned and uploaded as a digital .pdf file and was shared with the participants using the 'Shared screen' option of the Zoom platform.

For both groups, spoken discourse was elicited through the picture description task of the WAB-R. The instruction given to the participants before the picture description task was to describe everything they saw happening in the picture, using complete sentences (« *Décrivez en détail tout ce qui se passe sur cette image en utilisant des phrases complètes.* »). If the participants remained silent for more than ten seconds, they were prompted one time by the examiner with the following sentence: « Is there something you would like to add? » *(« Avez-vous quelque chose à ajouter ? »)*. The recording was then stopped.

Transcriptions

The procedures, including the thematic informativeness coding and analysis and the microstructural 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.B-D. or N.D.) using CHAT conventions (MacWhinney, 2000) for phonetic transcription, utterance segmentation, scoring for utterances, and identification of lexical errors. Full transcriptions were conducted, and the transcription was verbatim. Utterance segmentation was based on a combination of phonological, syntactic, and semantic criteria (Marini, et al., 2011a). Additional guidance for French users of this program was incorporated (Colin & Le Meur, 2016).

Spoken Discourse Measures

Thematic informativeness

Thematic units (TUs) were extracted from each sample and given a score of one point for a complete unit. Similar to Marini et al. (2011b), the focus of the present analysis aimed at the 16 TUs produced by at least 75% of 45 neurotypical Canadian French adults (see Brisebois et al., 2020 for more details about the selection of the TUs). For each participant, a maximum of 16 thematic units could be obtained, irrespective of the length/elaboration of the picture description.

Microstructural variables

Transcriptions underwent detailed analysis focusing on measures known to be sensitive to language impairment (Andreetta et al., 2012; Andreetta & Marini, 2015; Fergadiotis & Wright, 2016; Fromm et al., 2016). The CHAT manual (MacWhinney, 2000) was used to conduct utterances segmentation, transcription and scoring, with additional guidance for French speakers (Colin & Le Meur, 2016). Once the transcription was completed, the coding of morphological and grammatical information was conducted using the program called *mor,* which tags morphemes and words under each utterance in the transcripts. Subsequently, all microstructural variables were extracted for each sample using the program EVAL of CLAN (MacWhinney 2000; version of January 5th, 2021, updated September 30th, 2021).

Extracted productivity measures were defined as the total number of utterances and the number of words per minute. Extracted grammatical/syntactic complexity were measured by the mean length of utterance (MLU), the number of verbs/utterance, noun-to-verb ratio and propositional density (Brown et al., 2008). A separate script on CLAN was used for the Moving Average Token-Type Ratio (MATTR; Covington, 2007), which is commonly used to measure lexical diversity (e.g., Brisebois et al., 2020, 2021; Fergadiotis et al., 2011; Fergadiotis & Wright, 2011) and has the advantage of being independent of the length of the corpus. Finally, reformulations and repetitions were also extracted to measure discourse fluency.

Statistical analysis

All statistical analyses were completed using SPSS® v26.0, and the significance level was set at $p < .05$. Inter-rater coding reliability was assessed using two-way random effects intraclass correlations (ICC, Cronbach's alpha) with a consistency model. Independent samples *t*-tests were conducted for the total TUs and all microstructural variables, including the sample duration extracted from CLAN, to evaluate group differences. A Bonferroni correction was made for multiple comparisons, resulting in an alpha level of .005. Finally, Fisher exact tests were conducted to compare both groups on each 16 TUs.

Inter-rater reliability

Inter-rater reliability (IRR) was calculated separately for both groups. IRR was conducted on 10 samples (20%) of the *videoconference* group by a second-rater (N.D.) and 17 samples (36%) of the *in-person* group. IRR for the *in-person* group have been reported in a previous conducted on a larger sample of participants (Boucher et al., 2022), but have been recalculated using only participants from the present study and by adding confidence interval and standard error measurement. To determine consistency between raters, ICCs were calculated on the total number of TUs, the sample's duration, and all the microstructural variables extracted from CLAN. ICCs were interpreted following Cicchetti (1994): poor ($r < .40$), fair ($.40 \le r < .60$), good ($.60 \le r < .75$), or excellent ($r \ge .75$). The results of these analyses are reported in Supplementary Material 2. All variables met the threshold of excellent reliability, ICC ≥ .75. Best practice guidelines checklist from Stark et al. are reported in Supplementary Material 3.

RESULTS

Sample duration range was between 41 and 362 seconds (T1) for the *in-person group* and from 44 to 180 seconds for the *videoconference group*. The mean duration of the samples produced by the *in-person group (*mean = 90.28 ± 57.44) and the *videoconference group* (mean = 85.93 ± 38.46) were not significantly different (t=0.426, p=.664, standard error measurement 9.94).

Table 1 reports the means, standard deviations, ranges of each group of the total of TUs and all microstructural variables, as well as the statistical results. In sum, the total of TUs produced by the *in-person group* and the *videoconference group* were not significantly different (t=-1.91, p=.06). Similarly, the difference between the two groups was not significant (p>.05, two-tailed) for all microstructural variables, except for the number of verbs per utterance (t=-2.15, p=.034). However, the difference was not significantly different using the Bonferroni correction (p<.005, two-tailed). Figure 1 represents the distribution of the scores of duration, total of TUs and mean length of utterances (words), which are representative of all variables extracted.

Table 2 shows the number and the percentage of participants who produced each TU in both groups.

The results of the Fisher exact tests between the groups are also detailed in Table 1. In sum, no significant difference between *in-person* and *videoconference groups* was found for each TU ($p>.05$, two-tailed).

DISCUSSION

The present study aimed to validate remote videoconference assessment of spoken discourse using the picture description task from the WAB-R (Kertesz, 2006) in neurotypical Canadian French speakers. As expected, we provided evidence that there is no difference between *in-person* and *videoconference* administration of the WAB-R picture description task in thematic informativeness (i.e., a macrostructural measure) and an array of microstructural measures (i.e., productivity, grammatical/syntactic complexity, lexical diversity, and fluency measures) in neurotypical Canadian French speakers. We first demonstrated the validity of videoconference administration of this picture description task for quantitative analysis of spoken discourse production in neurotypical individuals.

