

# Construction of an instrument to assess oral language comprehension of children from 2 to 6 years

Construção de um instrumento para avaliar a compreensão da linguagem oral de crianças de 2 a 6 anos

Construcción de un instrumento para evaluar la comprensión del lenguaje oral en niños de 2 a 6 años

*Tassiana Barbeiro Fragoso de Sá\**

*Ricardo Franco de Lima\**

*Taís de Lima Ferreira Mattar\**

*Sylvia Maria Ciasca\**

## **Abstract**

**Objective:** This work aimed to describe the construction of an instrument to assess the oral language comprehension capability of 2 to 6-year-old children and get validity evidences based on content. **Method:** The research consisted in: gathering and characterizing existing national and international instruments used for the assessment of children oral language comprehension capability; elaborating a new instrument; and analysis of the instrument by five judges. Statistical analysis involved calculation of Percentage of Agreement, of Content Validity Ratio and of Content Validity Index. **Results:** Nineteen instruments were identified, which assess the oral language comprehension capability by using items formed by isolated words, simple and complex orders, and texts. It was subsequently developed an instrument consisting of eight stages containing eight items each. The task consists in the child manipulating figures in accordance with the orders of the evaluator with increasing level of difficulty. The judge's assessment indicated

\* UNICAMP, Campinas, São Paulo, Brazil.

### **Authors' contributions:**

TBFS conceived the work, developed the instrument and collaborated in all stages of writing and correcting the work. RFL oriented the work, conducted the statistical analyzes and collaborated in all the steps of writing and corrections of the work. TLFM co-oriented the study and assisted in writing the paper. SMC collaborated with writing and reviewing the work.

**Correspondence address:** Ricardo Franco de Lima - ricardolima01@yahoo.com.br

**Received:** 01/11/2017

**Accepted:** 09/03/2018

satisfactory results for the calculated indices for the instrument as a whole. Items with unsatisfactory rate were reviewed for inclusion in the final version of the instrument. **Conclusion:** This study described the first step of the construction process of the instrument. Validity evidences based on content were obtained. Further studies will be conducted to obtain other sources of content validity.

**Keywords:** Child language; Comprehension; Psychometrics; Neuropsychological tests; Language tests.

## Resumo

**Objetivo:** Descrever a construção de um instrumento para avaliar a compreensão da linguagem oral de crianças de 2 a 6 anos e obter evidências de validade baseadas no conteúdo. **Método:** A pesquisa consistiu em: levantamento e caracterização de instrumentos nacionais e internacionais destinados à avaliação da compreensão da linguagem oral de crianças; elaboração do instrumento; e análise do instrumento por cinco juízes. A análise estatística envolveu cálculo da porcentagem de concordância, razão de validade de conteúdo e índice de validade de conteúdo. **Resultados:** Foram levantados 19 instrumentos que avaliam a compreensão da linguagem oral utilizando palavras isoladas, ordens simples ou complexas e textos. Posteriormente, foi elaborado instrumento composto por oito fases com oito itens cada. A tarefa consiste na manipulação de figuras pela criança, conforme as ordens do avaliador com nível crescente de dificuldade. A análise dos juízes indicou resultados satisfatórios para os índices calculados para o instrumento como um todo. Os itens com índice insatisfatório foram revisados para compor a versão final do instrumento. **Conclusão:** O estudo descreveu a primeira etapa do processo de construção do instrumento e foram obtidas evidências de validade baseadas no conteúdo. Estudos posteriores serão conduzidos para obtenção de outras fontes de evidência de validade.

**Palavras-chave:** Linguagem infantil; Compreensão; Psicometria; Testes neuropsicológicos; Testes de linguagem.

## Resumen

**Objetivo:** Describir la construcción de un instrumento para evaluar la comprensión del lenguaje oral en niños de 2 a 6 años y obtener evidencias de validez basadas en el contenido. **Método:** La pesquisa se consistió en: levantamiento y caracterización de instrumentos nacionales e internacionales destinados a la evaluación de la comprensión del lenguaje oral en niños; elaboración de un nuevo instrumento; y análisis del instrumento por cinco jueces. El análisis estadístico envolvió cálculo de porcentaje de concordancia, proporción de validez de contenido e índice de validez de contenido. **Resultados:** Fueran identificados 19 instrumentos que evalúan la comprensión del lenguaje oral, utilizando palabras aisladas, ordenes simples o complejas y textos. Posteriormente, fue elaborado un instrumento compuesto por ocho frases con ocho ítems cada. La tarea consiste en la manipulación de figuras por el niño según las órdenes del evaluador, con nivel creciente de dificultad. El análisis de los jueces indicó resultados satisfactorios en lo que se refiere a los índices calculados para el instrumento como un todo. Los ítems con índice insatisfactorio fueran revisados para componer la versión final del instrumento. **Conclusión:** El estudio describió la primera etapa del proceso de construcción del instrumento e fueran obtenidas evidencias de validez basadas en el contenido. Estudios posteriores serán conducidos para obtención de otras fuentes de evidencia de validez.

