

Ana Carolina Paiva Bento-Gaz<sup>1</sup>  
 Debora Maria Befi-Lopes<sup>1</sup>

### Keywords

Speech  
 Language and hearing sciences  
 Evaluation  
 Language tests  
 Child language  
 Translating  
 Child

### Descritores

Fonoaudiologia  
 Tradução  
 Testes de linguagem  
 Linguagem infantil  
 Avaliação

### Correspondence address:

Ana Carolina Paiva Bento-Gaz  
 Rua Cipotânea, 51, Butantã, São Paulo  
 (SP), Brasil, CEP: 05360-160.  
 E-mail: carolbentogaz@gmail.com,  
 dmblopes@usp.br

Received: 04/08/2013

Accepted: 02/05/2014

# Adaptation of Clinical Evaluation of Language Functions – 4<sup>th</sup> Edition to Brazilian Portuguese

## *Adaptação do teste Clinical Evaluation of Language Functions – 4<sup>th</sup> Edition para o Português Brasileiro*

### ABSTRACT

**Purposes:** To translate and adapt the Clinical Evaluation of Language Functions – 4<sup>th</sup> Edition (CELF-4) to Brazilian Portuguese. **Method:** One hundred and sixty normal language development school children between the ages of seven and ten, half from public schools and the other half from private schools, both located on the east side of São Paulo. **Results:** CELF-4's translation and adjustment to Brazilian Portuguese language showed equivalence between the original and translated versions, which demonstrates that there were no significant changes in the test's form and content. Cronbach's  $\alpha$  test was used in order to verify CELF-4's subtests internal consistency, in other words, if every subtest measures consistently the evaluated constructors. In this analysis, we observed that by excluding right or wrong items, and problematic items from the pool (those different from the rest of the group), all analyzed subtest presented satisfactory internal consistency, except for the Word Association Task for eight years old. **Conclusion:** Most subtests, as well as the Pragmatic Profile and the Observational Evaluation Scale, were simply translated, dismissing significant adaptations. The alterations performed were due to morphosyntactic and phonological differences between both languages. CELF-4's translated and adapted version to Brazilian Portuguese was able to characterize the language performance in the studied population.

### RESUMO

**Objetivos:** Traduzir e adaptar o *Clinical Evaluation of Language Functions – 4<sup>th</sup> Edition* (CELF-4) para o Português Brasileiro. **Método:** Participaram deste estudo 160 escolares em desenvolvimento normal de linguagem, na faixa etária entre sete e dez anos, sendo 80 deles recrutados em escola pública e 80 em escola particular da Zona Leste da cidade de São Paulo. **Resultados:** Os procedimentos adotados na tradução do teste mantiveram a sua equivalência com a versão original, indicando que não houve mudanças significativas no conteúdo e no formato do teste. Foi utilizado o  $\alpha$  de Cronbach para verificar a consistência interna de cada subteste do CELF-4, ou seja, se cada subteste mede de forma consistente aquele construto que pretende avaliar. Nesta análise, observou-se que, ao se excluir apenas os itens acertados ou errados por todos os sujeitos e os itens problemáticos (aqueles discrepantes dos demais resultados encontrados no grupo) da totalidade da amostra, todos os demais subtestes analisados apresentaram consistência interna satisfatória, exceto o de Associação de Palavras para a faixa etária de oito anos. **Conclusão:** A maioria dos subtestes, assim como o Perfil Pragmático e a Escala de Avaliação Observacional, foram apenas traduzidos, não sendo necessárias adaptações significativas. As alterações realizadas ocorreram principalmente pela diferença morfosintática e fonológica das línguas inglesa e portuguesa. A versão traduzida e adaptada para o Português Brasileiro do teste de linguagem CELF-4 caracterizou a performance de linguagem da população estudada.

Study carried out at the Laboratory for Investigation of Language Development Alterations, Department of Speech-Language Pathology and Audiology, Physical Therapy, and Occupational Therapy, School of Medicine, Universidade de São Paulo – FMUSP – São Paulo (SP), Brazil.

(1) School of Medicine, Universidade de São Paulo – USP – São Paulo (SP), Brazil.