The field of speech-language pathology, and other healthcare professionals, slowly expanded the use of telerehabilitation in the past few years, but its use during the COVID-19 pandemic has grown exponentially. However, the validation of teleassessment has not increased as quickly, but it is essential to provide a proper administration of protocols that were first developed for in-person administration (Dekhtyar et al., 2020). The current study represents a first step aiming to validate remote administration of picture description as it was only conducted in neurotypical individuals. The present result confirms the validity of spoken discourse assessment using videoconferencing platforms in a natural environment (i.e., at home), as previously reported in simulated environment (Georgeadis et al., 2004; Turkstra et al., 2012). In these previous studies, participants were in an adjacent room and an assistant was able to set up the connection and hence causing a potential distraction in the room (Cherney & van Vuuren, 2012). Videoconference administration conducted when participants are at home (or another quiet environment of their choice) may promote the recruitment of groups that are more difficult to reach usually, such as people in remote areas, with atypical work schedules or with mobility challenges. Further studies will be required to investigate the need for caregiver assistance to assess picture description (and other types of spoken discourse) using videoconference in people with communication impairments to provide guidelines and recommendations to clinicians using this type of administration. For instance, Dekhtyar et al. (2020) reported that most of their 20 individuals with post-stroke aphasia required at least minimal assistance to set up the connection for the administration of the complete WAB-R (Kertesz, 2006).

As mentioned previously, our research team recently focused on thematic informativeness (i.e., TUs) as a quantitative measure of spoken discourse. Results of the present study extend the validity of the TU checklist we developed recently (Brisebois et al., 2020). We replicated the results obtained previously as the total of TUs produced by the *in-person group* and the *videoconference group* were similar. All 16 TUs were produced by at least 75% of the participants of each group and in similar proportions. Therefore, this study is the first to generate a TU checklist for the picture description of WAB-R in a relatively large control group sample of Canadian French speakers. One of the main reasons for our recent focus on thematic informativeness is that it is a clinician-friendly, easy-to-implement, and provides a prompt quantitative measure of spoken discourse. Further studies will be required to validate the reliability of the TU checklist in clinical settings, but these preliminary normative data could potentially be used for both diagnostic purposes and to assess the longitudinal changes of aphasia recovery.

Although the difference between the two groups was not significant, we noted that the administration of picture description by *videoconference* seems to demonstrate a slight advantage for the production of TUs compared to administration *in-person*. Three potential reasons have been identified to support this advantage. First, the image presented was probably larger on the computer screens, depending on each participant's screen, than its original form in the stimuli book. A larger image might have helped some individuals to see smaller details. For example, the kite was named by all participants in the videoconference group, whereas four participants in the in-person group did not refer to it (Ferraro et al., 1997). Second, the videoconference group was, on average, a little younger and a little more educated than the in-person group, although the difference was not significant. Despite perceptions, the overall results of studies about the impact of aging in spoken discourse production remain inconclusive. Our recent study using the same picture description task showed significant differences between the age groups (50-69 y.o. and 70-90 y.o.) for only three measures (i.e. repetitions, speech rate and efficiency; Boucher et al., 2022). However, the effect of age seems to be modulated by the sex of participants (Ardila & Rosselli, 1996). Namely, a steady decline in spoken discourse was observed in men only after the age of 50, with no comparable decline in women. The effect of education also seems to vary according to the measures extracted in the different discourse genres. For instance, the level of education did not affect the macrostructural parameters evaluated in conversation, but differences in microstructural variables (namely, shorter and less complete picture description) were observed in individuals with lower education levels (Mackenzie, 2000). Third, videoconference assessment allows a more familiar and more comfortable environment than our research offices in an acute care hospital. On the other hand, the room was quieter, and there were no distractions during the assessment for the *in-person group*.

This study has a few limitations, which should be acknowledged. First and foremost, the participants only completed one type of administration (*in-person* or *videoconference*). The pandemic of COVID-19 required us and others to quickly modify our research protocols and continue our recruitment and testing using videoconference. One group of participants was recruited before the pandemic, whereas the other was recruited during the pandemic, which explains why only one modality was used for each participant. Nonetheless, considering that no data are currently available in Canadian French for assessment during a videoconference and that an increasing number of clinicians and researchers will use videoconference assessments in the future, we believe that the present study is a practical starting point in the establishment of validation of remote assessment for the Canadian French population. Second, these results cannot be extended to all clinical populations, especially for people with more severe aphasia (e.g., Hill et al., 2009). The present results will need to be validated with individuals with communication impairments to provide specific guidelines to clinicians. Third, no cognitive screening was used for the in-person group, which could have contributed to the slight spoken discourse differences between groups. However, these participants were recruited via a participant's research bank from a geriatric center which gave us a list of participants without neurocognitive impairments. Also, none reported functional impairments in their daily life activities. Fourth, although the sample size is not very large, this study contributes to a more standardized assessment of spoken discourse variables in Canadian French, as previous work conducted by our research team (Boucher et al., 2022). Fifth, the sample had an overrepresentation of women (64% of the sample) and highly-educated individuals, similar to most studies conducted in the Québec French population (e.g., Bourgeois-Marcotte et al., 2015; Callahan et al., 2010; Marcotte et al., 2017). The recruitment of men and individuals with lower levels of education is challenging, and the use of remote assessment might favor their participation if we eliminate the travel time. Ideally, a stratified sampling method would have been preferable to ensure that every subgroup is correctly represented in the sample. In other words, based on the demographic statistics of Quebec, the population should be divided into subgroups based on relevant characteristics such as age, education, and gender, which will maximize the representativeness of the sample. Finally, the aim of the present study was not to investigate the stability of the measures over time. Thus, test-retest stability remains to be investigated in future studies as it was not accounted for in the present study.