**Palabras claves:** Lenguaje infantil; Comprensión; Psicometría; Pruebas neuropsicológicas; Pruebas del lenguaje.

## Introduction

Oral language is a complex system of symbols and rules that enables communication involving expression and comprehension<sup>1</sup>. The capacity of expression comprises codifying information organized systematically and conveyed in symbols (speech/sounds) to transmit concepts to a listener. In turn, the capacity of comprehension consists in processing and decoding the meanings of these symbols. Thus spoken information is understood<sup>2</sup>.

From the perspective of cognitive neuropsychology, there are models that describe the diverse elements that constitute the input and output of oral language and its functioning. One example is the neuropsycholinguistic model that deems language expression and comprehension a dynamic, integrated interaction between cerebral functioning and formal and functional processes of prosody, phonology, semantics, morphosyntax, and pragmatics<sup>1</sup>.

Another model is the processing of cognitive information, whose language components are input (input/ sensory records), short- and long-term memory, executive component, and output that engages learned cognitive capacities<sup>2</sup>. There is also the stage model of comprehension, which served as the basis for this study and specifically considers language comprehension<sup>3</sup>. Its authors propose that oral comprehension is not limited to one-way input of information but uses multiple levels involving bottom-up and top-down processing to understand and respond to expression.

Some studies report that impaired language comprehension can be associated with unfavorable therapeutic prognoses, which can lead to difficulties in reading and understanding text<sup>4</sup>. To understand such problems properly requires accurate characterization of the child's use of language resources. Accordingly, input and output must be evaluated. Once the compromised aspect is determined, a more precise diagnosis can be made and a more effective therapeutic plan, using early intervention to minimize aggravation, can be designed<sup>4-6</sup>.

Instruments from psychometric studies can assist in evaluating language in clinical settings, helping to establish a baseline and to manage treatment efficacy<sup>4</sup>. Literature reviews, however, note the dearth of this type of instrument for pre-

school children<sup>6-8</sup>. While some researchers have translated instruments from foreign languages, revising formats, scoring, and instrument implementation, such instruments are not always adapted to Brazilian culture and may fail to consider the needs of Brazilians<sup>9</sup>.

An instrument available in Brazil is the Token Test, which is designed to evaluate oral language comprehension.<sup>10</sup> One study used a condensed version of Token and Raven's colored progressive matrices to evaluate 109 children 7 to 10 years old, in grades 2 through 4<sup>11</sup>. Its authors concluded that the Token Test was capable of evaluating oral comprehension.

An example of an instrument developed in the context of this study is the Computerized Battery of Oral Language (BILOv3), whose validity has been researched<sup>5,12</sup>. BILOv3 was used in conjunction with the Peabody Image Vocabulary Test (TVIP) with 157 students, age 4 to 7. The study found significant correlations between BILOv3 and TVIP scores, indicating the former's convergent validity. Moreover, the instrument could distinguish the subjects' performance by age and school level, indicating criterion validity<sup>5</sup>. In a more recent study<sup>12</sup>, 474 children, age 6 to 11, were evaluated using BILOv3 to assess its criterion validity. The battery was administered collectively and effects of age, gender, and school on performance were obtained.

As available instruments have gaps that preclude their use with young children and are problematic in terms of vocabulary, the present study describes the development and content validation of an oral language comprehension screening instrument for children age 2 to 6.

## Method

This research was approved by the Faculty of Medical Sciences, UNICAMP Research Ethics Committee (Opinion No. 1,241,619 - 2015). It was undertaken in four stages. In the first, a survey of national and international instruments used to assess children's oral language comprehension was conducted to inform an operational description of the construct. In the second, the instruments were characterized in terms of main skills, content, stimuli and examiner task, as set forth in Table 1.

**Table 1.** Model form used to survey instrument characteristics.

Name of article or instrument	
Type of work:	
Author(s):	
Publication year:	
Place of publication:	
Objective:	
Age group:	
Time of implementation:	
Material/type of stimulus:	
Score:	
Evaluated skills:	Tasks used to evaluate skills:

The third step consisted of the development of the proposed instrument. To select the items, a survey of articles in national databases was conducted to identify the most frequent words in the vocabulary of children age 2 or older. The authors focused on recent works from across Brazil citing Portuguese words (nouns, articles, verbs) and selected those cited in more than one study.

In the fourth and final stage, the study's findings were submitted to the analysis of five evaluators: a speech therapist with a doctorate in medical sciences, a speech/language pathologist, a speech therapist with specialization in neuropsychology, a neuropsychologist with a master's degree in health and rehabilitation, and a psychologist/psychometrician with a doctorate in medical sciences.

Each evaluator was provided the instrument's implementation and scoring instructions and a scoring sheet, with each item to be awarded 2 points if adequate, 1 if partially adequate, and 0 if inadequate.

