Analysis of fundamental frequency, jitter, shimmer and vocal intensity in children with phonological disorders

Summary

Phonological Disorder is a disturbance of primary manifestation of undefined causes that makes speech become unintelligible. The analysis of vocal parameters becomes important in the process of diagnosis of this disorder, since voice disorders could interfere in the production of speech sounds. **Aim:** The objective of this study was to verify vocal characteristics related to the intensity and fundamental frequency -F0- and their disturbance indexes - jitter and shimmer - in children with phonological disorders. **Study design:** clinical prospective with transversal cohort. **Material and Method:** There were 40 children, 20 of them with phonological disorders and 20 with no speech and language disturbances. Phonological exams with the ABFW infantile language test and spontaneous speech were applied. The Computer Speech Lab was used to record and perform acoustic analyses of the vowels /a/, /e/, /i/, through the vocal parameters: fundamental frequency, intensity, jitter and shimmer. **Results:** F0 - vowel /e/ was smaller, on average, in the Phonological Disorder Group and it was 126 Hz in the Control Group. To shimmer and jitter there was no evidence that the means of the Phonological Disorder Group were different from the ones of the Control Group (p= 0.191, p=0.865, respectively). As for intensity, there was evidence that the average did not differ in the Phonological Disorder Group and the Control Group (p= 0.002). **Conclusion:** The frequency of the vowel /e/ was smaller in the Phonological Disorder Group. There was difference between the two groups regarding the means of intensity of vowels /a/, /e/ and /i/, smaller in the Phonological Disorder Group. No differences between the groups were found regarding the averages of jitter and shimmer.

Key words: language development disorders, speech acoustics, diagnosis.
INTRODUCTION

Phonological disorders are affections to primary manifestation of undefined cause, which may turn the speech into intelligible production leading to misunderstood messages.

According to the survey performed among speakers of Brazilian Portuguese by Andrade et al., about 5.6% of the children who were seen in the speech and hearing sector of Centro de Saude Escola Samuel Barnsley Pessoa had phonological and/or phonetic affections. Gierut stated that the phonological disorders affect approximately 10% of the population of American children in pre-school and school age, and it is sufficiently severe in many cases to require clinical intervention in 80% of the cases.

The literature refers that to the proficiency of the phonological system by children involves the development of perception and production of a phonetic inventory, as well as phonological rules. Phonological rules correspond to the regularities that occur in the Phonology of a language.

Systematic simplifications of phonological rules that affect a class or sequence of sounds are named phonological processes. Phonological processes are expected within normal development, however, as the child grows and develops, they no longer use them and acquire the rules similar to adults.

In children with typical development, phonological processes are naturally suppressed, whereas in children with phonological disorders, clinical intervention is required. To that end, a well-structured assessment with selection and administration of the right tests is mandatory.

Making use of such procedures, researchers intend to understand the characteristics of the phonological disorder, considering the unknown cause so as to classify possible subtypes. Following this line, Shriberg and Kwiatkowski identified some etiological factors for the phonological disorder, which allows the division into four subtypes: unknown origin, otitis media, speech development apraxia (SDA), and psychosocial involvement.

In 1999, Shriberg proposed a new classification that suggested the existence of five subgroups with different etiologies: delay in speech of genetic origin, delay in speech with repetitive otitis media with effusion, speech delay with SDA, speech delay with implication of psychosocial development and, finally, type of residual mistakes with history of distortions.

Factors related with the assessment for functional diagnosis

As phonological disorders are very common in the pre-school and school populations and they may be caused by different causes, investigations try to find specific descriptions for segmental and supra-segmental linguistic symptoms to relate them with specific characteristics of each cause correlated with the disorder. According to Lowe, segmental traits are referred to vowels and consonants that get together to form syllables, words and sentences, whereas supra-segmental or prosodic traits are elements that form the rhythm of a specific language or production.

To that end, in the speech assessment of phonological disorders, in addition to phonological analysis, we should include an analysis of cognitive, linguistic, neuromotor skills as well as the oral structure and functions, hearing, fluency, voice and supra-segmental aspects of speech. We should highlight that the process of functional diagnosis explores the search for etiological factors and tries to separate the subtypes of phonological disorders, which is important for more precise clinical intervention.

Aspects related with vocal quality

The segmental and supra-segmental aspects of vocal production are observed in the speech. The speech, considering that it is exclusive to a subject, becomes adjustable to the particularities of the speaker. Thus, each person may use variations of speed, height and intensity in their production. According to Peña-Brooks and Hedge, the decrease and increase in height are essential to give melody to the sentence. Stress pattern, related with the combination of increased intensity, greater duration and higher frequency in syllables of a sentence emphasize some parts of the production, ensuring rhythm to the spoken language.

