Noun phrase dictation as a writing assessment instrument: a psychometric analysis

Mirian Aratangy Arnaut
Christian Hudelot
Maria Mercedes Saraiva Hackerott
Clara Regina Brandão de Avila

Keywords
Writing
Students
Learning
Evaluation
Speech, Language and Hearing
Sciences

ABSTRACT

Purpose: Prepare a list of noun phrases able to characterize the writing performance of schoolchildren according to school year. Methods: A list of 25 noun phrases intended for dictation was prepared and applied to 275 Elementary School students (2nd to 5th grades). The unidimensionality underlying the proposed items was analyzed using Item Response Theory. Results: The list enabled differentiation between the performance of schoolchildren attending the 2nd and 3rd years and that of students attending the 4th and 5th years. Items on the list presented moderate difficulty level and medium discrimination capacity. Conclusion: The model showed good fit and was able to differentiate the performance of 2nd to 3rd grade students from that of 4th to 5th grade students. The list of noun phrases assessed writing in 4th to 5th grade students more accurately.

RESUMO

Objetivo: Construir uma lista de sintagmas nominais capaz de caracterizar desempenhos segundo o ano escolar. Método: Uma lista de 25 sintagmas nominais foi elaborada para ditado, aplicada a 275 estudantes típicos, do 2º ao 5º ano do Ensino Fundamental I. A unidimensionalidade subjacente aos itens propostos foi analisada por meio da Teoria de Resposta ao Item. Resultados: A lista diferenciou o desempenho dos dois anos iniciais do apresentado pelos finais. Em geral, seus itens mostraram grau moderado de dificuldade e médio poder de discriminação. Conclusão: As análises mostraram ajuste ao modelo proposto e diferenciaram desempenhos quando os anos escolares foram agrupados. A lista avaliou com mais precisão crianças de 4º e 5º anos.

Study conducted at Departamento de Fonoaudiologia, Universidade Federal de São Paulo – UNIFESP – São Paulo (SP), Brazil.

1 Universidade Federal de São Paulo – UNIFESP - São Paulo (SP), Brasil.
2 Université Nice Sophia Antipolis – Nice, França.
3 Universidade Paulista – UNIP - São Paulo (SP), Brasil.

Financial support: Bolsa de Doutorado Sanduíche - CAPES-PDSE – 10557/13-0 e CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico), processo nº 313645/2014.

Conflict of interests: nothing to declare.
INTRODUCTION

Learning of orthographically correct writing is usually assessed through dictation tasks or analysis of spontaneous handwriting of words, phrases, or texts. A survey of the writing evaluation studies published between 1996 and 2005 showed that 28.1% of the 32 analyzed articles presented validity evidence of the material applied and 59.4% of them were conducted by means of dictations, mostly of isolated words\(^{(1)}\). Assessment evidence the hypotheses that children raise about writing, as well as the spelling mistakes they make; however, some types of errors are not identified in the dictation of isolated items.

In addition to orthographic information, which is learned since the beginning of Elementary School, there are other aspects equally relevant to be considered, especially in the beginning of the learning process\(^{(2-4)}\). Throughout this initial period, schoolchildren perceive that words are represented graphically by letters and delimited by blanks, and that they can be combined with others to form sentence\(^{(5)}\). This is the first step to represent connected speech. To this end, it is also necessary to know how to analyze and neutralize certain conflicting aspects: in Brazilian Portuguese, as well as in other languages, syllables present varying duration that adjust at relatively constant intervals; the salience of this information is the stressed syllable, which has louder volume or longer duration than unstressed syllables, characterizing an accentual and non-syllabic rhythm\(^{(6-7)}\).

Therefore, in connected speech, syllables are pronounced in tonal groups, which do not necessarily correspond to a word\(^{(8-11)}\), and may not coincide with the graphic spaces that mark the boundaries of words. When these tonal groups are formed by two or more words, the coarticulation phenomenon occurs. In writing, this phenomenon is known as intervocabulary junction\(^{(12)}\).