In conclusion, we believe there is a need to validate videoconference assessment administration in Canadian's French population from Quebec. Further work should expand to larger sample sizes and various spoken discourse genres, including less structured tasks (e.g., story retelling, personal recounts) in both neurotypical individuals and individuals with communication impairments. Some modifications may have to be made for efficient videoconference administration to ensure equivalent performance to in-person test administration (Dekhtyar et al., 2020; Doub et al., 2021).

ACKNOWLEDGMENTS

We are very grateful to all the participants for their contribution to this study. Also, we wish to thank Johémie Boucher for her help in the data processing. This project was funded by the Heart and Stroke Foundation (grant-in-aid numbers G-16-00014039 and G-19-0026212) to K.M. 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".

DATA AVAILABILITY STATEMENT

The raw data presented in this article are not readily available because of the sensitivity of the video materials. The datasets analyzed during the current study are available from the corresponding author on reasonable request.

REFERENCES

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

Ardila, A., & Rosselli, M. (1996). Spontaneous language production and aging: Sex and educational effects. *International Journal of Neuroscience*, *87*(1–2), 71–78. https://doi.org/10.3109/00207459608990754

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

Boucher, J., Brisebois, A., Slegers, A., Courson, M., Désilets-Barnabé, M., Chouinard, A.-M., Gbeglo, V., Marcotte, K., & Brambati, S. M. (2022). Picture description of the Western Aphasia Battery: Reference data for the Canadian-French population. *American Journal of Speech-Language Pathology*, *31*(1), 257–270. https://doi.org/10.1044/2021_AJSLP-20-00388

Bourgeois-Marcotte, J., Wilson, M. A., Forest, M., & Monetta, L. (2015). TEFREP: Épreuve de répétition de phrases en franco-québécois. Développement, validation et normalisation. *Canadian Journal on Aging / La Revue Canadienne Du Vieillissement*, *34*(3), 391–396. https://doi.org/10.1017/s0714980815000173

Bratan, T., Aichinger, H., Brkic, N., Rueter, J., Apfelbacher, C., Boyer, L., & Loss, J. (2021). Impact of the COVID-19 pandemic on ongoing health research: an ad hoc survey among investigators in Germany. *BMJ Open* , *11*(e049086).

Brennan, D. M., Georgeadis, A. C., Baron, C. R., & Barker, L. M. (2004). The Effect of Videoconference-Based Telerehabilitation on Story Retelling Performance by Brain-Injured Subjects and Its Implications for Remote Speech-Language Therapy. *Telemedicine Journal and E-Health*, *10*(2). https://doi.org/10.1089/tmj.2004.10.147

Brisebois, A., Brambati, S. M., Boucher, J., Rochon, E., Leonard, C., Désilets-Barnabé, M., Desautels, A., & Marcotte, K. (2021). A longitudinal study of narrative discourse in post-stroke aphasia. *Aphasiology*. https://doi.org/10.1080/02687038.2021.1907295

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). https://doi.org/10.1080/02687038.2019.1705661

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.

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). https://doi.org/10.3758/BRM.40.2.540

Bryant, L., Ferguson, A., & Spencer, E. (2016). Linguistic analysis of discourse in aphasia: A review of the literature. *Clinical Linguistics & Phonetics*, *30*(7). https://doi.org/10.3109/02699206.2016.1145740

Callahan, B. L., Macoir, J., Hudon, C., Bier, N., Chouinard, N., Cossette-Harvey, M., Daigle, N., Fradette, C., Gagnon, L., & Potvin, O. (2010). Normative Data for the Pyramids and Palm Trees Test in the Quebec-French Population. *Archives of Clinical Neuropsychology*, *25*(3). https://doi.org/10.1093/arclin/acq013

Cassarino, L., Santoro, F., Gelardi, D., Panerai, S., Papotto, M., Tripodi, M., Cosentino, F. I. I., Neri, V., Ferri, R., Ferlito, S., Modica, D., Fisicaro, F., Pennisi, M., Bella, R., & Lanza, G. (2022). Post-stroke aphasia at the time of COVID-19 pandemic: a telerehabilitation perspective. *Journal of Integrative Neuroscience*, *21*(1), 008. https://doi.org/10.31083/j.jin2101008

Chenery, H. J., & Murdoch, B. E. (1994). The production of narrative discourse in response to animations in persons with dementia of the Alzheimer's type: Preliminary findings.

Aphasiology, *8*(2), 159–171. https://doi.org/10.1080/02687039408248648

Cherney, L., & van Vuuren, S. (2012). Telerehabilitation, Virtual Therapists, and Acquired Neurologic Speech and Language Disorders. *Seminars in Speech and Language*, *33*(03). https://doi.org/10.1055/s-0032-1320044

Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, *6*(4). https://doi.org/10.1037/1040-3590.6.4.284

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)*.

Dekhtyar, M., Braun, E. J., Billot, A., Foo, L., & Kiran, S. (2020). Videoconference Administration of the Western Aphasia Battery–Revised: Feasibility and Validity. *American Journal of Speech-Language Pathology*, *29*(2). https://doi.org/10.1044/2019_AJSLP-19-00023

Doub, A., Hittson, A., & Stark, B. C. (2021). Conducting a Virtual Study With Special Considerations for Working With Persons With Aphasia. *Journal of Speech, Language, and Hearing Research*, *64*(6), 2038–2046. https://doi.org/10.1044/2021_JSLHR-20-00392

Doyles, P. J., McNeil, M. R., Spencer, K. A., Goda, A. J., Cottrell, K., & Lustig, A. P. (1998). The effects of concurrent picture presentations on retelling of orally presented stories by adults with aphasia. *Aphasiology*, *12*(7–8). https://doi.org/10.1080/02687039808249558

Edwards, M., Stredler-Brown, A., & Houston, K. T. (2012). Expanding Use of Telepractice in Speech-Language Pathology and Audiology. *The Volta Review*, *112*(3). https://doi.org/10.17955/tvr.112.3.m.704

Fergadiotis, G., & Wright, H. H. (2011). Lexical diversity for adults with and without aphasia across discourse elicitation tasks. *Aphasiology*, *25*(11). https://doi.org/10.1080/02687038.2011.603898

