EXTRA
Cerebrovascular Diseases

Cerebrovasc Dis Extra 2016;6:27–31
DOI: 10.1159/000445778
Received: December 9, 2015
Accepted: March 23, 2016
Published online: April 20, 2016

1664–5456/16/0061–0027$39.50/0
www.karger.com/cee

Original Paper

Screening for Language Disorders in Stroke: German Validation of the Language Screening Test (LAST)

M. Koenig-Bruhin[a, d] T. Vanbellingen[a] R. Schumacher[b] T. Pflugshaupt[a]
J.M. Annoni[c] R.M. Müri[b, e] S. Bohlhalter[a] T. Nyffeler[a, e]

[a]Neurology and Neurorehabilitation Center, Luzerner Kantonsspital, Lucerne, [b]Division of Cognitive and Restorative Neurology, Department of Neurology, Inselspital, Bern University Hospital, and University of Bern, Bern, [c]University of Fribourg, Fribourg, [d]Spitalzentrum Biel, Biel, and [e]ARTORG Center for Biomedical Engineering Research, Gerontechnology and Rehabilitation Group, University of Bern, Bern, Switzerland

Key Words
Aphasia · Language · Stroke · Screening

Abstract
Background: Screening of aphasia in acute stroke is crucial for directing patients to early language therapy. The Language Screening Test (LAST), originally developed in French, is a validated language screening test that allows detection of a language deficit within a few minutes. The aim of the present study was to develop and validate two parallel German versions of the LAST. ***Methods:*** The LAST includes subtests for naming, repetition, automatic speech, and comprehension. For the translation into German, task constructs and psycholinguistic criteria for item selection were identical to the French LAST. A cohort of 101 stroke patients were tested, all of whom were native German speakers. Validation of the LAST was based on (1) analysis of equivalence of the German versions, which was established by administering both versions successively in a subset of patients, (2) internal validity by means of internal consistency analysis, and (3) external validity by comparison with the short version of the Token Test in another subset of patients. ***Results:*** The two German versions were equivalent as demonstrated by a high intraclass correlation coefficient of 0.91. Furthermore, an acceptable internal structure of the LAST was found (Cronbach's $\alpha = 0.74$). A highly significant correlation ($r = 0.74$, $p < 0.0001$) between the LAST and the short version of the Token Test indicated good external validity of the scale. ***Conclusion:*** The German version of the LAST, available in two parallel versions, is a new and valid language screening test in stroke.

Prof. Dr. med. Thomas Nyffeler
Neurology and Neurorehabilitation Center
Luzerner Kantonsspital
CH–6000 Lucerne (Switzerland)
E-Mail thomas.nyffeler@luks.ch

Cerebrovasc Dis Extra 2016;6:27–31
DOI: 10.1159/000445778

Introduction

Stroke patients with aphasia should receive language therapy within the first days after disease onset [1, 2]. Hence, very early screening for aphasia is important. In the acute stroke phase, however, patients' cognitive capacities are often reduced and administration of comprehensive aphasia diagnostic test batteries is thus not possible. Language screening tests with short administration times are therefore needed. In German, the Aachen Aphasia Bedside Test [3, 4], the Aphasie Schnelltest [5], the Kurze Aphasieprüfung [6], the short version of the Token Test [7], the Aphasie-Check-Liste [8], or the Bielefelder Aphasie Screening [9] have been designed to screen language functions in neurologic patients. These tests include naming, comprehension, repetition, word fluency, reading, and writing tasks. Administration and evaluation time is relatively long, ranging from 15 to 40 min. Recently, a more rapid screening test, the Language Screening Test (LAST), was published in French [10]. It includes confrontation naming of black-and-white pictured objects, repetition, automatic speech, and auditory comprehension of words and sentences. The test can be accomplished within approximately 2 min, and two equivalent parallel versions are available [10]. As a further advantage, it excludes tasks during which premorbid conditions such as dyslexia or illiteracy, and coexisting cognitive deficits such as disorders of attention or executive functions would influence the results [10].