Just as they can determine the emergence of word junctions, the phonological representation and processing of tonal groups\(^{(13)}\) based on auditory perceptions\(^{(14,15)}\) can also lead to undue segmentation of written words. Both are considered inaccuracies that should, along with other types of spelling mistakes, disappear throughout schooling. Evidently, the writing of sentences or text, spontaneous or under dictation, will demand abilities other than the orthographic one, namely, bias, semantic, morphosyntactic and memory skills, which also influence this production\(^{(16-19)}\). Moreover, it should not be forgotten that the number of errors decreases throughout schooling. This decrease is most significantly observed between the 3\(^{rd}\) and 4\(^{th}\) grades, which is a period that corresponds to the passage from alphabetic writing to the beginning of orthographic writing\(^{(20-29)}\).

Writing under dictation of isolated words does not enable analysis of possible segmentations or junctions. The conception and analysis of a list of noun phrases aiming to evaluate the writing of schoolchildren in the early years of Elementary School have proven interesting, considering that such list provides not only orthographic evaluation of error at the word level, but also analysis of the writing of tonal groups. Noun phrases are units of meaning that are larger than words and smaller than sentences\(^{(30)}\). Because they are short, they require minimum working memory, present simple morphosyntactic structure, and are easy to understand - characteristics that contribute to the application of dictation to different school years.

Some orthographic writing assessments present evidence of construct validation. However, it is necessary to collect direct psychometric measures on test scores, in the form of validation, reliability and analysis of the items, which is not often found in Brazilian studies\(^{(21,22)}\).

In this study, in order to ensure quality of the written evaluation items, the Item Response Theory (TRI) was used to identify the best test items in the dictation of noun phrases according to the levels of discrimination and difficulty shown in the analysis of writing of students in early Elementary School. This study was guided by the hypothesis that analysis of dictation would demonstrate that noun phrases constitute a good writing evaluation procedure and would show difference in performance between school years. Presentation of the preparation steps of a list of noun phrases for assessment of the writing of Brazilian children attending 2\(^{nd}\) to 5\(^{th}\) grades of Elementary School and verification of the fit of this instrument to the selected school years should confirm this hypothesis. The list of noun phrases also intends to be a clinical instrument of simple application, able to identify developmental characteristics of writing during the first school years, which will enable structured and reliable reproducibility of the orthographic evaluation.

METHODS

Sample selection

Study participants were 275 schoolchildren (150 girls; 54.54%) aged 6 years and 8 months to 11 years regularly enrolled from 2\(^{nd}\) to 5\(^{th}\) grades at eight Elementary Schools in the city of Sao Paulo, randomly selected respecting the proportion of 85% of public schools (4 municipal and 3 state schools) and 15% of private schools (1 school). After selecting and contacting the eight schools, an Informed Consent Form was signed by the respective principals, and teachers were then requested to indicate the children eligible for the study, i.e., those who did not present complaints or indicators of school difficulties, behavioral, neurological and/or sensory disorders. The children were not selected based on their school grades, but on reading task performance. From the teachers’ indication, the children were screened using an oral reading test, which enabled calculation of text reading rate values, with the purpose of ensuring a certain level of automatic word recognition and knowledge of orthographic writing, according to the school year. Using this procedure, those that presented reading rate below the following values for the respective school years were excluded from the survey: 2\(^{nd}\) grade = 47 words per minute (w.p.m.); 3\(^{rd}\) grade = 66 w.p.m.; 4\(^{th}\) grade = 71 w.p.m.; 5\(^{th}\) grade = 91 w.p.m\(^{(23)}\). The sample was distributed as follows: 66 students from 2\(^{nd}\) grade (mean age = 7y and 9m); 76 from 3\(^{rd}\) grade (mean age = 8y and 10m); 59 from 4\(^{th}\) grade (mean age = 9y and 9m), and 74 from 5\(^{th}\) grade (mean age = 10y and 6m).
Procedures

Pilot study

Focusing on the graphic representations of all the phonemes and types of syllables of Brazilian Portuguese, a panel composed of a speech-language pathologist and a linguist prepared a list of 34 noun phrases comprising sequences from three to seven words selected from word banks\textsuperscript{24,25}. In a pilot study conducted with 80 students from 2\textsuperscript{nd} to 5\textsuperscript{th} grades of Elementary School (20 from each school year, 10 from each school network), this list was dictated to gather information on evaluation of the instructions, calculation of the time spent on the test, and ideal number of children per application. From the dictation, statistical analysis of the corpus collected was conducted with application of the Item Response Theory (TRI). This analysis resulted in the exclusion of nine noun phrases that presented very low values of discrimination power or difficulty for the sample\textsuperscript{26}.