**Financial support:** Coordination for Improvement of Higher Education Personnel (CAPES).

**Conflict of interests:** nothing to declare.

## INTRODUCTION

The purpose of language assessment processes is to collect reliable information that is to be integrated and interpreted to make a judgment or a decision. This information contains the clinical history of the individual assessed, results of formal tests, and information collected from observations of natural situations<sup>(1-4)</sup>.

In regard to children with suspicion of language alterations, the assessment must determine whether the child presents any language alterations, identify the cause of the problem, identify the areas with deficit, describe regularities in the child's language behavior, and decide the conduct to be adopted<sup>(1,5-15)</sup>.

For the assessment process to be effective, it is necessary to use adequate instruments and procedures to verify the language patterns of the child in question. These instruments must meet the conditions of the age range and language possibilities, as well as enable the best possible access to the child's linguistic potential<sup>(16)</sup>.

In Brazil, only two tests that assess children's language are available on the market: the ABFW Child Language Test<sup>(17)</sup> and the Protocol of Behavioral Observation<sup>(18)</sup>. As pointed out by many authors, it is possible to observe the scarcity of formal tests to assess children's language in Brazil<sup>(16,17,19-21)</sup>.

Some authors suggest that the translation and adaptation of instruments available into other languages can assuage this paucity. Moreover, research studies with this purpose can propitiate transcultural studies that compare Brazilian to international findings. In this manner, study objects would be better characterized in the sense that several research centers would use the same instruments<sup>(22-25)</sup>.

Recently, in Brazil, some studies have been carried out with the purpose of translating and adapting formal tests of children's language into Brazilian Portuguese (BP)<sup>(21,26)</sup>. Although these studies have been completed, difficulties are still found in the country to diagnose children with language alterations and also to conduct research in this area, due to the lack of reliable reference standards about the linguistic abilities of children undergoing normal language development, especially for age ranges that surpass 7 years.

## PURPOSE

The objective of this study was to translate and adapt the *Clinical Evaluation of Language Functions – 4th Edition* (CELF-4)<sup>(27)</sup> into BP.

## METHODS

### Ethical aspects

This study was analyzed and approved by the ethics committee of the Clinics Hospital of Universidade de São Paulo's School of Medicine (approval report number 145/10).

It is important to highlight that, before elaborating the project, we contacted Pearson Books, which published the CELF-4<sup>(27)</sup>, with the purpose of obtaining authorization to translate the

instrument into BP. The publishing house and the authors of the test gave us the permission to translate it and apply it only for the purposes of this research study.

In addition, all the individuals recruited agreed to participate in the study by signing an informed consent form.

### Sample

The participants were 160 students undergoing normal language development, ranging from 7 to 10 years of age, recruited from public and private schools in the eastern region of the city of São Paulo. The individuals were divided in four age ranges, totaling 40 individuals per age range, 20 from public schools and 20 from private schools.

The inclusion criteria were absence of language-related complaints or previous speech therapy, and satisfactory school performance, as reported by their teachers. Furthermore, they were required to achieve an adequate performance on a phonology test<sup>(28)</sup> and on the School Performance Test (SPT)<sup>(29)</sup>, used to ensure that the individuals included in the study had no phonological alterations and achieved adequate performances on reading, writing, and arithmetic tasks. All these elements, along with good school performances, guaranteed that the children were undergoing typical language development. The choice for the aforementioned tests was made considering the age range of the group studied. The children had to present performances that met the expectations for their age range and schooling, based on the parameters of each evaluation and on the phonology test, and to fall within the reference standards for their schooling on the SPT. Considering that authors of previous studies<sup>(21,26)</sup> did not find differences between the sexes, this variable was not considered.

### Description of the test

The CELF-4, elaborated by Elisabeth Wiig, Wayne Secord, and Eleanor Semel in 2003, is an instrument used to assess communication and language disorders among students ranging from 5 to 21 years of age. According to the instructions manual, it must be administered individually; as it occurred in our study.

The instrument, which is administered in approximately 1 hour and 30 minutes, has previous editions and is available for speakers of English and Spanish.

The purposes of the test are to identify the existence of language disorders, to describe the nature of the language disorder, to evaluate abilities that underlie language (work memory, speech automation, and phonological awareness), and to evaluate contextual language and communication.

The CELF-4 test is composed of 2 image albums and 2 answer sheets, one for ages 5–8, and another for ages 9–21.

