Montreal–Toulouse Language Assessment Battery: Evidence of criterion validity from patients with aphasia

Karina Carlesso Pagliarin a,⁎, Karin Zazo Ortiz b, Simone dos Santos Barreto c, Maria Alice de Mattos Pimenta Parented, Jean-Luc Nespolous e, Yves Joanette f, Rochele Paz Fonseca g

a Federal University of Santa Maria, Santa Maria, Brazil
b Federal University of Sao Paulo, Sao Paulo, Brazil
c Federal University Fluminense, Rio de Janeiro, Brazil
d Federal University ABC, Sao Paulo, Brazil
e Université Toulouse II - Le Mirail, Toulouse, France
f Université de Montréal, Montréal, Canada
g Pontifical Catholic University of Rio Grande do Sul, Porto Alegre, Brazil

A R T I C L E   I N F O
Article history:
Received 9 February 2015
Received in revised form 28 July 2015
Accepted 29 July 2015
Available online 31 July 2015

Keywords:
Language tests
Aphasia
Assessment
Adults
Stroke

A B S T R A C T
Background: The Montreal–Toulouse Language Assessment Battery — Brazilian version (MTL-BR) provides a general description of language processing and related components in adults with brain injury.
Objective: The present study aimed at verifying the criterion-related validity of the Montreal–Toulouse Language Assessment Battery — Brazilian version (MTL-BR) by assessing its ability to discriminate between individuals with unilateral brain damage with and without aphasia.
Methods: The investigation was carried out in a Brazilian community-based sample of 104 adults, divided into four groups: 26 participants with left hemisphere damage (LHD) with aphasia, 25 participants with right hemisphere damage (RHD), 28 with LHD non-aphasic, and 25 healthy adults.
Results: There were significant differences between patients with aphasia and the other groups on most total and subtotal scores on MTL-BR tasks.
Conclusions: The results showed strong criterion-related validity evidence for the MTL-BR Battery, and provided important information regarding hemispheric specialization and interhemispheric cooperation. Future research is required to search for additional evidence of sensitivity, specificity and validity of the MTL-BR in samples with different types of aphasia and degrees of language impairment.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

The left hemisphere (LH) is traditionally considered dominant for language, and damage to this region is therefore associated with aphasia. Aphasia is characterized by a reduction or impairment in the ability to correctly process language form, structure, content, meaning, use or function [44].

Twenty to 50% of patients are diagnosed with aphasia in the acute period of ischemic stroke [15,23,27,46], and approximately 50% of these individuals exhibit some improvement in the first two years following the lesion. Such improvements may be observed starting on post-stroke day 10 [27].

Given the high likelihood of language impairment following strokes, patients must undergo comprehensive language assessment procedures which can identify dissociations between impaired and unimpaired processes. There are a number of instruments designed to provide a detailed assessment of linguistic competence (e.g. Boston Diagnostic Aphasia Examination, [24]; Western Aphasia Battery, [33]; Aachener Aphasia Test, [60]), and these have been adapted for use in a number of different languages [35,47,52,54]. These tests allow for the identification of associations and dissociations between key components of language, and can be used to classify subtypes of aphasia based on findings regarding impairments in comprehension, expression, naming, repetition, reading and writing.

These assessment batteries have been investigated as to their validity in assessing clinical samples, generally by comparative studies of patients with aphasia and neurologically healthy participants [35,38]. Some studies, however, have used patients with right hemisphere damage (RHD) as a control for individuals with LH Damage (LHD) [40,57]. Although patients with RHD may display some functional language impairment (pragmatic competence), they still tend to outperform patients with aphasia on language assessment tasks which do not involve paralinguistic or metalinguistic processes [57,59]. The use of clinical control groups also attests to the specificity of the assessment instruments studied, as it excludes that possibility that the group...
differences detected are due to the presence of brain damage in general rather than the location of the lesion. The use of this method may lead to a reduced rate of false positives and increase clinical validity [21]. When investigating language assessment tools, as well as including a clinical control group with RHD and neurologically healthy participants, it is also important to assess individuals with aphasia and those with LHD but no language impairment. Studies have shown that patients with LHD but no aphasia are very talkative, producing more words than normal subjects [2]. As such, these patients may display impairments on verbal tasks such as directed interviews. Furthermore, even patients who do not display clinical signs of aphasia may develop latent aphasia [56]. Given these findings, comparative studies may be especially important in evaluating the diagnostic accuracy of language assessment instruments. Criterion validity reflects accuracy with which a battery determines a person has a deficit from those with intact communication abilities [39].