Actual study

The list containing the 25 remaining noun phrases was dictated to the 275 study participants throughout the second school term. According to the results of the Pilot Study, participants were divided into groups of maximum 10 children per school year for the application. The dictation was performed in two stages for children attending the 2\textsuperscript{nd} and 3\textsuperscript{rd} grades and in a single stage for those attending the 4\textsuperscript{th} and 5\textsuperscript{th} grades. Each participant received a pencil and a lined paper sheet (with space to register their names, date, name of school, and school grade). The following instruction was given prior to dictation commencement:

\begin{quote}
I will say a small phrase out loud and you will write it down on your sheet of paper. If needed, I will repeat it. If you make a mistake, cross the word and write it again on the side. You must not erase any word.
\end{quote}

Errors were identified and computed. Correctness (1 point) and mistake (zero point) were assigned per word. When more than one mistake was found in the same word, they were all considered. The study was approved by the Research Ethics Committee of Universidade Federal de São Paulo – Escola Paulista de Medicina (no.1768/11). All participating schoolchildren signed an Informed Consent Form (ICF) prior to study commencement.

Table 1.

<table>
<thead>
<tr>
<th>Discrimination</th>
<th>Discrimination Standard error</th>
<th>Difficulty</th>
<th>Difficulty Standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>A festa do dia das mães (Mother's Day party)</td>
<td>0.540</td>
<td>0.125</td>
<td>-2.068</td>
</tr>
<tr>
<td>O ladrão de trens (Train thief)</td>
<td>0.423</td>
<td>0.097</td>
<td>-0.749</td>
</tr>
<tr>
<td>Essa roda da carroça de bois (Ox wagon wheel)</td>
<td>0.659</td>
<td>0.112</td>
<td>-0.220</td>
</tr>
<tr>
<td>A mesa de vidro transparente (Transparent glass table)</td>
<td>0.676</td>
<td>0.125</td>
<td>-0.880</td>
</tr>
</tbody>
</table>

Statistical analysis

Each item on the list of noun phrases was analyzed. For statistical analysis, the Item Response Theory (TRI)\textsuperscript{26} was applied in order to verify the discrimination capacity (very high = 1.69; high = 1.69 to 1.34; medium = 1.33 to 0.65; low = 0.64 to 0.35; very low = 0.34 to 0.01) and difficulty level (extremely difficult = >3.0; very difficult = 3.0 to 2.1; difficult = 2.0 to+1.01; moderate = 1.0 to -1.0; easy = -1.01 to -2.0; very easy = -2.01 to -3.0; extremely easy = <-3.0) of each item. Noun phrases that presented extreme response values, i.e., ≥95% and ≤5% of discrimination and difficulty, respectively, were eliminated.

In order to assess whether discrimination and difficulty levels were invariant throughout the school years, Multi-group Confirmatory Factor Analysis (Invariance Test) was conducted considering the items common to all school years, performed by root mean square error approximation (RMSEA) (>0.06); comparative fit index (CFI)(>0.95); and Tucker-Lewis index (TLI) (>0.90). To test the invariance for the variable “school year”, Group 1 was formed by the grouping of the 2\textsuperscript{nd} and 3\textsuperscript{rd} grades, whereas Group 2 formed by the grouping of the 4\textsuperscript{th} and 5\textsuperscript{th} grades, because the sample of each year alone was not large enough to conduct separate analyses. All statistical analyses were processed using the MPlus 7.11 software.

RESULTS

Results were obtained from the collection of the dictation of noun phrases applied to the randomized sample of 275 Elementary School students.

By means of the item exclusion procedure, eight noun phrases that presented extreme responses were withdrawn from the list, resulting in 17 noun phrases (Table 1).

Of the 17 remaining noun phrases, 15 (88.23%) showed medium and two (11.6%) presented low discrimination capacity. Analysis of the level of difficulty revealed the following numbers of noun phrases and values: one (5.8%), difficult; 13 (76.47%), moderate; two (11.6%), easy; one (5.8%), very difficult.