The aim of the present study was to develop and validate two parallel German versions of the LAST in a cohort of acute and subacute stroke patients. We predicted good internal consistency and a uniform structure of the scale. In addition, we expected a significant correlation between the LAST and the short version of the Token Test, indicating good external validity of the scale.

Methods

Test Description

The LAST consists of 15 items in five subtests: (1) naming of five black-and-white drawings presented on one page; (2) oral repetition of one word with 4 syllables and one sentence with 8 words (11 syllables); (3) automatic speech, i.e. counting from 1 to 10; (4) word-picture matching with four orally presented target words depicted within a set of eight pictures, with the four distractor pictures being either visually, semantically or phonologically related to one of the target words, and (5) execution of three verbal commands (simple, semicomplex, complex). A value of 1 is scored for correct answers given within 5 s, otherwise a value of 0 is scored. Two subscores are then obtained by addition, one for production (naming, repetition, and automatic speech; maximum score 8 points) and one for comprehension (word-picture matching and verbal command execution; maximum score 7 points). Total scores below 15 indicate a language deficit that needs further evaluation by a speech language therapist [10]. The French LAST showed a sensitivity of 0.98 and a specificity of 1 [10].

Translation

We aimed at maintaining the linguistic properties of the original version such as word frequency, word and sentence length, number of syllables, and consonant clusters in the German translation, as they have an impact on language performance [11]. The German word frequencies were chosen from the CELEX database [12]. In the subtest 'naming', two stimuli out of five were replaced with lexically matched items. In the repetition part, the German stimuli were consistent with the word and sentence length in syllables, words, and number of consonant clusters of the French original. The subtest 'automatic speech', i.e. counting from

1 to 10, was unchanged. In the word-picture matching subtask, the items 'cigarette' with the semantic distractor 'pipe' and 'eye' with the visual distractor 'fish' were identical to the French version. 'Käse' (cheese) was newly chosen as a phonologically complex distractor for the item 'Hase' (hare) and 'Kanne' (tea pot) was added as phonologically simple distractor for 'Tanne' (fir tree).

Patients and Procedure

One hundred and one patients with stroke, all native speakers of German, were consecutively recruited at the Neurology Department of the Lucerne Kantonsspital, at the Division of Cognitive and Restorative Neurology of the Department of Neurology, University Hospital Berne, and at the Spitalzentrum Biel. Exclusion criteria were psychiatric disorders, dementia, preexisting sensory loss, and disorders of vigilance. The study was approved by the local ethics committees.

All patients were tested with one of the two randomly allocated versions of the LAST in German (50 with version A, 51 with version B). Seventy-eight of the patients were additionally tested with the short version of the Token Test [7], one of the most common screening tests, to analyze external validity. Furthermore, both versions of the LAST were successively administered in another subset of 28 patients to evaluate equivalence of the two versions. The examiners were either physicians or language therapists. This was an important point to us, because in the clinical setting, the screening might be used by health professionals with different backgrounds.

Statistical Analysis

The latter was assessed by calculating the intraclass correlation coefficient (ICC). ICC values should be above 0.75 [13]. Internal validity of both LAST versions was assessed in three steps. First, normal score distribution for each item was explored by analyzing skewness and kurtosis values. In addition, floor and ceiling effects on item level were explored. Scores of 90% correct answers and higher would indicate a ceiling effect, while scores of 10% and below would represent a floor effect [14]. Second, construct validity was analyzed by means of principal component analysis, which explores the underlying dimensions of the LAST. Finally, Cronbach's α was calculated, a measure of internal consistency. A value above 0.70 is an indicator of an acceptable homogeneity of items within the total scale [15]. External validity was examined by using Pearson correlation analyses between the total scores of the LAST and the short version of the Token Test. Scales measuring similar concepts should show at least correlations of $r > 0.60$ (moderate correlation), indicating good convergent validity [16]. Data were analyzed with SPSS (IBM, Chicago, Ill., USA, version 23.0, http://spss.com). For all analyses, a two-sided p value of less than 0.05 was used as the threshold for statistical significance.