The instrument has 16 subtests that evaluate language subsystems and pragmatic profiles with a scale for observational evaluation, as shown in Chart 1.

The subtests that compose the CELF-4 are grouped in four levels of the evaluation process. Levels 1, 2, and 3 form the constructs or different scores generated by the test, as shown in Chart 2.

## Procedure

The process of translation and adaptation of the test was carried out as follows:

- Literal translation of the CELF-4 test (American English to BP);
- Back-translation;
- Selection of the individuals and application of the CELF-4;
- Analysis of theoretical, semantic, and cultural equivalence – before the application of the test on the individuals selected for the study, the researcher, the person in charge of back-translation (an experienced researcher in the area of children’s language, with a PhD in Speech-Language Pathology and Audiology, and fluent in English), and the supervisor of the present study (an experienced researcher in the area of children’s language, with a PhD in Speech-Language Pathology and Audiology) examined each of the items on the subtests that compose the test, as well as the material used to apply it (in this case, the albums), to verify potential sociocultural discrepancies both in the subtest items and in the graphic material. No discrepancies were found, mainly because the material in question is recommended for children and it is graphically simple. The necessary adjustments pertained to phonological and morphosyntactic aspects, as in any translation. These adjustments were made during the translation process, when the need for them became clear; this happened, for instance, when the words for the phonological awareness subtest were selected. It is important to clarify that the phonemes to be discriminated, as well as their position in a given word, were maintained.

The CELF-4 test, translated and adapted into BP, was administered, in a quiet room in their schools, to all the individuals who met the inclusion criteria.

The answers were corrected and interpreted in accordance with the instructions contained in the examiner’s manual, translated into BP.

The results of the test were classified according to the criteria adopted in the original version.

## RESULTS

The CELF-4<sup>(27)</sup> was translated without difficulties. The original format of the test was not altered, that is, all the items tested were kept along with all application instructions (start, scoring criteria, interpretation of results, and answer sheets). The majority of the subtests as well as the Pragmatic Profile and the Observational Evaluation Scale (OES) were only translated, as significant adaptations were not necessary.

Alterations were necessary in the case of the subtests of Word Structure, Semantic Relations, Sentence Assembly, Paragraph Comprehension, and Phonological Awareness. The changes made were mainly due to morphosyntactic and phonological differences between the English and the Portuguese languages.

## Descriptive statistical analyses and measures of reliability

The data were initially analyzed by descriptive statistics and measures of reliability, with the purpose of verifying whether the items of each scale presented good indices of internal consistency. The indices are used to analyze whether a given scale measures the construct to be analyzed in a consistent manner, that is, whether the items of a scale are consistent and adequate for the purpose in question. Following part of the same procedures used to validate the original material in English, we adopted Cronbach’s  $\alpha$  (for each scale and for each age, separately) as the measure of internal consistency in this study. The option for this type of analysis is considered because this statistical treatment is capable of measuring the internal consistency of the test. There is an important difference between this analysis and the one used in the original version, namely that while we had the data of 40 individuals in this study, in the American material the minimum number is 200 children per age range, which makes the analysis more encompassing. This factor, combined to the fact that the majority of the answers analyzed are binary (correct answer = 1; error = 0), substantially reduces the values of Cronbach’s  $\alpha$ , which can slightly affect this analysis, albeit without invalidating. It is noteworthy that the purpose of this study was not to validate the CELF-4, but to translate it and adapt it to BP. Clearly, in a validation study the number of individuals must be at least four times larger.

The acceptable values of Cronbach’s  $\alpha$  for tests that measure abilities are generally equal to or higher than 0.70. Therefore, we conducted new analyses for the scales that initially presented a lower value than the one mentioned above and removed the problematic items, with the purpose of reaching the minimum cutoff point. This does not mean that these items must be necessarily removed from the test (because of the limitations previously mentioned about sample size and the binary nature of the variables). Instead, they may eventually aid in understanding what types of items can be problematic for the sample that spoke BP in this study, and also indicate which items require special attention at the time of validating the translation and adaptation to BP presented here.

The interpretation of a “problematic item” does not pertain to the absolute difficulty of an item (whether it was easy or difficult for the children); instead, it has to do with whether it is at par with the level of difficulty of a subtest, that is, if it seems to be discrepant in relation to the others.