Distribution of the noun phrases according to level of discrimination was restricted to medium and low discrimination power, whereas the level of difficulty varied along difficult, moderate, easy, and very easy (Table 2).

In the analysis of configurational and scalar invariance of the list of noun phrases, TLI and CFI showed good fit of the model.
to the sample. Groups 1 and 2 could be compared because the model presented stability (Table 3).

Figure 1 depicts the skills closest to +3 (more difficult) and closest to -3 (easier) on the abscissa axis. Higher levels of information indicate the most accurate measure of the construct. The two groups determined by schooling were compared. In this way, the Total Information Curves (TIC) were constructed. Two points can be highlighted in both graphs: 1) the peak of

Table 1. Continued...

<table>
<thead>
<tr>
<th></th>
<th>Discrimination</th>
<th>Discrimination standard error</th>
<th>Difficulty</th>
<th>Difficulty standard error</th>
</tr>
</thead>
<tbody>
<tr>
<td>O caldo de cana gelado (Iced sugar cane juice)</td>
<td>0.671</td>
<td>0.121</td>
<td>-0.942</td>
<td>0.189</td>
</tr>
<tr>
<td>Uma chuteira de futebol vermelha (Red football boots)</td>
<td>0.658</td>
<td>0.135</td>
<td>-1.531</td>
<td>0.280</td>
</tr>
<tr>
<td>A gangorra quebrada (Broken seasaw)</td>
<td>0.784</td>
<td>0.126</td>
<td>-0.489</td>
<td>0.135</td>
</tr>
<tr>
<td>O canto do sabiá do campo (Mermaid song)</td>
<td>0.676</td>
<td>0.129</td>
<td>1.365</td>
<td>0.246</td>
</tr>
<tr>
<td>Meus óculos com dois graus (Two-degree glasses)</td>
<td>0.675</td>
<td>0.139</td>
<td>0.975</td>
<td>0.203</td>
</tr>
<tr>
<td>Qualquer bicicleta preta (Any Black bike)</td>
<td>0.984</td>
<td>0.177</td>
<td>-0.921</td>
<td>0.147</td>
</tr>
<tr>
<td>O perfume natural do jardim florido (Flowery garden’s natural scent)</td>
<td>1.013</td>
<td>0.203</td>
<td>-1.201</td>
<td>0.176</td>
</tr>
<tr>
<td>As portas iguais da cozinha (Same kitchen doors)</td>
<td>0.879</td>
<td>0.144</td>
<td>-0.876</td>
<td>0.151</td>
</tr>
<tr>
<td>O esquilo apressado (Rushed squirrel)</td>
<td>0.791</td>
<td>0.125</td>
<td>-0.090</td>
<td>0.123</td>
</tr>
<tr>
<td>A receita da empada (Patty recipe)</td>
<td>0.845</td>
<td>0.138</td>
<td>-0.173</td>
<td>0.120</td>
</tr>
<tr>
<td>Brigas entre certas cabras (Fights between certain goats)</td>
<td>0.991</td>
<td>0.158</td>
<td>0.106</td>
<td>0.109</td>
</tr>
<tr>
<td>Iscas de filé na tigela do inglês (Fillet baits in the Englishman bowl)</td>
<td>0.753</td>
<td>0.121</td>
<td>0.649</td>
<td>0.151</td>
</tr>
<tr>
<td>Raposas com sono e quietas (Sleepy quiet foxes)</td>
<td>0.760</td>
<td>0.122</td>
<td>0.530</td>
<td>0.142</td>
</tr>
</tbody>
</table>

Table 2. Percentage summary measures of the 17 noun phrases distributed according to levels of discrimination and difficulty

<table>
<thead>
<tr>
<th>Discrimination</th>
<th>N</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Very high</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>High</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Medium</td>
<td>15</td>
<td>88.23</td>
</tr>
<tr>
<td>Low</td>
<td>2</td>
<td>11.76</td>
</tr>
<tr>
<td>Very low</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Very difficult</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Difficult</td>
<td>1</td>
<td>5.88</td>
</tr>
<tr>
<td>Moderate</td>
<td>13</td>
<td>76.47</td>
</tr>
<tr>
<td>Easy</td>
<td>2</td>
<td>11.06</td>
</tr>
<tr>
<td>Very easy</td>
<td>1</td>
<td>5.88</td>
</tr>
<tr>
<td>Extremely easy</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 3. Analysis of the configurational and scalar invariance of noun phrases according to the school year