Fergadiotis, G., & Wright, H. H. (2016). Modelling confrontation naming and discourse performance in aphasia. *Aphasiology*, *30*(4). https://doi.org/10.1080/02687038.2015.1067288

Fergadiotis, G., Wright, H. H., & Capilouto, G. J. (2011). Productive vocabulary across discourse types. *Aphasiology*, *25*(10). https://doi.org/10.1080/02687038.2011.606974

Ferraro, F. R., Barth, J., Morton, M., & Whetham, T. (1997). Stimulus Size Effects in Boston Naming Test Performance in Younger and Older Adults. *Applied Neuropsychology*, *4*(4). https://doi.org/10.1207/s15324826an0404_8

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

Galski, T., Tompkins, C., & Johnston, M. V. (1998). Competence in discourse as a measure of social integration and quality of life in persons with traumatic brain injury. *Brain Injury*, *12*(9), 769–782. https://doi.org/10.1080/026990598122160

García, A. O., Brambati, S. M., Brisebois, A., Désilets-Barnabé, M., Rochon, E., Desautels, A., Marcotte, K., Osa García, A., Brambati, S. M., Brisebois, A., Désilets-Barnabé, M., Houzé, B., Bedetti, C., Rochon, E., Leonard, C., Desautels, A., & Marcotte, K. (2020). Predicting early post-stroke aphasia recovery from initial severity. *Frontiers in Neurology*, *11*. https://doi.org/10.3389/fneur.2020.00120

Georgeadis, A., Brennan, D., Barker, L., & Baron, C. (2004). Telerehabilitation and its effect on story retelling by adults with neurogenic communication disorders. *Aphasiology*, *18*(5–7). https://doi.org/10.1080/02687030444000075

Goodglass, H., Kaplan, E., & Barresi, B. (2001). *Boston Diagnostic Aphasia Examination (3rd ed.) (BDAE-3)* (3rd ed.). Pro-Ed.

Gordon, J. K. (1998). The fluency dimension in aphasia. *Aphasiology*, *12*(7–8). https://doi.org/10.1080/02687039808249565

Harkey, L. C., Jung, S. M., Newton, E. R., & Patterson, A. (2020). Patient Satisfaction with Telehealth in Rural Settings: A Systematic Review. *International Journal of Telerehabilitation*, *12*(2). https://doi.org/10.5195/ijt.2020.6303

Hill, A. J., Theodoros, D. G., Russell, T. G., Ward, E. C., & Wootton, R. (2009). The effects of aphasia severity on the ability to assess language disorders via telerehabilitation. *Aphasiology*, *23*(5). https://doi.org/10.1080/02687030801909659

Kaplan, E., Goodglass, H., & Weintraub, S. (2001). *Boston Naming Test (2nd Edition)*. Lippincott, Williams & Wilkins.

Kertesz, A. (2006). *Western Aphasia Battery- Revised*. Pearson.

Mackenzie, C. (2000). Adult spoken discourse: The influences of age and education. *International Journal of Language and Communication Disorders*, *35*(2), 269–285. https://doi.org/10.1080/136828200247188

MacWhinney, B. (2000). *The CHILDES Project: Tolls for Analzying Talk: Vol. 3rd Editio*. 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., Andreetta, S., del Tin, S., & Carlomagno, S. (2011). A multi-level approach to the analysis of narrative language in aphasia. *Aphasiology*, *25*(11), 1372–1392. https://doi.org/10.1080/02687038.2011.584690

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

Mashima, P. A., & Doarn, C. R. (2008). Overview of Telehealth Activities in Speech-Language

Pathology. *Telemedicine and E-Health*, *14*(10), 1101–1117. https://doi.org/10.1089/tmj.2008.0080

McCue, M., Fairman, A., & Pramuka, M. (2010). Enhancing Quality of Life through Telerehabilitation. *Physical Medicine and Rehabilitation Clinics of North America*, *21*(1), 195–205. https://doi.org/10.1016/j.pmr.2009.07.005

McNeil, M. R., Sung, J. E., Yang, D., Pratt, S. R., Fossett, T. R. D., Doyle, P. J., & Pavelko, S. (2007). Comparing connected language elicitation procedures in persons with aphasia: Concurrent validation of the Story Retell Procedure. *Aphasiology*, *21*(6–8). https://doi.org/10.1080/02687030701189980

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., Cummings, J. L., & Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: A Brief Screening Tool For Mild Cognitive Impairment. *Journal of the American Geriatrics Society*, *53*(4). https://doi.org/10.1111/j.1532-5415.2005.53221.x

Nicholas, L. E., & Brookshire, R. H. (1993). A System for Quantifying the Informativeness and Efficiency of the Connected Speech of Adults With Aphasia. *Journal of Speech, Language, and Hearing Research*, *36*(2). https://doi.org/10.1044/jshr.3602.338

Nicholas, L. E., & Brookshire, R. H. (1995). Presence, Completeness, and Accuracy of Main Concepts in the Connected Speech of Non-Brain-Damaged Adults and Adults With Aphasia. *Journal of Speech, Language, and Hearing Research*, *38*(1). https://doi.org/10.1044/jshr.3801.145

Palsbo, S. E. (2007). Equivalence of functional communication assessment in speech pathology using videoconferencing. *Journal of Telemedicine and Telecare*, *13*(1). https://doi.org/10.1258/135763307779701121

Pitt, R., Theodoros, D., Hill, A. J., & Russell, T. (2019). The impact of the telerehabilitation group aphasia intervention and networking programme on communication, participation, and quality of life in people with aphasia. *International Journal of Speech-Language Pathology*, *21*(5), 513–523. https://doi.org/10.1080/17549507.2018.1488990

Sherratt, S., & Bryan, K. (2019). Textual cohesion in oral narrative and procedural discourse: the effects of ageing and cognitive skills. *International Journal of Language & Communication Disorders*, *54*(1). https://doi.org/10.1111/1460-6984.12434