From their responses, a statistical analysis was performed in which a percentage of concordance (PC)<sup>9</sup> was obtained by dividing the number of evaluators rating an item adequate by the total number of evaluators and multiplying by 100. The minimal adequate rate was set at 80%<sup>13</sup>.

A content validity ratio (RVC)<sup>14</sup> was calculated by the following formula:

$$RVC = \frac{n_e - \frac{N}{2}}{\frac{N}{2}}$$

in which represents the number of evaluators rating an item adequate, and indicates the total number of evaluators. The RVC's minimal value for con-

cordance was set at 0.99 to indicate that it would be unlikely that it occurred by chance.

A content validity index (IVC)<sup>9</sup>, which measures the concordance rate among evaluators in regard to each item and the instrument as a whole was calculated by dividing the number of responses rated adequate by the total number of ratings, with a suggested minimal agreement of 0.80. The calculations were performed for each item, for each phase, and for the instrument as a whole.

## Results

In the first stage of the study, 19 instruments were identified that evaluate oral language comprehension: three Brazilian, nine adapted from foreign instruments, one translated, and six in foreign languages (see Table 2).

National and adapted instruments were the Language Development Assessment<sup>15</sup>, Peabody Picture Vocabulary Test<sup>16</sup>, Computerized Battery of Oral Language (BILOv3)<sup>35</sup>, Test of Auditory Vocabulary by Figures USP (TVfusp)<sup>17</sup>, Test of Language Development Primary (TOLD-P3) adapted for Brazilian Portuguese<sup>6</sup>, Clinical Evaluation of Language Fundamentals (CELF4) adapted for Brazilian Portuguese<sup>18</sup>, Token Test for Children (TTFC) adapted for Brazilian Portuguese<sup>11</sup>, Cultural Adaptation of the Test of Narrative Language (TNL) for Brazilian Portuguese<sup>19</sup>, Test of Early Language Development (TELD-3), translated into Brazilian Portuguese<sup>20</sup>, Preschool Language Assessment Instrument (PLAI-2), translated into Brazilian Portuguese<sup>21</sup>, Illinois Test of Psycholinguistic Abilities (ITPA), translated into Brazilian

Portuguese<sup>22</sup>, and Behavioral Observation Protocol (PROC)<sup>23</sup>.

The foreign instruments were the Battery for Assessment of Initial Skills for Reading and Writing (BACLE)<sup>24</sup>, Preschool Language Scale (PLS)<sup>25</sup>, Test of Auditory Comprehension and Language (TACL)<sup>26</sup>, Oral and Written Language Scales (OWLS-II)<sup>27</sup>, Reynell Developmental Language Scales (RDLS)<sup>28</sup>, and Test for Reception of Grammar (TROG)<sup>29</sup>.

The national instruments are intended for preschool-age children and focus on the evaluation of oral language comprehension through sentences, and the international instruments for children at least 3 years old. Two instruments are intended solely for the evaluation of the comprehension of isolated words (TVIP, TVfusp). On the other hand, there are more comprehensive batteries that evaluate diverse aspects of language in terms of comprehension, expression, and even such capacities as attention and memory (BACLE, RDLS, PLS, ITPA and CELF4).

It was determined that the content that characterized the stimuli comprised isolated words, simple orders, complex orders and texts, with the printed figure the most common stimulus in most instruments. The tasks consisted in performing motor responses (pointing at or manipulating toys and objects) in response to the evaluator's oral requests.

From data collected in the previous phase, the instrument was further refined and included: (a) preliminary manual with objective, materials, general and specific instructions for implementation, correction, and scoring; (b) a scoring sheet with 64 items (8 phases of 8 items each) Table 3 provides items included in the second phase.

The subjects' task is to point or manipulate the 8 to 16 figures arranged before them at the evaluator's oral request. In regard to the most frequent words in children's oral vocabularies, some

studies found nouns most frequent in children two or older<sup>30,31</sup>; another, verbs<sup>32</sup>; and yet another, adjectives and adverbs in children three or older.<sup>33</sup> Accordingly, the instrument's items are composed of phrases that increase gradually in complexity in terms of length and syntax. In the initial phase, sentences are composed of three elements: verb, article, and noun, for example, "Take the dog". In the second, the sentences consist of four: verb, article, and two distinct nouns, for example, "Take the dog and the apple". In the third, the sentences are also composed of four elements but vary the syntax: verb, article, adjective, and noun, for example "Take the big house". In the fourth and final phase, the sentences consist of eight elements, for example, "Put the car in the house, the cat with the doll, and give me the apple".

Each response to an item is awarded 1 point if correct, 0.5 points if partially correct, and 0 if incorrect. Scores are totaled for each phase and for the instrument as a whole. Score criteria are set forth in the instrument's implementation manual. In all phases, a response is deemed adequate when the subject fulfills the evaluator's request regardless of sequence. Specific criteria for the intermediate rating are provided for each phase, but, in general, it is accorded for partial fulfillment of the evaluator's request. Finally, when the subject fails to meet even these criteria, the response is deemed inadequate.