Results

We evaluated 101 stroke patients, aged between 33 and 91 years (mean = 71.37, SD = 12.28), 57% of whom were male. Mean time interval between stroke and examination was 11 days (SD = 16). LAST scores for version A (mean = 11.54, SD = 3.61, range 0–15) did not significantly differ from those of version B (mean = 11.92, SD = 3.46, range 2–15; independent samples t test: $t = -0.374$, $p = 0.710$). Moreover, the correlation of LAST-A and -B scores in the subgroup of 28 patients showed that the two versions were equivalent with an ICC of 0.91 (confidence interval 0.81–0.96).

Regarding the internal validity of the LAST on the item level, no floor or ceiling effects were found. In the subtests 'automatic speech' and 'word-picture matching', most of the subjects (88 and 84%) reached maximum scores. These subtests thus appeared to be easier than the others. The maximum total score of 15 was reached by 25.7% of the patients. Principal component analysis revealed that the LAST measures one construct, which is demonstrated by one component with an eigenvalue above 1.0, explaining 57.7% of the variance. Communalities were between 0.32 and 0.67, also indicating substantial common variance. The internal consistency of the 15 items was acceptable (Cronbach's $\alpha = 0.74$).

A strong significant correlation between the LAST and the short version of the Token Test ($r = 0.74$, $p < 0.0001$) was found, indicating good external validity of the scale. Comparing the short version of the Token Test with the LAST subtests, highest correlations were found for repetition ($r = 0.673$, $p < 0.01$) and for verbal command execution ($r = 0.624$, $p < 0.01$), whereas word-picture matching showed the lowest correlation ($r = 0.472$, $p < 0.01$).

Discussion

This study shows that the German version of the LAST is a reliable and valid screening test for language in acute and subacute stroke. The two German versions of the LAST are equivalent. The test has acceptable internal validity and a strong significant correlation between the LAST and the short version of the Token Test indicated good external validity.

Compared to other screening tests validated in German, the LAST has several advantages: it is very brief to administer and avoids the impact of preexisting dyslexia, illiteracy and hand paresis. Furthermore, two equivalent parallel versions are available, which minimizes possible retest effects. Health professionals with different backgrounds can easily use the test, as shown in the study by Flamand-Roze et al. [10]. Finally, the interpretation is straightforward: patients scoring below 15 points need to be referred to a speech and language therapist for an in-depth evaluation.

An acceptable internal consistency of 0.74 for the German version of the LAST was found. Compared to the original study [10], in which they found a value of 0.88, our lower value could be due to the heterogeneity of the sample (including acute and subacute patients), explaining a larger error variance in our scores.

The short version of the Token Test was taken to analyze external validity, since this test is also used as a screening test, identical to the LAST. Not only is the short version of the Token Test a measure of aphasia severity and oral comprehension, it also examines verbal short-term memory [17]. This might explain the relatively high correlation we found between the repetition subtest of the LAST and the Token test.

The present study has some limitations. First, 74% of the 101 patients were diagnosed as having a possible language deficit, i.e. they scored below 15 points in the LAST. The prevalence of acute aphasia after stroke is, however, estimated at 30% [18], so there might be a selection bias in our sample. Indeed, many of our patients were evaluated by language therapists because of suspected communication disorders. Second, the high percentage of correct answers in the subtests 'automatic speech' and 'word-picture matching' might reflect the low level of difficulty of these tasks. Counting from 1 to 10 and identification of frequent words such as 'house' or 'tomato' are easy to accomplish, only patients with severe aphasia are likely to fail in these tasks. Future research might use Rasch analysis in order to identify the distribution of the degree of difficulty of the tasks [19]. Finally, we did not evaluate the interrater reliability. However, this has already been done in the original French study, with excellent interrater reliability of 0.998 [10].