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

Stark, B. C., Bryant, L., Themistocleous, C., den Ouden, D.-B., & Roberts, A. C. (2022). Best practice guidelines for reporting spoken discourse in aphasia and neurogenic communication disorders. *Aphasiology*, 1–24. https://doi.org/10.1080/02687038.2022.2039372

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). https://doi.org/10.1044/2020_AJSLP-19-00093

Steel, J., & Togher, L. (2019). Social communication assessment after traumatic brain injury: a narrative review of innovations in pragmatic and discourse assessment methods. *Brain Injury*, *33*(1), 48–61. https://doi.org/10.1080/02699052.2018.1531304

Theodoros, D., Hill, A., Russell, T., Ward, E., & Wootton, R. (2008). Assessing Acquired Language Disorders in Adults via the Internet. *Telemedicine and E-Health*, *14*(6). https://doi.org/10.1089/tmj.2007.0091

Turkstra, L. S., Quinn-Padron, M., Johnson, J. E., Workinger, M. S., & Antoniotti, N. (2012). In-Person Versus Telehealth Assessment of Discourse Ability in Adults With Traumatic Brain Injury. *Journal of Head Trauma Rehabilitation*, *27*(6). https://doi.org/10.1097/HTR.0b013e31823346fc

Wales, D., Skinner, L., & Hayman, M. (2017). The Efficacy of Telehealth-Delivered Speech and Language Intervention for Primary School-Age Children: A Systematic Review. *International Journal of Telerehabilitation*, 55–70. https://doi.org/10.5195/ijt.2017.6219

Wood, E., Bhalloo, I., McCaig, B., Feraru, C., & Molnar, M. (2021). Towards development of guidelines for virtual administration of paediatric standardized language and literacy assessments: Considerations for clinicians and researchers. *SAGE Open Medicine*, *9*, 205031212110505. https://doi.org/10.1177/20503121211050510

Worrall, L., Sherratt, S., Rogers, P., Howe, T., Hersh, D., Ferguson, A., & Davidson, B. (2011). What people with aphasia want: Their goals according to the ICF. *Aphasiology*, *25*(3), 309–322. https://doi.org/10.1080/02687038.2010.508530

Figure legend

Figure 1. Violin plots showing the distribution of the data and the probability density of the three measures produced during picture description in the *in-person group* and the *videoconference group*. Black dots refer to the group mean and the pointranges represents one standard deviation. Panel A represents the distribution of duration, Panel B represents the distribution of the total score of thematic units and Panel C represents the distribution of the mean length of utterances (words).

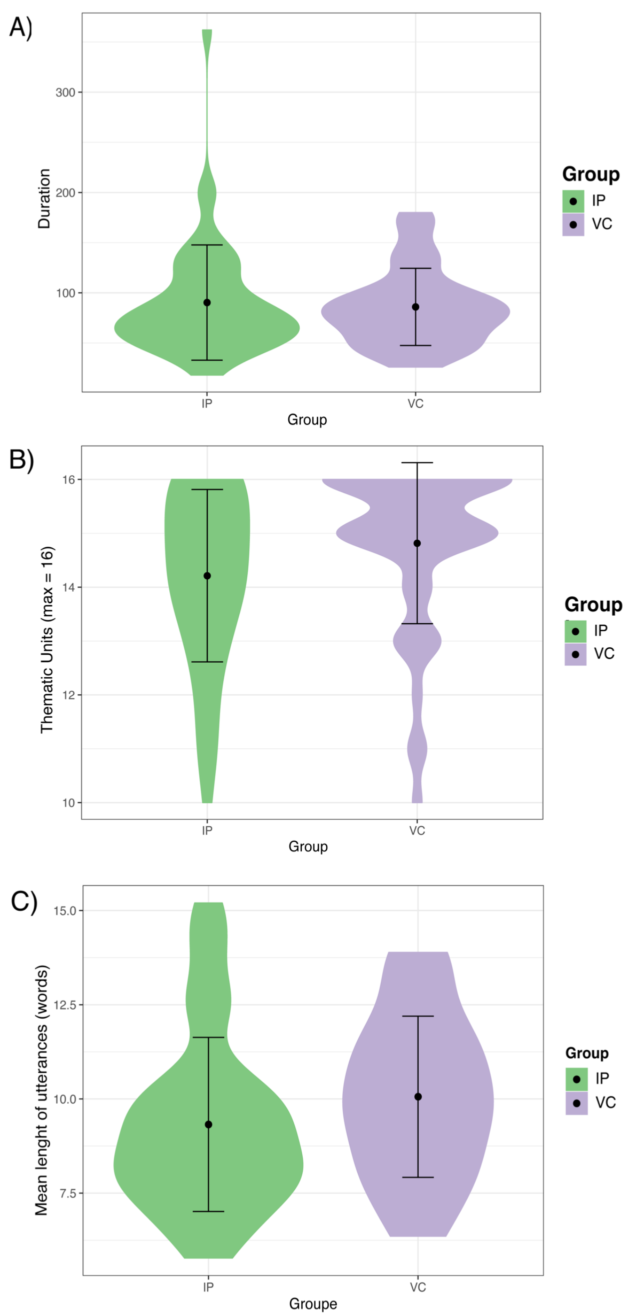


Table 1. Descriptive statistics of the total thematic units score and microlinguistic variables and t-tests comparing group difference between the in-person group and the videoconference group.