Abnormal prosodic characteristics have been described in the literature as a trait of speech development apraxia. As mentioned by Shriberg, this condition is a subtype of phonological disorders. Thus, the assessment of prosody may contribute to the identification of this condition.

Shriberg et al. presented a perceptual procedure (PVSP - Prosody-Voice Screening Profile) to assess the prosody and voice of spontaneous speech. These authors consider essential to assess vocal characteristics of children with phonological disorders, including aspects such as pitch, loudness and vocal quality.

The study of the speech prosodic aspects may be conducted by means of assessment of fundamental frequency (f0) variation, duration of syllables, words and other units and intensity.

According to Behlau and Pontes, vocal intensity is directly related with subglottic pressure of the air column. Subglottic pressure, in turn, depends on factors such as amplitude of vibration and tension of vocal folds, more specifically the glottic resistance.

Variations of intensity, however, also depend on frequency. To Behlau and Pontes, high voices tend to be more intense, because the increase in laryngeal tonus generates higher glottic resistance and, consequently, more intensity.

The analysis of vocal parameters become important in the process of diagnosis of phonological disorders, given...
that vocal disorders and articulation normally coexist, given that the former may lead to difficulties of sound in the latter.

The voice may be assessed subjectively (perceptual-auditory assessment) and/or objectively to help with acoustic analysis equipment. According to Gurgueira, the acoustic analysis enables determination of number, frequency and amplitude (intensity) of vibrations to form a complex sound.

The most important vocal acoustic parameters for clinical use are measurements of noise, vocal extension profile, acoustic spectrography, fundamental frequency and perturbation index - jitter and shimmer.

According to Behlau et al., fundamental frequency is determined physiologically by the number of cycles that the vocal folds make in a second, and they are the natural result of the length of these structures.

Jitter and shimmer represent the variations that occur in the fundamental frequency. Whereas jitter indicates the variability or perturbation of fundamental frequency, shimmer refers to the same perturbation, but it is related to amplitude of sound wave, or intensity of vocal emission. Jitter is affected mainly because of lack of control of vocal fold vibration and shimmer with reduction of glottic resistance and mass lesions in the vocal folds, which are related with presence of noise at emission and breathiness.

Some studies were developed to compare perceptually and acoustically the voice of children with and without communication affections, such as the one performed by Shriberg and Kwiatkowski, who used the perceptual procedure PVSP - Prosody-Voice Screening Profile to compare a group of 64 children aged 3 to 6 years, with phonological disorders, whose speech mistakes were severe enough to interfere in the intelligibility and a group of 71 children aged 3 to 5 years with normal speech development, studied by Miller. Both groups presented similar results in speech rate, pitch and resonance. However, 17.8% of the children with phonological disorders versus 1.4% of the normal children were classified as having perceptual and discussable involvement of stress pattern; 30.7% versus 2.8% in loudness, and 48.8% vs. 23.8% in laryngeal vocal quality aspects.

There are few studies about supra-segmental acquisition and relations between vocal parameters and phonological disorders in children. Studies in this area are extremely important to enable advances in search for etiology and ensure more effective intervention and early identification of the disorder.

Thus, the present study intended to study the vocal characteristics related to intensity and fundamental frequency and their perturbations indexes - jitter and shimmer, in children with phonological disorders.

MATERIAL AND METHODS

The present study was approved by the Ethics and Research Committee HC-FMUSP (n° 00/09220-3) and sponsored by FAPESP (process nº 02/03102-4). The people responsible for the children signed the Informed Consent Term.

The study comprised 40 children aged 4 to 10.2 years, both genders, residents in Sao Paulo.

The group with phonological disorder (GTF) comprised 29 subjects, 9 female and 11 male subjects, seen in the Laboratory of Phonology (LIF), Department of Physical Therapy, Speech and Hearing Therapy, Occupation Therapy, Medical School, University of Sao Paulo. They were selected after the diagnostic process that included language tests ABFW28, spontaneous speech test, oral myofunctional system analysis, phonological awareness and audiological assessment.

The control group (GC) comprised 20 subjects without speech or language affections, from schools in the region of Butanta, Sao Paulo, 9 female and 11 male subjects. To select then, we used a questionnaire that the parents answered to check whether they had any complaint concerning speech or language. Next, we applied the phonology test of language ABFW29 and the spontaneous speech test. Thus, we included in the GC only subjects who presented phonological development within the expected level for the age without complaints related to language development.