Owing to the fact that the majority of subtests used binary data, the items that were very easy or very difficult (with correct or wrong answers provided by all the children) did not present any variance and were therefore eliminated, even from the initial analyses.

Upon exclusion of the items with correct answers given by the entire sample, as well as the problematic items, the subtests analyzed presented satisfactory internal consistency, except for Word Association in the case of the 8-year age range (Table 1).

The other subtests that compose the CELF-4 test presented Cronbach’s  $\alpha$  equal to or higher than 0.70, that is, a satisfactory internal consistency.

**Chart 1.** Description of the subtests that compose the *Clinical Evaluation of Language Functions – 4<sup>th</sup> Edition*

|  |
|--|
| Subtests of the Clinical Evaluation of Language Functions – 4 <sup>th</sup> Edition  |
| Concepts by Following Commands (CFC) – 54 items<br>5–12 years<br>Task: The student points to objects in the stimulus book in response to commands given orally.  |
| Word Structure (WS) – 32 items<br>5–8 years<br>Task: The student completes a sentence (cloze procedure) with the target structure.   |
| Sentence Repetition (SR) – 32 items<br>5–21 years<br>Task: The student imitates the sentences presented by the examiner.   |
| Sentence Formulation (SF) – 28 items<br>5–21 years<br>Task: The student formulates sentences about the visual stimulus presented using target words or sentences.  |
| Word Classes (1 and 2) — (WC-1 and WC-2)<br>21 items (WC-1) and 24 items (WC-2)<br>5–7 years (WC-1) and 8–21 years (WC-2)<br>Task: The student chooses two words that are related and describes their relation.  |
| Sentence Structure (SS) – 26 items<br>5–8 years<br>Task: The student points to the figure that illustrates the sentence given.   |
| Expressive Vocabulary (EV) – 27 items<br>5–9 years<br>Task: The student identifies an object, person, or activity portrayed in the stimulus book.  |
| Word Definition (WD) – 24 items<br>9–21 years<br>Task: The student defines a word that was uttered and uses it in a sentence.  |
| Paragraph Comprehension (PC) – 15 items<br>5–21 years<br>Task: The student answers questions about the paragraph presented orally by the examiner. The questions aim for the main idea contained in the paragraph, as well as details and sequential, inferential, and predictive information. |
| Semantic Relations (S Rel) – 21 items<br>9–21 years<br>Task: After hearing a sentence, the student selects two correct options out of four in response to the target question.   |
| Sentence Assembly (SA) – 19 items<br>9–21 years<br>Task: The student produces two sentences that are semantically and grammatically correct with words or a group of words presented visually and orally, contained in the stimulus book.  |
| Phonological Awareness (PA) – 85 items<br>5–21 years<br>Task: The student rhymes words and segments, and mixes and identifies sounds and syllables within words.   |
| Rapid Automatic Naming (RAN) – 3 items<br>5–21 years<br>Task: While the examiner uses a timer, the student names colors, shapes, and combinations of shapes and colors.  |
| Word Association (WA) – 3 items<br>5–21 years<br>Task: The student pronounces words within a specific category during one minute.  |
| Number Repetition (1 and 2) – (NR-1 and NR-2) – 15 items<br>5–21 years<br>Task: The student repeats numbers in the order given or backwards.   |
| Familiar Sequences (1 and 2) – (FS-1 and FS-2)<br>12 items (FS-1) and 8 items (FS-2)<br>5–16 years (FS-1) and 17–21 years (FS-2)<br>Task: The student names the days of the week, and counts backwards and in other orders while being timed.  |
| Pragmatic Profile (PP) – 52 items<br>5–21 years<br>Task: The examiner elicits information from parents and teachers about the student's social language skills.  |
| Observational Evaluation Scale (OES) – 40 items<br>5–21 years<br>Task: Parents, teachers and the student evaluate the student's interaction in the classroom and his/her communication skills.   |