Variable	***In-person* Group (n=47)**		***Videoconference* group (n=49)**		**Statistics**		
	Mean (SD)	Range	Mean (SD)	Range	**t**	***p***	**Standard error measurement**
TU_{total}	14.21 (1.60)	10 – 16	14.81 (1.50)	11 – 16	- 1.91	.059	.32
Total number of utterances	23.91 (12.30)	8 – 71	22.76 (10.65)	9 – 63	.24	.811	2.35
Words per minute	157.38 (36.74)	78.07 – 246.67	.914–.995	105.00 – 239.27	- 1.17	.246	6.69
Mean length of utterances (words)	9.32 (2.30)	5.78 – 15.20	10.05 (2.13)	6.36 – 13.89	- 1.62	.109	.45
Mean length of utterances (morphemes)	10.66 (2.59)	7.26 – 17.00	11.49 (2.37)	7.14 – 15.59	- 1.64	.104	.51
Verbs per utterance	.48 (.08)	.17 –.86	.46 (.08)	.29 – 1.30	-2.15	.034*	.06
Noun/verb ratio	14.42 (3.50)	6.62 – 23.42	14.85 (4.18)	6.25 – 22.90	- .55	.587	.79
Propositional density	.35 (.05)	.22 – .46	.37 (.05)	.28 – .46	- 1.51	.133	.01
MATTR	.95 (.01)	.92 – 1.00	.95 (.01)	.91 – .98	.44	.659	.003
Reformulation	105.87 (52.83)	51 – 279	115.96 (56.98)	2 – 309	- .90	.371	11.22
Repetition	3.36 (2.47)	0 – 11	3.08 (2.50)	0 – 9	.55	.582	.51

* not significant using the adjusted p-value following the Bonferroni correction ($p < .005$)
SD = Standard Deviation; TU = Thematic Unit; MATTR = Moving-Average Type-Token Ratio

Table 2. Descriptive statistics of each thematic units and Fisher exact test comparing group difference between the in-person group and the videoconference group.

	In-person group		*Videoconference* group			
Thematic unit	**n (max = 47)**	**%**	**n (max=49)**	**%**	**Fisher exact test**	***p***
Cerf-volant (kite)	45	95.7	49	100.0	2.13	.144
Chien (dog)	45	95.7	48	98.0	.39	.533
Pêcher/pêcheur (to fish/fisherman)	45	95.7	48	98.0	.39	.533
Château de sable/sable (sand castle/sand)	45	95.7	47	95.6	.002	.966
Homme/monsieur/papa (man, Mr., dad)	45	95.7	46	93.9	.17	.681
Bateau (boat)	44	93.6	48	98.0	1.13	.287
Maison/chalet (house. country house)	44	93.6	48	98.0	1.13	.290
Garçon (boy)	44	93.6	45	91.8	.11	.737
Femme/madame/maman (woman/Mrs./mom)	43	91.4	46	93.9	.20	.653
Fille (girl)	43	91.4	45	91.8	.004	.951
Lire (to read)	43	91.4	45	91.8	.004	.951
Boire/breuvage (to drink, beverage)	42	89.4	43	87.8	.06	.805
Pique-nique (picnic)	41	87.2	47	95.9	.237	.124
Radio/écouter de la musique (radio/to listen to music)	39	82.9	42	85.7	.410	.523
Voiture/auto (car)	37	78.7	37	75.5	.14	.708
Lac/rivière/mer/eau (lake, river, sea, water)	36	76.5	42	85.7	1.31	.253

Supplemental Material S1. Study procedure document for the remote assessment.

Original Version (in French)

Guide de passation virtuelle des épreuves

Recrutement

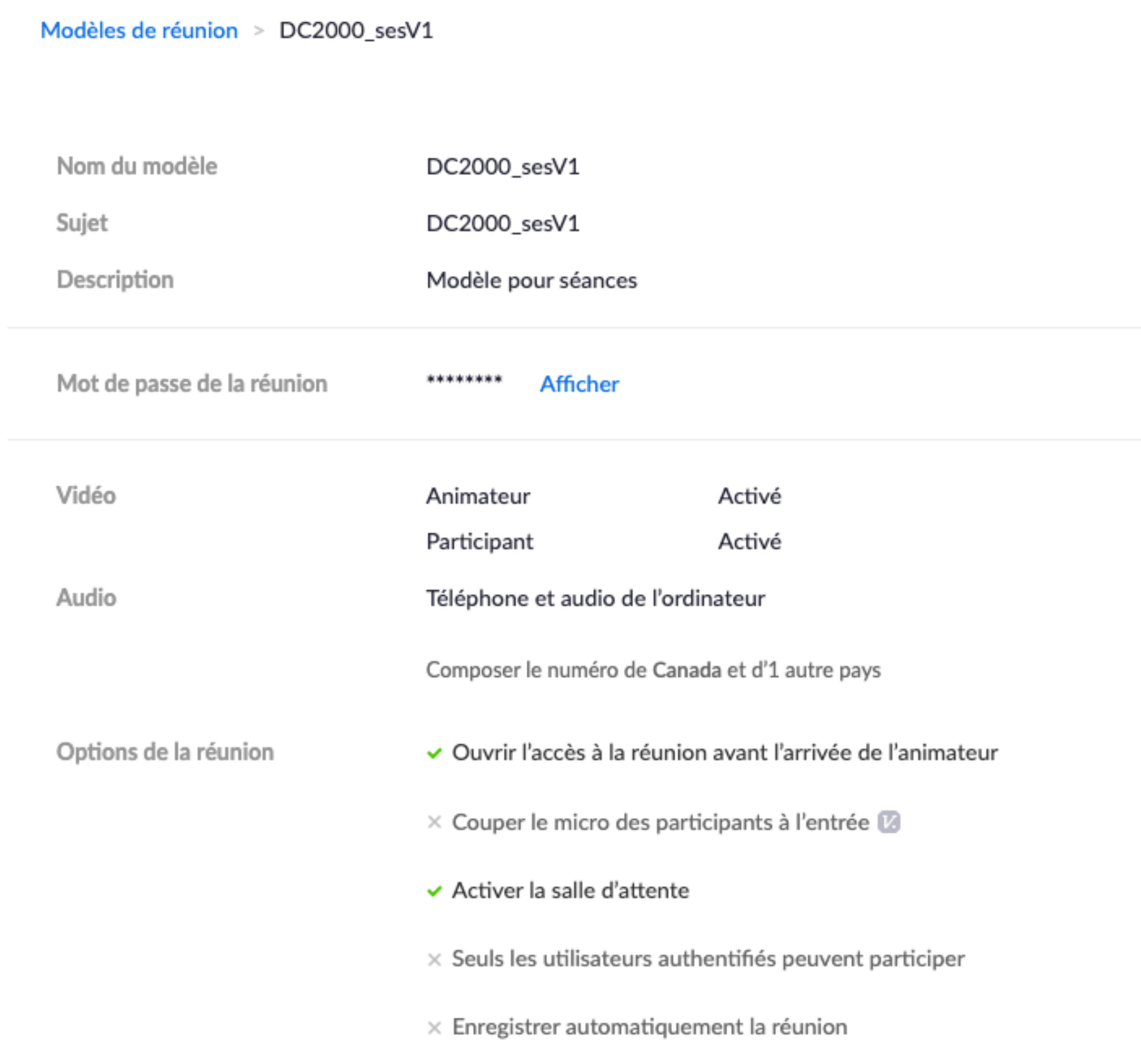
- Annonces affichées sur Facebook.
- Le participant doit nous contacter via courriel ou téléphone (utiliser le script téléphonique pour les personnes qui nous contactent par téléphone).
- Envoyer le formulaire d'information et de consentement par courriel avec l'adresse du laboratoire.
  - Le participant doit prendre connaissance du document, le signer, puis nous le retourner par courriel.