All tests were recorded in digital audio tape (DAT Foster D-5 Digital Master Record) and recorded in digital video (Sony CCD-TRV66).

To acoustically analyze the studied parameters - fundamental frequency (F0), jitter, shimmer and intensity - we used the Computer Speech Lab (CSL), manufactured by Kay Elemetrics - model 4300B, and unidirectional microphone brand Shure, model SM-58.

PROCEDURES

The analysis of F0 (Hz), jitter (%), shimmer (dB) and intensity (dB HL) was performed with isolated production and sustained vowels /a/, /e/, and /i/.

Subjects remained seated and were instructed to produce vowels at comfortable intensity and height, with the microphone at 10cm from their mouth. Each vowel was produced 3 times, and they were directly recorded by the Computerized Speech Lab (CSL).

To neutralize the effects of vocal attach, the beginning of the recording of each vowel was discarded. Samples were edited to present the same duration in all subjects, or 2 seconds.

Each one of the three sustained vowel productions were individually analyzed for each one of the studied parameters. The final value of the measures of each of the vowels was the mean obtained from the analysis of each one of the vowels separately.

After data collection, we adjusted in the computer the command of the sampling rate for 10,000 Hz.
To reach values of F0, jitter and shimmer, vowel samples were made through commands of CSL. The calculation of intensity was also made in CSL, which provided intensity in three moments - beginning of emission (0 second), half emission (1 second) and end of emission (2 seconds). The final value of the obtained intensity was the mean of initial, mid and final values of each vowel.

RESULTS

The descriptive analysis demonstrated that F0 of vowels /a/ and /i/ were on average similar in both groups, whereas F0 of vowel /e/ was smaller on average for GTF (237 Hz for GC and 126 Hz for GTF). Jitter in all vowels presented similar distribution for both groups. As to shimmer, only in vowel /e/ it was higher than the median in GTF (0.339 dB for GC and 0.486 dB for GTF).

As to intensity, we observed that vowels /a/, /e/ and /i/ presented smaller values than the median for GTF. Results may be better visualized in Tables 1 and 2.

To perform the inference analysis of F0 we used the non-parametric Mann-Whitney test. Table 3 shows a level of significant of 5%, concluding that only fundamental frequency of vowel /e/ was different between GTF and GC. Figure 1 is a box plot related to the comparison between GC and GTF and variable F0.

Table 1. Descriptive analysis - GC

<table>
<thead>
<tr>
<th>Variables</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>Standard deviation</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Q1</th>
<th>Q3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fo /a/ (Hz)</td>
<td>20</td>
<td>243</td>
<td>240</td>
<td>24</td>
<td>206</td>
<td>281</td>
<td>225</td>
<td>264</td>
</tr>
<tr>
<td>Fo /e/ (Hz)</td>
<td>20</td>
<td>237</td>
<td>241</td>
<td>31</td>
<td>148</td>
<td>285</td>
<td>218</td>
<td>262</td>
</tr>
<tr>
<td>Fo /i/ (Hz)</td>
<td>20</td>
<td>229</td>
<td>245</td>
<td>46</td>
<td>153</td>
<td>281</td>
<td>171</td>
<td>265</td>
</tr>
<tr>
<td>Jitter /a/ (%)</td>
<td>20</td>
<td>1,551</td>
<td>1,338</td>
<td>0,906</td>
<td>0,909</td>
<td>5,158</td>
<td>1,043</td>
<td>1,653</td>
</tr>
<tr>
<td>Jitter /e/ (%)</td>
<td>20</td>
<td>1,678</td>
<td>1,331</td>
<td>1,303</td>
<td>0,814</td>
<td>6,326</td>
<td>1,078</td>
<td>1,547</td>
</tr>
<tr>
<td>Jitter /i/ (%)</td>
<td>20</td>
<td>1,113</td>
<td>1,008</td>
<td>0,368</td>
<td>0,642</td>
<td>2,239</td>
<td>0,868</td>
<td>1,313</td>
</tr>
<tr>
<td>Shimmer /a/ (dB)</td>
<td>20</td>
<td>0,610</td>
<td>0,541</td>
<td>0,187</td>
<td>0,393</td>
<td>1,013</td>
<td>0,445</td>
<td>0,743</td>
</tr>
<tr>
<td>Shimmer /e/ (dB)</td>
<td>20</td>
<td>0,375</td>
<td>0,339</td>
<td>0,170</td>
<td>0,185</td>
<td>0,918</td>
<td>0,262</td>
<td>0,461</td>
</tr>
<tr>
<td>Shimmer /i/ (dB)</td>
<td>20</td>
<td>0,465</td>
<td>0,426</td>
<td>0,243</td>
<td>0,129</td>
<td>1,298</td>
<td>0,363</td>
<td>0,524</td>
</tr>
<tr>
<td>Intensity - /a/ (dB)</td>
<td>20</td>
<td>73,8</td>
<td>73,8</td>
<td>2,2</td>
<td>69,7</td>
<td>78,9</td>
<td>72,0</td>
<td>75,3</td>
</tr>
<tr>
<td>Intensity - /e/ (dB)</td>
<td>20</td>
<td>75,8</td>
<td>75,7</td>
<td>2,2</td>
<td>70,8</td>
<td>79,4</td>
<td>74,4</td>
<td>77,5</td>
</tr>
<tr>
<td>Intensity - /i/ (dB)</td>
<td>20</td>
<td>75,5</td>
<td>75,3</td>
<td>2,4</td>
<td>71,5</td>
<td>81,5</td>
<td>74,1</td>
<td>77,1</td>
</tr>
</tbody>
</table>