The results of the evaluators' analyses for each item are found in Table 4. In aggregate, the items in phases one through five attained the following scores: PC = 100%, RVC = 1.0, and IVC = 1.0. In phase six, these indices were PC = 88%, RVC = 0.75, and IVC = 0.88, and in phase seven, PC = 83%, RVC = 0.65, and IVC = 0.83. Finally, in phase 8, the scores were PC = 80, RVC = 0.60, and IVC = 0.80. For the instrument as a whole, they were PC = 94%, RVC = 0.88, and IVC = 0.94.

Table 2. Synthesis of instruments identified in literature.

Instrument	Objective	Age	Comprehension Task	Time	Score/Classification
Assessment of Language Development (ADL)	Classify language development.	1-6 years	Pointing at pictures or manipulating toys in response to oral requests.	20-40 min.	Classification in: spectrum of normal, mild, moderate, or severe disorder.
Battery for Assessment of Initial Skills for Reading and Writing (BACLE)	Assess acquisition of reading and writing skills.	Preschool or children with educational difficulties.	Point at or manipulate figures in response to oral requests.	60 min.	Total score provides descriptive and qualitative classification.p
Computerized Battery of Oral Language (BILOV3)	Assess oral language comprehension.	2-6 years	Point at figures in response to oral requests.	40 min.	Age-referenced score.
Clinical Evaluation of Language Functions (CELF4) adapted for Brazilian Portuguese	Identify presence and type of language disorder and assess underlying language ability.	5-21 years	Respond verbally to oral requests.	90 min.	Age-referenced score.
Oral and Written Language Scales (OWLS-II)	Assess comprehension and expression of oral language.	3-21 years	Point at or manipulate figures in response to oral requests.	45-110 min.	Age-referenced score.
Preschool Language Assessment Instrument (PLAI-2)	Assess communicative capacity.	3-5 years	Point at figures in response to oral requests.	30 min.	Age-referenced score.
Preschool Language Scale (PLS)	Assess comprehension and expression of language, attention, social communication, and vocal development.	2 weeks-7 years	Questionnaires to parents and manipulate figures in response to oral requests.p	45-60 min.	Age-referenced score.
Behavioral Observation Protocol (PROC)	Assess development of communicative and cognitive skills	12-18 months	Analyze a 30-40 minute video.	Time of film and analysis.	Age-referenced score.
Reynell Developmental Language Scales (RDLS)	Assess oral language comprehension.	1-6 years	Point at objects in response to oral requests.	--	Score converted to mental age.
Test of Auditory Vocabulary by Figures USP (TVfusp)	Assess receptive vocabulary.	7-10 years	Point at figures in response to oral requests.	90 min.	Age-referenced score.
Peabody Image Vocabulary Test (TVIP)	Assess receptive vocabulary.	2 ½-14 years	Point at figures in response to oral requests.	20 min.	Age-referenced score.
Test for Reception of Grammar (TROG)	Assess comprehension of oral language.	3-11 years	Point at objects in response to oral requests.	10-20 min.	Age-referenced score.
Illinois Test of Psycholinguistic Abilities (ITPA) translated into Brazilian Portuguese	Assess comprehension of oral language.	2-10 years	Point at figures in response to oral requests.	45-60 min.	Age-referenced score.
Test of Auditory Comprehension and Language (TACL)	Assess comprehension and expression of language.	3-12 years	Point at or manipulate figures in response to oral requests.	20-30 min.	Age-referenced score.
Test of Early Language Development, third edition (TELD-3) adapted for Brazilian Portuguese	Assess comprehension and expression of language.	2 to 7 years and 11 months	Point at figures in response to oral requests.	30 min.	Age-referenced score.
Test of Language Development, Primary (TOLD3) adapted for Brazilian Portuguese.	Assess comprehension and expression of language in regard to semantics, syntax, and phonology.	4-7 years	Point at figures in response to oral requests.	90 min.	Age-referenced score.
Test of Narrative Language (TNL)	Assess comprehension of oral language.	5-11 years	Point at figures or respond after an oral story.	15-25 min.	Age-referenced score.
Token Test for Children (TIFC) adapted for Brazilian Portuguese	Assess comprehension of oral language.	7-10 years	Point at or manipulate figures in response to oral requests.	10-15 min.	Age-referenced score.