**Chart 2.** Description of the levels of the evaluation process of the *Clinical Evaluation of Language Functions – 4<sup>th</sup> Edition* and scores obtained

|  |  |  |
|--|--|--|
| Level 1  |  |  |
| <i>Identifying the problem and determining eligibility</i>   |  |  |
| Overall Language Score (OLS) – <i>sum of the gross scores from the subtests described below</i>                                  |  |  |
| 5–8 years  | 9–12 years   | 13–21 years  |
| Subtests:<br>Concepts by and Following Commands<br>Word Structure<br>Sentence Repetition<br>SF                                   | Subtests:<br>Concepts by and Following Commands<br>Sentence Repetition<br>Sentence Formulation<br>Word Classes-2 (total)           | Subtests:<br>Sentence Repetition<br>Sentence Formulation<br>Word Classes-2 (total)<br>Word Definition      |
| Level 2  |  |  |
| <i>Describing the nature of the disorder</i>   |  |  |
| Receptive Language Score (RLS)   |  |  |
| 5–8 years  | 9–12 years   | 13–21 years  |
| Subtests:<br>Concepts by Following Commands<br>Word Classes-1/2 (receptive)<br>Sentence Structure                                | Subtests:<br>Concepts by Following Commands<br>Word Classes-1/2 (receptive)<br>Sentence Structure                                  | Subtests:<br>Word Classes-2 (receptive)<br>Semantic Relation<br>Paragraph Comprehension                    |
| Expressive Language Score (ELS)  |  |  |
| 5–8 years  | 9–12 years   | 13–21 years  |
| Subtests:<br>Word Structure<br>Sentence Repetition<br>Sentence Formulation   | Subtests:<br>Sentence Repetition<br>Sentence Formulation<br>Word Classes-2 (expressive)  | Subtests:<br>Sentence Repetition<br>Sentence Formulation<br>Word Classes-2 (expressive)                    |
| Content Language Score (CLS)   |  |  |
| 5–8 years  | 9–12 years   | 13–21 years  |
| Subtests:<br>Concepts by Following Commands<br>Word Classes-1/2 (total)<br>Expressive Vocabulary                                 | Subtests:<br>Word Classes-2 (total)<br>Expressive Vocabulary (9 years)<br>Word Definition (10–12 years)<br>Paragraph Comprehension | Subtests:<br>Word Definition<br>Sentence Assembly<br>Paragraph Comprehension                               |
| Structure Language Score (SLS)   | Memory Language Score (MLS)  |  |
| 5–8 years  | 9–12 years   | 13–21 years  |
| Subtests:<br>Word Structure<br>Sentence Repetition<br>Sentence Formulation<br>Sentence Structure                                 | Subtests:<br>Sentence Repetition<br>Concepts by Following Commands<br>Sentence Formulation   | Subtests:<br>Sentence Repetition<br>Sentence Formulation<br>Semantic Relation                              |
| Level 3  |  |  |
| <i>Evaluating subjacent clinical behaviors</i>   |  |  |
| Working Memory Score (WMS)   |  |  |
| 5–8 years  | 9–12 years   | 13–21 years  |
| Subtests:<br>Phonological Awareness<br>Word Association<br>Rapid Automatic Naming<br>Number Repetition-1<br>Familiar Sequences-1 | Subtests:<br>Phonological Awareness<br>Word Association<br>Rapid Automatic Naming<br>Number Repetition-1<br>Familiar Sequences-1   | Subtests:<br>Word Association<br>Rapid Automatic Naming<br>Number Repetition-1/2<br>Familiar Sequences-1/2 |
| Level 4  |  |  |
| <i>Evaluating language in contexts</i>   |  |  |
| 5–8 years  | 9–12 years   | 13–21 years  |
| Subtests:<br>Pragmatic Profile<br>Observational Evaluation Scale   | Subtests:<br>Pragmatic Profile<br>Observational Evaluation Scale   | Subtests:<br>Pragmatic Profile<br>Observational Evaluation Scale   |

## DISCUSSION

In order not to annul their use and applicability, the process of translation and adaptation of instruments published in other language must be thorough. Moreover, transcultural research studies are possible only if an equivalence between the original and the translated versions of the instrument is ensured.

In this study, the translation and adaptation of the CELF-4 followed a method already endorsed in other studies<sup>(30-33)</sup> and recommended by the World Health Organization<sup>(34)</sup>. Thus, the steps followed were: (1) literal translation; (2) back-translation; (3) analysis of theoretical, semantic, and cultural equivalence; (4) test application; and (5) final version of the translation.