Attendre de recevoir la copie signée du formulaire d'information et de consentement avant de passer à l'étape suivante.

Avant l'évaluation

- Envoyer au participant le lien vers le sondage pré-entrevue sur la plateforme sécurisée RedCAp.
- POUR LES PARTICIPANTS NON-FAMILIERS AVEC ZOOM, d'abord appeler le participant au téléphone pour lui offrir un accompagnement pour l'ouverture de l'appareil (ordi ou iPad) et l'application Zoom. Leur envoyer par courriel les documents et liens suivants.
  - Guide d'utilisation de zoom pour iPAD ou ordinateur pour les participants qui n'ont pas d'expérience avec zoom.
  - Lien pour télécharger Zoom https://zoom.us/download
  - Liens vers des tutoriels d'utilisation de Zoom. https://support.zoom.us/hc/fr?flash_digest=f3fd58514a60ffc5c80803437ca77b4325fc5a29
- APRÈS AVOIR ACCOMPAGNÉ LES PARTICIPANTS NON-FAMILIERS AVEC ZOOM ET POUR LES PARTICIPANTS FAMILIERS AVEC ZOOM , faire un test Zoom avec la personne quelques jours avant l'évaluation. Donner une date et une heure précise au participant. Pendant ce test :
  - Ré-expliquer le déroulement de l'étude.
  - Pratiquer le partage d'écran.
  - Montrer au participant l'enregistrement de la séance.

- Planifier la rencontre virtuelle en indiquant la durée au participant.
- Créer le lien Zoom en nommant la rencontre avec le code du participant.
- Envoyer le lien Zoom au participant à l’avance.



Le jour de l’évaluation

- Se connecter 10 minutes avant le début de la rencontre ou appel téléphonique pour accompagner le participant avec l’ouverture de l’appareil et de Zoom.

- S’assurer que la fenêtre avec le visage du participant est toujours bien visible.

- S’assurer que le participant est dans un environnement calme et peu bruyant. Lui rappeler au besoin.

- S’assurer que le participant vous entend bien et que vous l’entendez bien.

- S’assurer que le participant vous voit bien et que vous voyez clairement le participant.

- **Confirmer le début de l’enregistrement avec le participant. Débuter l’enregistrement.** Idéalement, découper l’enregistrement d’emblée en arrêtant et recommençant l’enregistrement à chacune des tâches. On peut débuter et cesser l’enregistrement de la séance avec le raccourci shift+ ⌘+ C. Faire un seul vidéo est également une option, mais il faudra le segmenter par la suite.

- Expliquer clairement chacune des consignes et valider la compréhension du participant.
    -

Tâche de description du pique-nique de la Western Aphasia Battery-Revised

- Partager votre écran avec l’image de la scène du pique-nique.

- Veuillez-vous assurer que la fenêtre avec le visage du participant est toujours agrandie et clairement visible à des fins d'enregistrement afin de voir clairement son visage.

- Désactivez votre microphone lorsque la consigne est terminée pour éviter tout bruit de fond supplémentaire pendant que le participant parle.

Après avoir complété l’évaluation

- Faire compléter le sondage « Rétroaction suite à participation » au participant en envoyant le lien.
    -
- Télécharger les vidéos et les déposer dans l’espace sécurisé prévu en utilisant la nomenclature des fichiers vidéos suivante :

ParticipantAnnéeNuméro_ses-Vnumérodetesting_task-nomdelatâche
Ex. : participant no.1 groupe contrôle testé en 2020, 1ere évaluation, tâche de description d’image du WAB = DC2001_ses-V1_task-WAB

Translation in English

Videoconference assessment guide

Recruitment

- Ads displayed on Facebook.
- The participant must contact us via email or telephone (use the telephone script for people who contact us by telephone).
- Send the information and consent form by email with the laboratory email address
    - The participant must read the document, sign it, then return it to us by email.

Wait until the reception the signed copy of the consent form before proceeding to the next step.

Before the assessment

- Send the participant the link to the pre-interview survey on the secure platform RedCAp.
- FOR PARTICIPANTS NOT FAMILIAR WITH ZOOM, first call the participant on the phone to offer him support for opening the device (computer or iPad) and the Zoom application. Email them the following documents and linksGuide d'utilisation de zoom pour iPAD ou ordinateur pour les participants qui n'ont pas d'expérience avec zoom.
    - Link to download Zoom https://zoom.us/download
    - Link to Zoom tutorials https://support.zoom.us/hc/fr?flash_digest=f3fd58514a60ffc5c80803437ca77b4325fc5a29
- AFTER ACCOMPANIING PARTICIPANTS NOT FAMILIAR WITH ZOOM AND FOR PARTICIPANTS FAMILY WITH ZOOM, do a Zoom test with the person a few days before the assessment. Give a specific date and time to the participant. During this test:
    - Re-explain the course of the study.
    - Practice screen sharing.
    - Show the participant the recording of the session.
- Plan the virtual meeting by indicating the duration to the participant.
- Create the Zoom link by naming the meeting with the participant code.
- Send the Zoom link to the participant in advance.