**Table 3.** Examples of items in instrument's second phase.

	Order	Responses	Scores		
1	Take the dog and the apple.		0	0.5	1
2	Take the ball and the fish.		0	0.5	1
3	Take the car and the doll.		0	0.5	1
4	Take the house and the apple.		0	0.5	1
5	Give me the cat and the ball.		0	0.5	1
6	Give me the doll and the car		0	0.5	1
7	Give me the dog and the banana.		0	0.5	1
8	Give me the fish and the house.		0	0.5	1
		Nº adequate			

**Chart 1.** Concordance among evaluators.

Phase 1	J1	J2	J3	J4	J5	PC	RVC	IVC
Item 1	2	2	2	2	2	100	1.0	1.0
Item 2	2	2	2	2	2	100	1.0	1.0
Item 3	2	2	2	2	2	100	1.0	1.0
Item 4	2	2	2	2	2	100	1.0	1.0
Item 5	2	2	2	2	2	100	1.0	1.0
Item 6	2	2	2	2	2	100	1.0	1.0
Item 7	2	2	2	2	2	100	1.0	1.0
Item 8	2	2	2	2	2	100	1.0	1.0
Phase 2	J1	J2	J3	J4	J5	PC	RVC	IVC
Item 1	2	2	2	2	2	100	1.0	1.0
Item 2	2	2	2	2	2	100	1.0	1.0
Item 3	2	2	2	2	2	100	1.0	1.0
Item 4	2	2	2	2	2	100	1.0	1.0
Item 5	2	2	2	2	2	100	1.0	1.0
Item 6	2	2	2	2	2	100	1.0	1.0
Item 7	2	2	2	2	2	100	1.0	1.0
Item 8	2	2	2	2	2	100	1.0	1.0
Phase 3	J1	J2	J3	J4	J5	PC	RVC	IVC
Item 1	2	2	2	2	2	100	1.0	1.0
Item 2	2	2	2	2	2	100	1.0	1.0
Item 3	2	2	2	2	2	100	1.0	1.0
Item 4	2	2	2	2	2	100	1.0	1.0
Item 5	2	2	2	2	2	100	1.0	1.0
Item 6	2	2	2	2	2	100	1.0	1.0
Item 7	2	2	2	2	2	100	1.0	1.0
Item 8	2	2	2	2	2	100	1.0	1.0
Phase 4	J1	J2	J3	J4	J5	PC	RVC	IVC
Item 1	2	2	2	2	2	100	1.0	1.0
Item 2	2	2	2	2	2	100	1.0	1.0
Item 3	2	2	2	2	2	100	1.0	1.0
Item 4	2	2	2	2	2	100	1.0	1.0
Item 5	2	2	2	2	2	100	1.0	1.0
Item 6	2	2	2	2	2	100	1.0	1.0
Item 7	2	2	2	2	2	100	1.0	1.0
Item 8	2	2	2	2	2	100	1.0	1.0

<b>Phase 5</b>	<b>J1</b>	<b>J2</b>	<b>J3</b>	<b>J4</b>	<b>J5</b>	<b>PC</b>	<b>RVC</b>	<b>IVC</b>
Item 1	2	2	2	2	2	100	1.0	1.0
Item 2	2	2	2	2	2	100	1.0	1.0
Item 3	2	2	2	2	2	100	1.0	1.0
Item 4	2	2	2	2	2	100	1.0	1.0
Item 5	2	2	2	2	2	100	1.0	1.0
Item 6	2	2	2	2	2	100	1.0	1.0
Item 7	2	2	2	2	2	100	1.0	1.0
Item 8	2	2	2	2	2	100	1.0	1.0
<b>Phase 6</b>	<b>J1</b>	<b>J2</b>	<b>J3</b>	<b>J4</b>	<b>J5</b>	<b>PC</b>	<b>RVC</b>	<b>IVC</b>
Item 1	2	2	2	2	2	100	1.00	1.0
Item 2	1	2	2	2	2	80	0.60	0.80
Item 3	2	2	2	2	2	100	1.00	1.0
Item 4	1	2	2	2	2	80	0.60	0.80
Item 5	1	2	2	2	2	80	0.60	0.80
Item 6	2	2	2	2	2	100	1.00	1.0
Item 7	1	2	2	2	2	80	0.60	0.80
Item 8	1	2	2	2	2	80	0.60	0.80
<b>Phase 7</b>	<b>J1</b>	<b>J2</b>	<b>J3</b>	<b>J4</b>	<b>J5</b>	<b>PC</b>	<b>RVC</b>	<b>IVC</b>
Item 1	2	2	1	2	2	80	0.60	0.80
Item 2	2	2	1	2	2	80	0.60	0.80
Item 3	2	2	1	2	2	80	0.60	0.80
Item 4	2	2	1	2	2	80	0.60	0.80
Item 5	2	2	2	2	2	100	1.0	1.0
Item 6	1	2	2	2	2	80	0.60	0.80
Item 7	1	2	2	2	2	80	0.60	0.80
Item 8	1	2	2	2	2	80	0.60	0.80
<b>Phase 8</b>	<b>J1</b>	<b>J2</b>	<b>J3</b>	<b>J4</b>	<b>J5</b>	<b>PC</b>	<b>RVC</b>	<b>IVC</b>
Item 1	2	2	1	2	2	80	0.60	0.80
Item 2	2	2	1	2	2	80	0.60	0.80
Item 3	2	2	1	2	2	80	0.60	0.80
Item 4	2	2	1	2	2	80	0.60	0.80
Item 5	2	2	2	2	2	100	1.0	1.0
Item 6	2	2	1	2	2	80	0.60	0.80
Item 7	1	2	1	2	2	60	0.20	0.60
Item 8	1	2	2	2	2	80	0.60	0.80

Legend: J–judge/evaluator PC–percentage of concordance, RVC–content validity ratio, IVC–content validity index.