One of the greatest difficulties reported by health professionals who work with adults with language impairment in Brazil is the lack of standardized language assessment instruments validated for use in Brazilian Portuguese [45]. The reality of Brazilian speech clinic is the use of translated tests and informal assessments of language, because there are no commercialized instruments.

The Montreal–Toulouse Language Assessment Battery – Brazilian version (MTL-BR), originally known as the Protocole Montréal–Toulouse d'examen linguistique de l'aphasie (MT-86; [41]) has been adapted in order to fill this gap in language assessment [28]. The aim of the MTL-BR is to provide diagnostic information regarding language disorders associated with LHD. Although it does not aim to classify aphasia according to subtypes or severity, it may contribute to the clinical evaluation of such variables [45].

Although traditional instruments for language assessment have made great contributions to the literature, most were originally developed in the English language [29]. The MTL-BR is the only assessment tool originally developed in a Latin language (French) which is also able to assess all possible relationships between linguistic inputs and outputs, as well as praxis and arithmetical skills. Although praxis and arithmetical tasks are not part of language assessment, they were included in the battery due to the high rate of comorbidity between aphasia, apraxia and acalculia.

The adaptation of aphasia assessment instruments provides a significant contribution to the study of language impairment following acquired brain damage. Furthermore, crosslinguistic research allows for the identification of universal language features, and can greatly contribute to the enhancement of psycholinguistic models [35].

An important part of crosslinguistic adaptation of assessment instruments is the search for evidence of the validity and reliability of adapted instruments [8,29,48]. Some psychometric properties have been established for MTL-BR [43] like concurrent validity (correlation between 0.33–0.71) and reliability (test–retest mean 0.52, Cronbach's alpha between 0.79 and 0.90). Therefore, the present study aimed to obtain the criterion validity of the MTL-BR Battery by comparing the performance of clinical groups (patients with RHD, LHD with aphasia and LHD without aphasia) to a control group with no brain damage.

2. Method

2.1. Participants

A total of 104 adults divided into three clinical groups and one healthy control group took part in the study. The clinical group was composed of 26 participants with LHD and aphasia (aphasia group — Group 1). The clinical control groups consisted of 25 patients with RHD (Group 2) and 28 individuals with LHD non-aphasic (LHDna — Group 3), while a third control group was comprised of 25 neurologically healthy participants (control group — Group 4). Participants were matched for age (total sample: M = 58.50, SD = 12.39) and number of years of education (total sample: M = 9.98, SD = 5.76). The groups did not differ by age or education.

Inclusion criteria consisted of being first language Brazilian Portuguese speakers (monolinguals) over 19 years of age, right handed according to the Edinburgh Handedness Inventory [5,42], having no uncorrected sensory deficits (visual or auditory), no symptoms of depression prior to stroke and no current or prior abuse of alcohol or drugs (self-report, [20]).

Control participants also had no history of neurological illnesses as well as no signs of dementia (according to the Mini Mental State Examination — MMSE, adapted by [9]; with cut off points by [34]); and no symptoms of depression (Geriatric Depression Scale (15 points) — GDS-15 [161], adapted by [11]; scores below 4). It is important to highlight that GDS-15 was administered to all participants in order to obtain the same measure for all groups [16,37].

All clinical groups had been previously diagnosed with ischemic unilateral strokes (one, or at most two, strokes in the same hemisphere) confirmed by neurological examination and computer tomography and/or magnetic resonance imaging. All patients in the aphasia group had been previously diagnosed with aphasia by a speech pathologist with expertise in the assessment and treatment of acquired language disorders, using non structured strategy to evaluate auditory comprehension, naming performance, reading, writing, repetition, disfluencies,agrammatism, paraphasias and other aphasic manifestations observed in language examination. Exclusion criteria consisted of bilateral stroke and head trauma.