In this manner, the procedures followed for the translation and adaptation of the CELF-4 were compatible with the guidelines described in the literature with respect to the care demanded in the process of translating and adapting a foreign instrument.

While analyzing the equivalence between the translated and the original versions of the CELF-4, we came across difficulties concerning the items that evaluate morphosyntactic aspects of oral language. In the process of translating the items of the TELD-3 – Test of Early Language Development<sup>(11)</sup>, difficulties were found pertaining to those that investigated syntactic and morphological aspects of oral language<sup>(26)</sup>, similarly to the ones we encountered while translating the CELF-4. This seems obvious, considering the structural differences among languages, especially between Germanic languages (as the original language of both tests) and Latin languages (which is the case of BP).

Although changes were necessary in the items that evaluated phonological and morphosyntactic aspects, we can affirm that these adaptations were not significant, as they did not interfere with the internal consistency of the items evaluated and did not modify what the items of the test had the purpose of measuring, as proved by our results.

Considering that the analysis of the effectiveness of an instrument is part of its translation process, the translated instrument has to be administered in its new context for its validity to be analyzed. In the same manner, to consider a test as valid, it needs to be clinically accepted and be able to measure what it purports to evaluate with exactitude.

In this study, we used Cronbach’s  $\alpha$  to verify the internal consistency of each subtest of the CELF-4, that is, whether each subtest consistently measured the construct that it purports to evaluate. In this analysis, we observed that upon excluding the items answered correctly or wrongly by the entire sample, as well as the items considered problematic, all the subtests analyzed presented satisfactory internal consistency, except for Word Association in the 8-year age range. This aspect can be understood as an isolated finding or be attributed to the fact that selecting the students by age range reduced the sample, and also to the fact that only a few items compose this subtest. Larger samples, along with thorough analyses of each item that was excluded in the analysis of reliability, will aid in taking the decision to remove any subtests or subtest items that composed the CELF-4 from the version in BP. Only after applying the test on a large scale will we be able to analyze whether it is necessary to reformulate the translation and adaptation of the subtest items for the purpose of making it more appropriate to the Brazilian context.

At this point, we are not yet able to analyze the sensitivity and specificity of the CELF-4 (translated and adapted to BP), which are fundamental elements for its validation. For this purpose, the test needs to be applied on a large scale, that is, with a much larger number of individuals in different age ranges and from different regions of the country.

It is also worth highlighting that to analyze these parameters, it is important that the version of the CELF-4 translated and adapted to BP must be applied on children with language alterations so as to verify its capability to detect alterations in the aspects of language that it purports to evaluate.

## CONCLUSION

Our purpose in this study was to translate and adapt the CELF-4<sup>(27)</sup> to BP. At the end of the study, we observed that:

1. The translation of the CELF-4 was carried out without difficulties that hindered its viability, and the entire original format of the test was maintained (start, scoring criteria, interpretation of the results, and answer sheets);

**Table 1.** Analysis of the items (sum of scores) and of the internal consistency of the Word Association subtest\*

| Item                          | 7 years       |                |               | 8 years       |                |               | 9 years       |                |               | 10 years      |                |               |
|-------------------------------|---------------|----------------|---------------|---------------|----------------|---------------|---------------|----------------|---------------|---------------|----------------|---------------|
|                               | Public (n=20) | Private (n=20) | Removed Items | Public (n=20) | Private (n=20) | Removed Items | Public (n=20) | Private (n=20) | Removed Items | Public (n=20) | Private (n=20) | Removed items |
| WA 1                          | 214           | 270            |               | 270           | 375            | x             | 284           | 358            |               | 296           | 349            |               |
| WA 2                          | 212           | 249            |               | 265           | 263            |               | 237           | 298            |               | 267           | 300            |               |
| WA 3                          | 100           | 88             | x             | 142           | 163            |               | 169           | 198            |               | 164           | 198            |               |
| Initial Cronbach ( $\alpha$ ) |               | 0.666          |               |               | 0.119          |               |               | 0.737**        |               |               | 0.634          |               |
| Items (n)                     |               | 3              |               |               | 3              |               |               | 3              |               |               | 3              |               |
| Final Cronbach ( $\alpha$ )   |               | 0.688          |               |               | 0.433          |               |               | –              |               |               | –              |               |
| Items (n)                     |               | 2              |               |               | 2              |               |               | –              |               |               | –              |               |