## Discussion

The study's objective was to describe the development of an oral language comprehension assessment instrument for children 2 to 6 and to obtain evidence of its validity based on content. To this end, steps described in the literature on psychometry<sup>13</sup> and procedures based on international studies<sup>34</sup> were followed.

The study's initial survey indicates a dearth of psychometric instruments in speech therapy that are Brazilian, easy to use, and present vocabulary frequently used by children age 2 or older, thus confirming the findings of previous research<sup>6,7</sup>.

In general, uniformity was noted in studies on the paradigm of instruments used to assess oral language, distinguishing themselves in terms of evaluated components, types of stimuli, and instructions. Thus, some instruments evaluate the comprehension of isolated words<sup>16,17</sup> and others of more complex texts<sup>5,19</sup>, while still others go further to probe data on metalanguage and grammar comprehension<sup>29</sup>.

Only one of the 19 identified instruments used oral requests from the evaluator with children expected to respond with point right answer or manipulations of figures arranged before them, task requiring such capacities as attention, visual discrimination of figures-background, praxis, and vocabulary<sup>25</sup>.

As reported in the literature, the rate of language impairment in children increases from age three<sup>35</sup>. Accordingly, this study considers children as young as 2 in the hope that its proposed instrument can aid in assessing early language difficulties.

In constructing the instrument's items, the words most frequently used from the onset of the development of language were used as a basis<sup>33</sup>, since the selected theoretical model<sup>3</sup> indicates that prior knowledge is essential to understanding oral language. Thus it is hypothesized that using words known by the subject could reduce potential disruptions and difficulties during task execution<sup>31</sup>.

Phases and related items considered included implementation time, range, disruption criteria; instrument sensitivity, and correspondence. As research indicates that fatigue may interfere with evaluation<sup>36</sup>, implementation time should be short. It is anticipated that the range of difficulty from the simplest items ("Take the ball") to the more

complex ("Put the banana in the house, the car with the cat, and give me the doll"), can correspond with the subjects' diverse ages and that the scores for each phase can inform the instrument's disruption criteria as described in other studies<sup>25-27</sup>.

All previous instruments used dichotomous criteria (correct/incorrect) deeming a response correct only when the subject performed the task precisely as directed<sup>10,24</sup>. The instrument proposed in this study, however, uses polymorphic criteria (adequate/partially adequate/inadequate), since the literature indicates that this approach can enhance the instrument's internal consistency<sup>37</sup> and sensitivity<sup>38</sup>. Thus its implementation manual describes these criteria and provides examples.

The number of figures arranged before the subject in all cases is between 8 and 16, depending on the phase. Increasing the number of possibilities that the subject can select in responding to the evaluator's request decreases the likelihood that an adequate response occurred by chance<sup>39</sup>.

To enhance effectiveness, it is recommended that the instrument's content should be assessed by between 5 and 16 evaluators<sup>14,34</sup>. Such professionals should have clinical experience, have conducted and published relevant research, and have expertise in the instrument's conceptual structure and construction of questionnaires and scales<sup>14</sup>. As noted, the instrument proposed in this study was assessed by five such evaluators, according to the criteria described herein.

To strengthen the evidence based on the instrument's content, three indices (RVC, PC, and IVC) were used in the statistical analysis of the evaluators' assessment as limitations have been reported in their sole use<sup>9</sup>. In the study's first- five phases, the three indices found adequate parameters. In its sixth phase, RVC found five items below the minimum (0.99)<sup>14</sup>.

The criteria for scoring individual items were the motives for the evaluators' questions in this phase. Initially, in all phases, 0.5 points were awarded if the subject matched the noun or adjective in the request (only the dog or only a small animal, for example). Following the evaluators' suggestions, this criterion has been revised, and subjects will only be accorded points if they take the complete figure with its noun and adjective (big dog, small cat, etc.)

In this same phase, another question arose in regard to the effect of spatial orientation skills

(above/below) on items such as “Put the little cat on top of the house.” As reported in the literature, these concepts are acquired by children beginning at age three<sup>33,40</sup>. Accordingly, children age 2 can be considered as subjects in the first two phases, which do not require these concepts.

In phases 7 and 8, RVC found additional items fell below the reference value with seven items deemed partially adequate in phase 7. The evaluators noted that the items incorporated words (*no, na, com*) that could confuse the child regarding what is requested in “Put the apple in the house and the cat with the dog.” They was suggested that the words be revised to (*ao lado da, beside or junto a, next to*).