The LHDna group was composed only of patients with scores below –2.0 standard deviations in the oral language assessment subtests of the Brazilian Brief Neuropsychological Assessment Instrument NEUPSILIN [18,19]. This cutoff was based on the normative data for the NEUPSILIN [19]. This instrument consists of 32 subtests which assess cognitive functions such as temporal and spatial orientation, attention, perception, memory, mathematical abilities, language, praxis, problem solving and executive functions. It is the only brief instrument to evaluate oral language impairments in addition to other neuropsychological deficits. The oral language subtests of the NEUPSILIN evaluate naming (2 objects and 2 pictures), repetition (8 words and 2 non-words) and automatic speech (counting aloud from one to ten and naming the months of the year, in order), as well as the auditory comprehension of words and sentences. These tests corroborated the results of the speech pathology assessments, which found no evidence of aphasia in any of the participants included in the LHDna group.

2.2. Procedures

The present study was approved by the university Research Ethics Committee (protocol number 04908/09). Participation was voluntary and participants were not paid, and all participants or their legal representa-tives provided written informed consent prior to participation.

Clinical control groups were recruited from public and private hospitals in the state of Rio Grande do Sul (RS), or through doctor referral. The aphasia group was recruited from speech therapy clinics, where they were undergoing rehabilitation for aphasia. All participants had been previously assessed and diagnosed with aphasia, and were in the chronic stage of stroke. Members of the control group were selected from university, work and community settings.

Assessments were conducted during a single session lasting approximately 90 min. If participants became fatigued during the session, the assessment was interrupted and resumed after one week. All participants were administered a sociodemographic and health questionnaire. These instruments were administered by five Speech pathologists and five psychologists, all of whom had been trained in the administration of the MTL-BR.

The MTL-BR was scored by a single researcher (first author) to ensure consistency, and 20% of cases were analyzed in parallel by a speech...
pathologist who was blinded to patient groups. Interrater reliability was assessed using the percentage of exact agreement between raters, and was found to be consistently between 80% and 100%.

Language was assessed using the MTL-BR Battery [45], which contains 22 subtests (41 scores) assessing language processes involved in communication, oral expression and comprehension (words, sentences, text and discourse), reading (words, sentences and text), writing (words, sentences and narratives), repetition, naming, praxis and mathematical skills. The tasks in the MTL-BR Battery are shown in Appendix A.

On most tasks, correct answers receive a score of one, except in the directed interview, automatic speech, spoken and written naming, auditory comprehension and written narrative. In the directed interview, written and spoken naming, and oral and written narratives, wrong answers receive a score of zero, incomplete or “almost correct” responses (in the naming task, for instance) receive a score of one, while correct answers are given a score of two. Automatic speech is assessed based on form (phonemic errors) and content (omissions). In the two tasks involving a narrative, the number of words produced is assessed, as is the number of information units (IU) present (microstructures): bank/bakery, robbery, robbers, police officer, car, running, waiting, calling, person and money/gun. Patients receive one point for each IU in the answer. Three main scene elements are also assessed (macrostructures): the robbery itself (main scene), someone waiting for the robbers and someone calling the police (1 point each).

2.3. Data analysis

Quantitative data were compared using one-way ANOVA, while the sex distribution in each group was assessed using a Chi-squared test. Between-group comparisons of data were carried out using One-Way ANOVA (after normalizing the distribution curves). Depression scores were entered as a covariate in a one-Way ANCOVA, and Bonferroni post-hoc tests were performed. Analyses were conducted using the SPSS Software, version 20.0 for Windows (SPSS Inc., Chicago, USA).

3. Results

Participants’ sociodemographic and clinical data are displayed in Table 1. As displayed in Table 1, groups were only significantly different in MMSE and depression scores. These findings were to be expected, as neurological disorders often have impairments in the cognitive and language abilities assessed by the MMSE. Furthermore, patients with neurological conditions may have difficulties in social and communicative tasks, which may lead to significant symptoms of depression [7].

Mean scores for each group on the 22 MTL-BR tasks are displayed in Table 2. The univariate analyses showed significant group differences in all MTL-BR tasks analyzed, except for auditory word comprehension. Post-hoc analyses suggested that the aphasia group performed worse than controls in all tasks, and that some scores differed between patients with aphasia and those in the LHDna and RHD groups.