In conclusion, the German versions of the LAST are reliable and valid. They can easily be administered at the bedside in approximately 2 min, which is a clear advantage for its use in emergency rooms, stroke units and subacute care settings.

Acknowledgements

We thank Tristan Laville for the drawings and Dr. C. Flamand-Roze for helpful discussions.

References

1 Godecke E, Hird K, Lalor EE, Rai T, Phillips MR: Very early poststroke aphasia therapy: a pilot randomized controlled efficacy trial. Int J Stroke 2012;7:635–644.
2 Koenig-Bruhin M, Kolonko B, At A, Annoni J-M, Hunziker E: Aphasia following a stroke: recovery and recommendations for rehabilitation. Swiss Arch Neurol Psychiatry 2013;164:292–298.
3 Biniek R, Huber W, Glindemann R, Willmes K, Klumm H: The Aachen Aphasia Bedside Test – criteria for validity of psychologic tests (in German). Nervenarzt 1992;63:473–479.
4 Biniek R, Huber W, Willmes K, Glindemann R, Brand H, Fiedler M, Annen C: A test for the detection of speech and language disorders in the acute phase after stroke. Development and clinical application (in German). Nervenarzt 1991;62:108–115.
5 Kroker C: Aphasie-Schnell-Test AST. Leverkusen, Steiner Verlag, 2000.
6 Lang C, Dehm A, Dehm B, Leuschner T: Kurze Aphasieprüfung. Frankfurt, Swets, 1999.
7 De Renzi E, Vignolo LA: The Token Test: a sensitive test to detect receptive disturbances in aphasics. Brain 1962;85:665–678.
8 Kalbe E, Reinhold N, Ender U, Kessler J: Aphasie-Check-Liste. Köln, Prolog, 2002.
9 Richter K, Wittler M, Hielscher-Fastabend M: Bias – Bielefelder Aphasie-Screening. Hofheim, NAT Verlag, 2006.
10 Flamand-Roze C, Falissard B, Roze E, Maintigneux L, Beziz J, Chacon A, Join-Lambert C, Adams D, Denier C: Validation of a new language screening tool for patients with acute stroke: the Language Screening Test (LAST). Stroke 2011;42:1224–1229.
11 Ivanova MV, Hallowell B: A tutorial on aphasia test development in any language: key substantive and psychometric considerations. Aphasiology 2013;27:891–920.
12 Baayen RH, Piepenbrock R, van Rijn H: The CELEX lexical database (CD-ROM). Linguistic Data Consortium, Philadelphia, University of Pennsylvania, 1993. Cited in: Blanken G, Döppler R, Schlenck KJ: Wortproduktionsprüfung. Hofheim, NAT Verlag, 1999.
13 Butts MM, Michels LC: The sources of four commonly reported cutoff criteria: what did they really say? Organ Res Methods 2006;9:202–220.
14 Warner RM: Applied Statistics: From Bivariate through Multivariate Techniques. Los Angeles, Sage Publications, 2013.
15 Bland JM, Altman DG: Cronbach's alpha. BMJ 1997;314:572.
16 Fleiss JL: Reliability of measurement; in Fleiss JL (ed): The Design and Analysis of Clinical Experiments. New York, John Wiley and Sons, 1986, pp 1–31.
17 Lesser R: Verbal and non-verbal components in the Token Test. Neuropsychologia 1976;14:79–85.
18 Engelter ST, Gostynski M, Papa S, Frei M, Born C, Ajdacic-Gross V, Gutzwiller F, Lyrer PA: Epidemiology of aphasia attributable to first ischemic stroke: incidence, severity, fluency, etiology, and thrombolysis. Stroke 2006;37:1379–1384.
19 Andrich D: Rasch Models for Measurement. London, Sage Publications, 1988.