Initial Cronbach: the items answered correctly by all individuals (x) were removed

Final Cronbach: the items answered correctly by all individuals (x) and the problematic items (x) were removed

\*Owing to the fact that there are only 3 items in this scale, values between 0.6 and 0.7 can be considered satisfactory. However, in the case of the 8-year age range, the scale seems poor

\*\*Values considered as satisfactory ( $\alpha \geq 0.70$ )

**Caption:** WA = Word Association

2. The majority of the subtests, as well as the Pragmatic Profile and the OES, were only translated and did not need any significant adaptations;
3. The translated and adapted BP version of the CELF-4 test proved itself adequate to characterize the language performance of the population studied in the abilities evaluated by the test;
4. The majority of the subtests presented satisfactory internal consistency, which means that the subtests that compose the CELF-4 indeed measure what they purport to evaluate;
5. The translated and adapted BP version of the CELF-4 must be applied on a large scale for the purpose of validating the test in BP. Its sensitivity and specificity must also be analyzed.

*\*ACPBG was responsible for data collection, tabulation and analysis as well as manuscript elaboration; DMBL was responsible for the project, study outline, and overall supervision of the stages of manuscript elaboration and writing.*

## REFERENCES

1. Shipley KG, McAfee JG. Assessment in speech-language pathology: a resource manual. 3<sup>rd</sup> Edition. New York: Delmar Learning; 2004. 518 p.
2. Webster RI, Majnemer A, Platt RW, Shevell MI. The predictive value of a preschool diagnosis of developmental language impairment. *Neurology*. 2004;63(12):2327-31.
3. Wetherby AM, Watt N, Morgan L, Shumway S. Social communication profiles of children with autism spectrum disorders late in the second year of life. *J Autism Dev Disord*. 2007;37(5):960-75.
4. Shumway S, Wetherby AM. Communicative acts of children with autism spectrum disorders in the second year of life. *J Speech Lang Hear Res*. 2009;52(5):1139-56.
5. Bloom L, Lahey M. Language development and language disorders. New York: McMillan; 1978. 689 p.
6. Lahey M. Who shall be called language disordered? Some reflections and one perspective. *J Speech Hear Disord*. 1990;55(4):612-20.
7. Boone DR, Plante E. Comunicação humana e seus distúrbios. Porto Alegre: Artes Médicas; 1994.
8. Plante E, Vance R. Selection of preschool language tests: a data-based approach. *Lang Speech Hear Serv Sch*. 1994;28:15-24.
9. Reed VA. An introduction to children with language disorders. New York: McMillan; 1994.
10. Merrel A, Plante E. Norm-referenced test interpretation in the diagnostic process. *Lang Speech Hear Serv Sch*. 1997;28(1):50-8.
11. Hresko W, Reid D, Hammill D. Test of Early Language Development (TELD). 3<sup>rd</sup> edition. Austin, TX: PRO-ED; 1999.
12. Paul R. Principles of assessment. In: Paul R. Language disorders from infancy through adolescence: assessment & intervention. 2<sup>nd</sup> Edition. St Louis, MO: Mosby; 2001. p.20-52.
13. Befi-Lopes DM. Alterações no desenvolvimento da linguagem – princípios de avaliação, diagnóstico e intervenção. In: Limongi SCO (Org.). Fonoaudiologia Informação para a Formação: Procedimentos Terapêuticos em Linguagem. Rio de Janeiro: Editora Guanabara Koogan; 2003. p.1-12.
14. Nelson HD, Nygren P, Walker M, Panoscha R. Screening for speech and language delay in preschool children: systematic evidence review for the US Preventive Services Task Force. *Pediatrics*. 2006;117(2):298-319.
15. Lung FW, Shu BC, Chiang TL, Lin SJ. Efficient developmental screening instrument for 6- and 18-month-old children in the Taiwan Birth Cohort Pilot Study. *Acta Paediatr*. 2008;97(8):1093-8.