4. Discussion

The present results suggest that the MTL-BR battery has adequate criterion validity. Most total and subtotal scores differed between individuals with and without aphasia; that is, scores on the MTL-BR can be used to detect language impairment associated with damage to the dominant hemisphere for this function. Therefore, the MTL-BR battery appeared to achieve its diagnostic goal.

The only task in which group differences were not observed was the auditory word comprehension task. This task simply requires the auditory comprehension of a single word, and the selection of a corresponding image among five distractors. Since most participants with aphasia did not have severe comprehension deficits, that is, difficulties understanding single words, the task may not have been sensitive enough to discriminate between the patients with and without aphasia included in the present sample.

As expected, the aphasia group performed worse than control participants on the MTL-BR tasks.

Patients with aphasia also tended to obtain lower scores than the clinical control groups, corroborating the results of other studies in the literature [22,25,35,40,51]. The data displayed in Table 2 show that some participants in the LHDna and RHD groups performed similarly to individuals with aphasia. These results corroborate Between-group comparisons of data were carried out using One-Way ANOVA previous findings that unilateral brain damage may impair performance even in the absence of aphasia [56].

No differences were found between patients with aphasia and those with RHD in the oral and written narrative tasks (IU and scene elements in both tasks, and the number of words produced in the oral narrative), or in the oral comprehension subtest. These results may be explained by the importance of the RH for pragmatic and semantic processing [30], and for complex discourse [50]. The oral text comprehension task also failed to distinguish patients with aphasia from those with RHD. However, this task relies heavily on attention and working memory processes, which are often impaired in individuals with RHD [59].

There were no significant differences between patients with RHD and control subjects on any scores on the MTL-BR. Interestingly, this pattern was evident even in tasks which were expected to demand more RH involvement, such as narratives (oral and written), semantic verbal fluency and oral text comprehension. One possible explanation for this finding is that these tasks do not contain paralinguistic or metalinguistic features which would require RH participation, such as inference processing and fluency tasks with low-productivity criteria [31]. Additionally, several studies have found that only approximately 50% of unselected patients with RHD exhibit communication disorders [4, 10,32]. As such, in the present sample, patients with RHD may not

Table 1
Socio-demographic and clinical data for participants in each group.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Stroke</th>
<th>RHD</th>
<th>LHDna</th>
<th>Aphasia</th>
<th>Control</th>
<th>F/χ²</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>59.3 ± 14.2</td>
<td>57.7 ± 12.5</td>
<td>60.7 ± 10.5</td>
<td>56.6 ± 12.5</td>
<td>.538</td>
<td>.657</td>
<td></td>
</tr>
<tr>
<td>Education</td>
<td>9.1 ± 5.0</td>
<td>10.4 ± 5.6</td>
<td>9.7 ± 5.6</td>
<td>10.1 ± 6.1</td>
<td>.261</td>
<td>.853</td>
<td></td>
</tr>
<tr>
<td>Sex female/male</td>
<td>12/11</td>
<td>13/15</td>
<td>8/18</td>
<td>14/11</td>
<td>3.459</td>
<td>.016</td>
<td></td>
</tr>
<tr>
<td>MMSE</td>
<td>25.2 ± 3.3</td>
<td>25.3 ± 3.0</td>
<td>20.4 ± 6.6</td>
<td>27.6 ± 2.4</td>
<td>.000</td>
<td>.990</td>
<td></td>
</tr>
<tr>
<td>GDS-15</td>
<td>.7 ± .8</td>
<td>.4 ± .8</td>
<td>.7 ± .8</td>
<td>.1 ± .3</td>
<td>3.616</td>
<td>.009</td>
<td></td>
</tr>
<tr>
<td>Time since stroke (months)</td>
<td>34.6 ± 40.0</td>
<td>27.9 ± 32.5</td>
<td>41.5 ± 45.0</td>
<td>–</td>
<td>.696</td>
<td>.502</td>
<td></td>
</tr>
</tbody>
</table>

Note. RHD = right hemisphere damage; LHDna = left hemisphere damage non-aphasic; MMSE = Mini Mental State Examination; GDS-15 = Geriatric Depression Scale; χ² = Chi-square test was used to analyze the sex variable; F = F-statistics, M = mean; SD = standard deviation.