Finally, in phase 8, seven items were found partially adequate, with item 7 presenting unsatisfactory values for all three indices. The evaluators raised questions regarding words and spatial concepts noted in previous phases, such as the use of *no* and *na*, citing as an example: “Give me the cat and the banana and put the dog at the foot of the doll.”

Considering the items for each phase, as well as the instrument as a whole, the indices found favorable results according to criteria described in the literature<sup>9,13,14</sup>, and their analysis provides evidence of the instrument’s validity based on content. Moreover, their assessments were deemed relevant and their recommendations are incorporated in the instrument’s current version.

Further research will be undertaken using the updated instrument in a pilot and target sample focused on children with complaints related to the acquisition and development of oral language. The research will further examine evidence of the instrument’s validity based on its internal structure, response process, and relationship with external criteria, as well as other indicators of reliability. While the development and scoring of the instrument’s items conform to the classical theory of tests, further examination of their validity can be complemented by the theory of item response. Based on recent studies of construct validity, it is anticipated that the instrument may help classify a child’s profile for diagnostic purposes and follow-up intervention as needed.

## Conclusion

In accord with its objectives, this study describes the development and construction of an instrument to evaluate the comprehension of oral

language in children. Evidence of validity based on content were obtained in regard to items in each phase and to the instrument as a whole. According to the evaluator’s analysis, some items were not found suitable and modifications were made accordingly to the current version of the instrument. Further studies will be undertaken to examine other sources of evidence of validity for the instrument.

## References

1. Chevrie-Muller C, Narbona J. A Linguagem da Criança: Aspectos normais e patológicos. 2ª ed. Porto Alegre: Artes Médicas; 2005.
2. Sternberg RJ. Psicologia Cognitiva. Porto Alegre: Artes Médicas; 2000.
3. Bishop DVM, Edmundson A. Language-Impaired 4-Year-Olds Distinguishing Transient from Persistent Impairment. *Speech Lang Hear*. 1987; 52(2): 156-173.
4. Crestani AH, Oliveira LD, Vendruscolo JF, Ramos-Souza AP. Distúrbio específico de linguagem: a relevância do diagnóstico inicial. *Rev CEFAC*. 2013; 15(1): 228-237.
5. Joly MCRA, Istome AC. Estudo correlacional da bateria informatizada de linguagem oral (versão 3) com o vocabulário receptivo. *Cad. Pós-Grad Distúrbios do Desenvolvimento da Universidade Presbiteriana Mackenzie*. 2013; 13(2): 8–21.
6. Lindau TA, Lucchesi FDM, Rossi NF, Giacheti CM. Instrumentos sistemáticos e formais de avaliação da linguagem de pré-escolares no Brasil: uma revisão de literatura. *Rev CEFAC*. 2015; 17(2): 656-662.
7. Gurgel LG, Joly MCRA, Plentz RDM, Reppold CT. Instrumentos de Avaliação da compreensão da linguagem oral em crianças e adolescentes: uma revisão sistemática da literatura. *Neuropsicologia Latinoamericana*. 2010; 2(1): 1-10.
8. Pagliarin KC, Oliveira CR, Silva BM, Calvette LF, Fonseca RP. Instrumentos para avaliação da linguagem pós-lesão cerebrovascular esquerda. *Rev CEFAC*. 2013; 15(2): 444-454.
9. Alexandre NMC, Coluci MZO. Validade de conteúdo nos processos de construção e adaptação de instrumentos de medidas. *Cien Saude Colet*. 2011; 16(7): 3061-3068.
10. De Renzi E, Vignolo LA. The Token test: a sensitive test to detect receptive disturbances in aphasics. *Int J Neurol Brain Disord*. 1982; 85(4): 665–678.
11. Malloy-Diniz LF, Bentes RC, Figueiredo PM, Brandão-Bretas D, Costa-Abrantes S, Parizzi AM et al. Normalización de una batería de tests para evaluar las habilidades de comprensión del lenguaje, fluidez verbal y denominación en niños brasileños de 7 a 10 años: resultados preliminares. *Rev Neurol*. 2007; 44(5): 275-280.
12. Reppold CT, Gurgel LG, Parise LF, Serafim A, Dias AS, Joly MCR A. Evidências de Validade de Critério da BILOv3 em Crianças Gaúchas. *Psic Teor Pesq*. 2015; 31(2): 155-162.
13. Pasquali L. Psicometria: teoria dos testes na psicologia e na educação. 5 ed. São Paulo: Vozes; 2013.
14. Pacico JC, Hutz CS. Validade. In: Hutz CS, Bandeira SR, Trentini CM, organizadores. *Psicometria*. Porto Alegre: Artmed; 2015. p.71-84.