<table>
<thead>
<tr>
<th>Model</th>
<th>Parameter number</th>
<th>Chi-squared</th>
<th>Degree of freedom</th>
<th>p-value</th>
<th>RMSEA</th>
<th>CFI</th>
<th>TLI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Groups 1 and 2</td>
<td>Configurational</td>
<td>68</td>
<td>253.744</td>
<td>238</td>
<td>0.2307</td>
<td>0.022</td>
<td>0.983</td>
</tr>
<tr>
<td></td>
<td>Scalar</td>
<td>53</td>
<td>269.708</td>
<td>253</td>
<td>0.2247</td>
<td>0.022</td>
<td>0.982</td>
</tr>
<tr>
<td></td>
<td>Scalar vs. Configurational</td>
<td>16.801</td>
<td>15</td>
<td>0.3309</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Captions: RMSEA - root mean square error of approximation (>0.06); CFI - comparative fit index (>0.95); TLI Tucker-Lewis index (>0.90); Group 1: 2nd and 3rd grades; Group 2: 4th and 5th grades.
The syllables with VV and VVC structures were eliminated with the removal of the noun phrases analyzed by IRT. However, these syllabic structures composed more complex syllables and could, therefore, be assessed somehow (for example, in the words chuteira (football boot) and bois (oxes), respectively).

It should be emphasized that performance assessment tools should present, as important characteristic, strength or robustness to discriminate or differentiate responses of the individuals evaluated. The more robust the instrument, the more balanced the analysis of discrimination capacity and difficulty level of the items that compose such instrument. The results of the analysis corroborated this idea, in that it showed predominance of noun phrases with characteristics of medium discrimination power and moderate difficulty level (Tables 1 and 2).

The list of noun phrases was adequate for the evaluation of the writing of the total sample, which was confirmed by the configurational and scalar invariance analyses (Table 3), whose TLI and CFI showed good fit of the model to the sample ($\chi^2_{2455}=16.801 \ (p=0.3309)^{21}$). Although it was expected that the list would evaluate the same writing abilities of each investigated school year, it was necessary to group the initial (2nd and 3rd grades) and the final (4th and 5th grades) years. In addition to the statistical procedure (and sample size), grouping was probably determined by similarities of learning level between 2nd and 3rd, and 4th and 5th grades. In fact, it is expected that the alphabetical phase of writing learning will be completed by the end of 3rd grade, and will lead to learning and mastery of orthographic writing at the end of 5th grade.

To evaluate the amount of information underlying the construct that can be explained by their respective items as a whole divided into several levels, a Total Information Curve (TIC) was drawn, in which the highest levels of information indicate the most precise measure of the construct. This TIC reinforces the standard found in IRT. Peak of information on the TIC was around zero for school year 3rd grades. In addition to the statistical procedure (and sample size), grouping was probably determined by similarities of learning level between 2nd and 3rd, and 4th and 5th grades. In fact, it is expected that the alphabetical phase of writing learning will be completed by the end of 3rd grade, and will lead to learning and mastery of orthographic writing at the end of 5th grade.

To evaluate the amount of information underlying the construct that can be explained by their respective items as a whole divided into several levels, a Total Information Curve (TIC) was drawn, in which the highest levels of information indicate the most precise measure of the construct. This TIC reinforces the standard found in IRT. Peak of information on the TIC was around zero for school year 3rd grades. In addition to the statistical procedure (and sample size), grouping was probably determined by similarities of learning level between 2nd and 3rd, and 4th and 5th grades. In fact, it is expected that the alphabetical phase of writing learning will be completed by the end of 3rd grade, and will lead to learning and mastery of orthographic writing at the end of 5th grade.