16. Befi-Lopes DM. Prova de verificação do vocabulário: aspectos da efetividade como instrumento de diagnóstico [Tese]. São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2002.
17. Andrade CRF, Befi-Lopes DM, Fernandes FDM, Wertzner HF. ABFW: teste de linguagem infantil nas áreas de fonologia, vocabulário, fluência e pragmática. Barueri: Pró-Fono; 2004.
18. Zorzi JL, Hage SRV. PROC – Protocolo de Observação Comportamental: avaliação de linguagem e aspectos cognitivos infantis. São José dos Campos: Pulso; 2004. 93 p.
19. Hage SRV. Distúrbio específico do desenvolvimento da linguagem: subtipos e correlações neuroanatômicas [tese]. Campinas: Faculdade de Ciências Médicas da Universidade Estadual de Campinas; 2000.
20. Corrêa LMS, Freitas MC, Lima CMC. Crianças com queixas de linguagem e procedimentos usuais de avaliação de habilidades lingüísticas. *Calidoscópio*. 2003;1(1):43-68.
21. Broggio FTO. Desempenho de crianças típicas de 4 a 8 anos de idade no test of language development primary 3 adaptado para o Português Brasileiro [dissertação]. São Paulo: Faculdade de Filosofia, Letras e Ciências Humanas, Universidade de São Paulo; 2005. 281p.
22. Dumenci L, Erol N, Achenbach TM, Simsek Z. Measurement structure of the Turkish translation of the Child Behavior Checklist using confirmatory factor analytic approaches to validation of syndromal constructs. *J Abnorm Child Psychol*. 2004;32(3):335-40.
23. Excoffier E, Vila G, Taupiac E, Mouren-Simeoni MC, Bouvard MP. Dimensional approach of social behaviour deficits in children. Preliminary validation study of the French version of the Children's Social Behaviour Questionnaire (CSBQ). *Encephale*. 2007;33(4 Pt 1):585-91.
24. Peña ED. Lost in translation: methodological considerations in cross-cultural research. *Child Dev*. 2007;78(4):1255-64.
25. Lauth B, Magnusson P, Ferrari P, Pétursson H. An Icelandic version of the Kiddie-SADS-PL: translation, cross-cultural adaptation and inter-rater reliability. *Nord J Psychiatry*. 2008;62(5):379-85.
26. Giusti E. Performance de crianças falantes do Português Brasileiro no Test of Early Language Development (TELD 3) [tese]. São Paulo: Faculdade de Filosofia Letras e Ciências Humanas, Universidade de São Paulo; 2007. 181 p.
27. Wing EH, Second WA, Semel E. Evaluation of Language Function (CELF-4). 4<sup>th</sup> edition. San Antonio, TX: Pearson; 2003.
28. Wertzner HF. Fonologia. In: Andrade CRF, Befi-Lopes DM, Fernandes FDM, Wertzner HF. ABFW: teste de linguagem infantil nas áreas de fonologia, vocabulário, fluência e pragmática. Barueri: Pró-Fono; 2004.
29. Stein LM. Teste de Desempenho Escolar (TDE): manual para aplicação e interpretação. São Paulo: Casa do Psicólogo; 1994.
30. Guillemin F, Bombardier C, Beaton D. Cross-cultural adaptation of health-related quality of life measures: literature review and proposed guidelines. *J Clin Epidemiol*. 1993;46(12):1417-32.
31. Bullinger M, Afonso J, Apolone G, Leplège A, Sullivan M, Wood-Dauphinee S, et al. Translating health status questionnaires and evaluating their quality: the IQOLA project approach. *International Quality of Life Assessment*. *J Clin Epidemiol*. 1998;51(11):913-23.
32. Ruperto N, Ravelli A, Pistorio A, Malattia C, Cavuto S, Gado-West L. Cross-cultural adaptation and psychometric evaluation of the Childhood Health Assessment Questionnaire (CHAQ) and the Child Health Questionnaire (CHQ) in 32 countries. Review of the general methodology. *Clin Exp Rheumatol*. 2001;19(4 Suppl 23):51-9.
33. Tubert-Jeannin S, Pegon-Machat E, Gremeau-Richard C, Lecuyer MM, Tsakos G. Validation of French version of the Child-OIDP index. *Eur J Oral Sci*. 2005;113(5):355-62.
34. World Health Organization (WHO). Management of substance abuse: process of translation and adaptation of instruments. Geneva: WHO; 2011.