15. Menezes MLM. Avaliação do desenvolvimento da linguagem – ADL. Rio de Janeiro; 2004.
16. Capovilla FC, Capovilla AG, Nunes LR, Araújo I, Nunes DR, Nogueira D, Bernat AB. Versão brasileira do Teste de Vocabulário por Imagens Peabody. *Distúrb Comun*. 1997; 8(2): 151-162.
17. Capovilla FC, Prudêncio ER. Teste de vocabulário auditivo por figuras: normatização e validação preliminares. *Aval Psic*. 2006; 5(2): 189-203.
18. Bento-Gaz ACP, Befi-Lopes DM. Adaptation of Clinical Evaluation of Language Functions - 4th Edition to Brazilian Portuguese. *CoDAS* 2014; 26(2): 131-137.
19. Rossi NF, Lindau, TA, Gillam RB, Giacheti CM. Adaptação cultural do Test of Narrative Language (TNL) para o Português Brasileiro. *CoDAS* 2016; 28(5),507-516.
20. Giusti E, Befi-Lopes DM. Performance de sujeitos falantes do Português e do Inglês no Test of Early Language Development. *Pró-Fono Revista de Atualização Científica* 2008; 20(1): 13-18.
21. Lindau TA, Rossi NF, Giacheti CM. Preschool Language Assessment Instrument, second edition, in Brazilian Portuguese-speaking children. *CoDAS* 2014; 26(4),328-330.
22. Bogossian MAD, Santos MJ. Manual do examinador: teste de habilidades psicolinguísticas. Rio de Janeiro: EMPSI; 1977.
23. Hage, SRV, Pereira TC, Zorzi JL. Protocolo de Observação Comportamental - PROC: valores de referência para uma análise quantitativa. *Revista CEFAC* 2012;14(4): 677-690.
24. Pereira RS, Rocha RM. Bateria de avaliação de pré-competências para a aprendizagem da leitura e escrita (BACLE). 3 ed. Rio de Janeiro: Associação Ester Janz; 2013.
25. Miyamoto RT, Hay-McCutcheon MJ, Kirk KI, Housto DM, Bergeson-Dana T. Language skills of profoundly deaf children who received cochlear implants under 12-months of age: a preliminary study. *Acta Otolaryngol*. 2008; 128(4): 373-377.
26. Carrow-Woolfolk E, Allen EA. Test of Auditory Comprehension and Language – TAACL. EUA: PRO-ED; 2014.
27. Carrow-Woolfolk E. Oral and Written Language Scale - Second Edition (OWLS-II) listening comprehension/oral expression manual. EUA: Wewstern Psychological Services; 2011.
28. Anastácio-Pessan FL. Evolução da Nomeação após Fortalecimento de Relações Auditivo-Visuais em Crianças com Deficiência Auditiva e Implante Coclear. [Dissertação]. Bauri (SP): Universidade Estadual Paulista; 2011
29. Bishop, DVM. Test for Reception of Grammar Version 2, TROG-2. Department of Experimental Psychology, University of Oxford. London: The Psychological Corporation; 2003.
30. Silva CTS. Primeiras palavras: um estudo a partir dos inventários de desenvolvimento comunicativo. *Rev Cont Ling*. 2010; 4(1): 9-20.
31. Savoldi A, Ceron MI, Keske-Soares M. Quais são as melhores palavras para compor um instrumento de avaliação fonológica? *Audiol Commun Res*. 2013; 18(3): 194-202.
32. Andersen EML. Representações lexicais subjacentes: verbos e léxico inicial. *ReVEL* 2008; 6(11): 1-31.
33. Prates LPCS, Martins VO. Distúrbios da fala e da linguagem na infância. *Rev Assoc Med Minas Gerais*. 2011; 21 (4): 54-60.
34. American Educational Research Association - AERA, American Psychological Association - APA and National Council on Measurement in Education – NCME. New York; 1999.
35. Mandrá PP, Diniz MV. Caracterização do perfil diagnóstico e fluxo de um ambulatório de Fonoaudiologia hospitalar na área de Linguagem infantil. *Rev Soc Bras Fonoaudiol* 2011; 16(2): 121-125.
36. Pawlowski JT, Marcell C, Bandeira DR. Discutindo procedimentos psicométricos a partir da análise de um instrumento de avaliação neuropsicológica breve. *Psico-USF* 2007;12(2): 211-219.
37. Santos MA, Primi R. Desenvolvimento de um teste informatizado para avaliação do raciocínio, da memória e da velocidade do processamento. *Estudos de Psicologia (Campinas)* 2005; 22(3): 241-254.
38. Neves CE. Experiência de *microrrays* e teoria de resposta ao item [Dissertação]. São Paulo (SP): Universidade de São Paulo; 2010.
39. Couto G, Primi R. Teoria de resposta ao item (TRI): conceitos elementares dos modelos para itens dicotômicos. *Boletim de Psicologia* 61(134):1-15.
40. Ishii C, Miranda CS, Isotani SM, Perissinoto J. Caracterização de comportamentos linguísticos de crianças nascidas prematuras aos quatro anos de idade. *Rev CEFAC* 2006; 8(2): 147-154.