* = p ≤ .05.
** = p ≤ .001.
have developed communication difficulties severe enough to interfere with performance on MTL-BR tests.

Participants in the LHDna group generally performed worse than control participants, although between-group differences were only significant for the phonemic verbal fluency task, which is mediated by verbal (phonemic) processes associated with the LH as well as by executive components [26,55]. It is possible that the general presence of brain damage influenced performance in the LHDna group.

Although the present study produced promising evidence of the criterion validity of the MTL-BR battery, a few limitations deserve consideration, such as the exclusive assessment of patients with ischemic strokes, since individuals with hemorrhagic and bilateral strokes, or those with severe aphasia, were excluded from the sample. The aphasic group indicates a specific group of patients although they had reasonable cognitive function and good comprehension but they were severe enough to warrant treatment. Additionally, the time since stroke varied widely among participants. Over 61% of patients were assessed at least six months after the lesion, and half these individuals had suffered a stroke over two years prior to assessment. This data must be considered when assessing stroke patients, as spontaneous recovery of language abilities often occurs in the first year post-stroke [36]. However, most patients in the present study were assessed after this period. Another point is that it was not possible to blind the assessor. Future research should consider these limitations, as well as investigate other psychometric properties of the MTL-BR battery.

The MTL-BR was developed to provide a standardized tool for language assessment in the Brazilian population. All tasks contributed to the detection of aphasia, and the comparison between control groups and patients with previously diagnosed language impairments allowed for the evaluation of the criterion validity of the MTL-BR. The fact that significant differences were found between patients with aphasia and controls on 40 out of 41 scores on the MTL-BR confirms the criterion validity of the instrument.
5. Conclusion

The MTL-BR may contribute to patient rehabilitation by guiding intervention strategies, and may also provide important information for caregivers and health workers caring for patients with language impairment. However, additional studies are still required to verify whether the MTL-BR may be used to monitor the recovery of language abilities (follow-ups) and to assess rehabilitation outcomes.

Acknowledgments

The authors thank CNPq (grant number: 141271/2010-8) and CAPES (grant number: 12247/12-0), whose scholarships contributed to the development of this manuscript.

Appendix A

1. Directed interview: consists of 13 open questions asked by the researcher (e.g. How old are you? What is your birth date? How is the weather today?).
2. Automatic speech: the responder is asked to count aloud from one to twenty, to name the days of the week in order, and to sing “Happy Birthday”.
3. Auditory comprehension: the task contains five written words and 14 simple and complex sentences, for a total of 19 items. The examiner reads each item aloud, and the participant is asked to point to the picture corresponding to the test item among five distractors.
4. Oral narrative: the participant is shown a picture of a robbery at a bakery and must write a story about the scene.
5. Written text comprehension: the participant is asked to read aloud 15 written words and five distractors.
6. Sentence copying: the participant is asked to copy a sentence in his own handwriting.
7. Dictation: the examiner reads aloud nine words (real words, pseudowords, regular and irregular words, and foreign idioms) and two sentences, which the participant was asked to write.
8. Repetition: the participant is asked to repeat 33 items, including words, pseudowords and sentences.
9. Reading: the participant was asked to read 33 words aloud, including real words, pseudowords, foreign idioms and sentences.
10. Semantic verbal fluency: the participant is given 90 s to elicit as many animal names as possible.
11. Naming: the participant is asked to name 15 images (12 nouns and 3 verbs) displayed by the researcher.
12. Non-verbal praxis: the participant is verbally instructed by the researcher to conduct six non-verbal gestures.
13. Object manipulation: the participant is asked to point to body parts such as the left or right arm.
14. Written naming: the participant is asked to write the names of 15 verbs displayed by the researcher.
15. Body part recognition and left–right orientation: the participant is asked to point to body parts such as the left or right arm.
16. Written naming: the participant is asked to write the names of 15 images (12 nouns and 3 verbs) displayed by the researcher.
17. Oral text comprehension: The examiner reads a passage to the participant and asks a series of questions about the story, which must be answered orally.
18. Number dictation: the participants are asked to write to dictation six numbers read aloud by the researcher.
19. Number reading: the participant is asked to read aloud six numbers.
20. Written narrative: the participant is shown a picture of a robbery at

References