To evaluate the amount of information underlying the construct that can be explained by their respective items as a whole divided into several levels, a Total Information Curve (TIC) was drawn, in which the highest levels of information indicate the most precise measure of the construct. This TIC reinforces the standard found in IRT. Peak of information on the TIC was around zero for school year 3rd grades. In addition to the statistical procedure (and sample size), grouping was probably determined by similarities of learning level between 2nd and 3rd, and 4th and 5th grades. In fact, it is expected that the alphabetical phase of writing learning will be completed by the end of 3rd grade, and will lead to learning and mastery of orthographic writing at the end of 5th grade.

To evaluate the amount of information underlying the construct that can be explained by their respective items as a whole divided into several levels, a Total Information Curve (TIC) was drawn, in which the highest levels of information indicate the most precise measure of the construct. This TIC reinforces the standard found in IRT. Peak of information on the TIC was around zero for school year 3rd grades. In addition to the statistical procedure (and sample size), grouping was probably determined by similarities of learning level between 2nd and 3rd, and 4th and 5th grades. In fact, it is expected that the alphabetical phase of writing learning will be completed by the end of 3rd grade, and will lead to learning and mastery of orthographic writing at the end of 5th grade.

To evaluate the amount of information underlying the construct that can be explained by their respective items as a whole divided into several levels, a Total Information Curve (TIC) was drawn, in which the highest levels of information indicate the most precise measure of the construct. This TIC reinforces the standard found in IRT. Peak of information on the TIC was around zero for school year 3rd grades. In addition to the statistical procedure (and sample size), grouping was probably determined by similarities of learning level between 2nd and 3rd, and 4th and 5th grades. In fact, it is expected that the alphabetical phase of writing learning will be completed by the end of 3rd grade, and will lead to learning and mastery of orthographic writing at the end of 5th grade.

To evaluate the amount of information underlying the construct that can be explained by their respective items as a whole divided into several levels, a Total Information Curve (TIC) was drawn, in which the highest levels of information indicate the most precise measure of the construct. This TIC reinforces the standard found in IRT. Peak of information on the TIC was around zero for school year 3rd grades. In addition to the statistical procedure (and sample size), grouping was probably determined by similarities of learning level between 2nd and 3rd, and 4th and 5th grades. In fact, it is expected that the alphabetical phase of writing learning will be completed by the end of 3rd grade, and will lead to learning and mastery of orthographic writing at the end of 5th grade.

To evaluate the amount of information underlying the construct that can be explained by their respective items as a whole divided into several levels, a Total Information Curve (TIC) was drawn, in which the highest levels of information indicate the most precise measure of the construct. This TIC reinforces the standard found in IRT. Peak of information on the TIC was around zero for school year 3rd grades. In addition to the statistical procedure (and sample size), grouping was probably determined by similarities of learning level between 2nd and 3rd, and 4th and 5th grades. In fact, it is expected that the alphabetical phase of writing learning will be completed by the end of 3rd grade, and will lead to learning and mastery of orthographic writing at the end of 5th grade.

To evaluate the amount of information underlying the construct that can be explained by their respective items as a whole divided into several levels, a Total Information Curve (TIC) was drawn, in which the highest levels of information indicate the most precise measure of the construct. This TIC reinforces the standard found in IRT. Peak of information on the TIC was around zero for school year 3rd grades. In addition to the statistical procedure (and sample size), grouping was probably determined by similarities of learning level between 2nd and 3rd, and 4th and 5th grades. In fact, it is expected that the alphabetical phase of writing learning will be completed by the end of 3rd grade, and will lead to learning and mastery of orthographic writing at the end of 5th grade.

To evaluate the amount of information underlying the construct that can be explained by their respective items as a whole divided into several levels, a Total Information Curve (TIC) was drawn, in which the highest levels of information indicate the most precise measure of the construct. This TIC reinforces the standard found in IRT. Peak of information on the TIC was around zero for school year 3rd grades. In addition to the statistical procedure (and sample size), grouping was probably determined by similarities of learning level between 2nd and 3rd, and 4th and 5th grades. In fact, it is expected that the alphabetical phase of writing learning will be completed by the end of 3rd grade, and will lead to learning and mastery of orthographic writing at the end of 5th grade.