Day of the assessment

- Connect 10 minutes before the start of the meeting or phone call to accompany the participant with the opening of the device and Zoom.
  -
- Make sure that the window with the participant's face is always clearly visible.
- Ensure that the participant is in a calm and quiet environment. Remind him if necessary.
- Make sure the participant can hear you and that you can hear them.
- Make sure the participant can see you clearly and that you can also see the participant clearly.
  -
- **Confirm the start of recording with the participant. Start recording.** Ideally, cut the recording from the start by stopping and restarting the recording at each task. You can start and stop recording the session with the shortcut shift+ ⌘+ C. Making a single video is also an option, but it will have to be segmented afterwards.
  -
- Clearly explain each of the instructions and validate the participant's understanding.

Western Aphasia Battery-Revised Picnic Description Task

- Share your screen with the picnic scene image.
- Please ensure that the window with the participant's face is always enlarged and clearly visible for recording purposes in order to see their face clearly.
- Mute your microphone when the briefing is over to avoid additional background noise while the participant is speaking.

Agfter completing the assessment

- • Have the participant complete the “Feedback following participation” survey by sending the link.
-
- Download the videos and place them in the secure space provided using the following video file nomenclature:

ParticipantYearNumber_ses-Vsessionnumber_task-nameofthetask

Ex. : participant no.1 control group tested in 2020, first assessment, Western Aphasia Battery-Revised Picnic Description Task = DC2001_ses-V1_task-WAB

Supplementary Material 2. Best Practice Guidelines checklist from Stark et al. (2022)*

*Stark, BC & Bryant, L; Themistocleous, H; den Ouden, D-B; Roberts, A (2021). Best Practice Guidelines for Reporting Spoken Discourse in Aphasia and Neurogenic Communication Disorders. *Doi: 10.1080/02687038.2022.2039372.* Visit https://osf.io/y48n9/ for updates on the project.

Category	Item Number	Reporting Standard	Included (Mark 'x')	Page Number(s)
Information about the discourse sample	1	Define "discourse"	x	
	2	Define "utterance" (or other unit, e.g., turn unit)	x	
	3*	Number of words in sample		
Information about how the discourse sample was collected	4	Describe elicitation task	x	
	5	Exact instructions used to elicit discourse sample	x	
Information about the persons included in the collection of the discourse sample	6	Demographic information about primary speaker [the person whose discourse is of interest]	x	
	7	Information about the primary speaker's neurological condition	x	
Methodology and rater agreement	8	Inter-rater reliability for each analyzed variable/measure	x	
	9	Reliability statistics used	x	
	10	Details on the number (percentage) of files used for determining reliability/agreement	x	
	11*	Reliability (point to point agreement) for transcription (orthographic or other)		
Analysis	12	Type of transcription (e.g., orthographic, phonetic)	x	
	13	Detailed description of any perceptual rating scale used, including providing a copy of the scale if not previously published	n/a	
	14	Details of the annotation system, formal (e.g., CHAT) or informal (created by the clinician/examiner)	x	
	15	Whether transcription was verbatim (e.g., including all behaviors such as fillers) or whether information was excluded in the transcription process.	x	
	16	Completeness of transcription (full, partial, transcribing errors only)	x	
	17*	Details of any software used for transcribing/annotating/generating data (e.g., SALT, CLAN, ELAN)	x	
	18*	Who/what transcribed the sample (by a human, by a machine/software, hybrid human and software)	x	
Information about the individual discourse variables/behaviors reported	19	What is being used as primary outcome measure(s) (e.g., linguistic information, speech information, etc)	n/a	
	20*	Theoretical rationale for selecting variable/behavior/outcome measure(s)		
	21	Operational definition for each variable/behavior/outcome(s)	x	

Note. Asteriks denote RECOMMENDED standards. All others are NECESSARY. Per a priori set criteria, NECESSARY reporting items are those noted as "highly" or "extremely" necessary by >70% participants in Round 3 of the expert consensus process. RECOMMENDED items are those recommendations noted as "highly" or "extremely" necessary by >65% in Round 2, but which did not reach "highly" or "extremely" necessary by >70% participants in Round 3, or those rated "highly" or "extremely" necessary by >70% in Round 1 that were not carried forward to subsequent rounds of the consensus process.

n/a= not applicable

Supplementary Material 3.

Inter-rater reliability for all variables express as two-way random effect intraclass correlations (ICC) for the videoconference group.

Variable	ICC	95% Confidence interval	Standard error measurement
Duration (sec)	.995	.979 – .999	10.77
TU_{total}	.976	.902 – .994	.48
Total number of utterances	.953	.812 – .988	4.07
Words per minute	.979	.914 – .995	12.11
Mean length of utterances (words)	.751	-.003 – .983	1.67
Verbs per utterance	.948	.789 – .987	.09
Noun/verb ratio	.990	.961 – .998	.34
Propositional density	.989	.956 – .997	.009
MATTR	.915	.657 – .979	.006
Reformulation	.940	.758 – .985	1.76
Repetition	.837	.345 – .960	3.31

TU = Thematic Unit ; MATTR = Moving-Average Type-Token Ratio

Inter-rater reliability for all variables express as two-way random effect intraclass correlations (ICC) for the in-person group.

Variable	ICC	95% Confidence interval	Standard error measurement
Duration (sec)	.991	.976 – .997	7.85
TU_{total}	.910	.751 – .967	1.22
Total number of utterances	.864	.625 – .951	8.04
Words per minute	.985	.958 – .995	11.48
Mean length of utterances (words)	.685	.129 – .886	1.78
Verbs per utterance	.934	.819 – .976	.12
Noun/verb ratio	.993	.982 – .998	.56
Propositional density	.987	.963 – .995	.01
MATTR	.965	.902 – .987	.008
Reformulation	.876	.658 – .955	1.18
Repetition	.907	.744 – .966	1.75

TU = Thematic Unit ; MATTR = Moving-Average Type-Token Ratio