To evaluate the amount of information underlying the construct that can be explained by their respective items as a whole divided into several levels, a Total Information Curve (TIC) was drawn, in which the highest levels of information indicate the most precise measure of the construct. This TIC reinforces the standard found in IRT. Peak of information on the TIC was around zero for school year 3rd grades. In addition to the statistical procedure (and sample size), grouping was probably determined by similarities of learning level between 2nd and 3rd, and 4th and 5th grades. In fact, it is expected that the alphabetical phase of writing learning will be completed by the end of 3rd grade, and will lead to learning and mastery of orthographic writing at the end of 5th grade.

To evaluate the amount of information underlying the construct that can be explained by their respective items as a whole divided into several levels, a Total Information Curve (TIC) was drawn, in which the highest levels of information indicate the most precise measure of the construct. This TIC reinforces the standard found in IRT. Peak of information on the TIC was around zero for school year 3rd grades. In addition to the statistical procedure (and sample size), grouping was probably determined by similarities of learning level between 2nd and 3rd, and 4th and 5th grades. In fact, it is expected that the alphabetical phase of writing learning will be completed by the end of 3rd grade, and will lead to learning and mastery of orthographic writing at the end of 5th grade.

Assessment of writing is a fundamental procedure to know the conditions and pace of learning. This study sought to prepare an instrument that could invariably evaluate the writing of Elementary School children and be easily applied, clinically or educationally, and enable reproducibility of the procedure in research.

Writing of noun phrases enables observation of whether schoolchildren perceive the boundaries between words\(^{[9-11]}\). It is true that other skills such as vocabulary development and morphosyntactic awareness play a role in this process\(^{[40]}\); however, it is considered that more basic issues, such as phonological and auditory processing, underlie undue junctions and segmentations, and should be identified\(^{[13,15]}\). Emergence of undue junctions or segmentations may indicate, therefore, the presence of deficits of primary abilities for the learning of reading and writing\(^{[27]}\).

Analysis of the items on the list of noun phrases determined the exclusion of some of them, which showed extreme values of discrimination and difficulty. Thus, most of the 17 remaining noun phrases, whose words maintained the graphical representation of all the phonemes of Brazilian Portuguese and most types of syllables, showed medium discrimination capacity, and only two presented low capacity. Regarding the level of difficulty, most noun phrases showed moderate level, and the others were distributed as follows: one with difficult, two with easy, and one with very easy level (Tables 1 and 2).

DISCUSSION

Assessment of writing is a fundamental procedure to know the conditions and pace of learning. This study sought to prepare an instrument that could invariably evaluate the writing of Elementary School children and be easily applied, clinically or educationally, and enable reproducibility of the procedure in research.

Writing of noun phrases enables observation of whether schoolchildren perceive the boundaries between words\(^{[9-11]}\). It is true that other skills such as vocabulary development and morphosyntactic awareness play a role in this process\(^{[40]}\); however, it is considered that more basic issues, such as phonological and auditory processing, underlie undue junctions and segmentations, and should be identified\(^{[13,15]}\). Emergence of undue junctions or segmentations may indicate, therefore, the presence of deficits of primary abilities for the learning of reading and writing\(^{[27]}\).

Analysis of the items on the list of noun phrases determined the exclusion of some of them, which showed extreme values of discrimination and difficulty. Thus, most of the 17 remaining noun phrases, whose words maintained the graphical representation of all the phonemes of Brazilian Portuguese and most types of syllables, showed medium discrimination capacity, and only two presented low capacity. Regarding the level of difficulty, most noun phrases showed moderate level, and the others were distributed as follows: one with difficult, two with easy, and one with very easy level (Tables 1 and 2).
ACKNOWLEDGEMENTS

The authors are grateful to Adriana Kida, professor at the Departamento de Fonoaudiologia - UNIFESP-EPM, for her collaboration in the preparation of the list of noun phrases, and to Hugo Cogo-Moreira, professor at the Departamento de Psiquiatria - UNIFESP-EPM, for the statistical analysis.

REFERENCES


Author contributions

MAI was responsible for the study design, collection, analysis and interpretation of the data, and writing of the manuscript. CH was in charge of the critical revision of the data analysis and the manuscript; MMSSH assisted with the preparation of the test items, design of the procedures, and revision of the manuscript. CRB4 was responsible for the study design, analysis and interpretation of the data, and final approval of the manuscript.