Key:
GC - control group
F0 - fundamental frequency
N - total number of subjects
Q1 - first quartile
Q3 - third quartile

Table 2. Descriptive analysis - GTF

<table>
<thead>
<tr>
<th>Variables</th>
<th>N</th>
<th>Mean</th>
<th>Median</th>
<th>Standard deviation</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Q1</th>
<th>Q3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fo /a/ (Hz)</td>
<td>20</td>
<td>246</td>
<td>246</td>
<td>23</td>
<td>193</td>
<td>284</td>
<td>232</td>
<td>262</td>
</tr>
<tr>
<td>Fo /e/ (Hz)</td>
<td>20</td>
<td>126</td>
<td>126</td>
<td>12</td>
<td>100</td>
<td>146</td>
<td>118</td>
<td>136</td>
</tr>
<tr>
<td>Fo /i/ (Hz)</td>
<td>20</td>
<td>228</td>
<td>237</td>
<td>42</td>
<td>171</td>
<td>285</td>
<td>183</td>
<td>265</td>
</tr>
<tr>
<td>Jitter /a/ (%)</td>
<td>20</td>
<td>1,609</td>
<td>1,247</td>
<td>0,899</td>
<td>0,810</td>
<td>3,825</td>
<td>0,986</td>
<td>2,063</td>
</tr>
<tr>
<td>Jitter /e/ (%)</td>
<td>20</td>
<td>1,965</td>
<td>1,395</td>
<td>1,914</td>
<td>0,673</td>
<td>8,732</td>
<td>0,985</td>
<td>1,722</td>
</tr>
<tr>
<td>Jitter /i/ (%)</td>
<td>20</td>
<td>1,331</td>
<td>1,107</td>
<td>0,816</td>
<td>0,706</td>
<td>4,243</td>
<td>0,834</td>
<td>1,370</td>
</tr>
<tr>
<td>Shimmer /a/ (dB)</td>
<td>20</td>
<td>0,610</td>
<td>0,576</td>
<td>0,251</td>
<td>0,282</td>
<td>1,369</td>
<td>0,420</td>
<td>0,735</td>
</tr>
<tr>
<td>Shimmer /e/ (dB)</td>
<td>20</td>
<td>0,501</td>
<td>0,486</td>
<td>0,235</td>
<td>0,195</td>
<td>1,123</td>
<td>0,303</td>
<td>0,653</td>
</tr>
<tr>
<td>Shimmer /i/ (dB)</td>
<td>20</td>
<td>0,526</td>
<td>0,486</td>
<td>0,216</td>
<td>0,150</td>
<td>0,864</td>
<td>0,337</td>
<td>0,709</td>
</tr>
<tr>
<td>Intensity - /a/ (dB)</td>
<td>20</td>
<td>72,2</td>
<td>72,2</td>
<td>2,02</td>
<td>69,1</td>
<td>76,2</td>
<td>70,5</td>
<td>73,8</td>
</tr>
<tr>
<td>Intensity - /e/ (dB)</td>
<td>20</td>
<td>73,5</td>
<td>74,5</td>
<td>2,9</td>
<td>68,1</td>
<td>80,2</td>
<td>71,0</td>
<td>75,1</td>
</tr>
<tr>
<td>Intensity - /i/ (dB)</td>
<td>20</td>
<td>72,7</td>
<td>72,6</td>
<td>3,3</td>
<td>66,1</td>
<td>78,0</td>
<td>70,3</td>
<td>75,1</td>
</tr>
</tbody>
</table>

Key:
GC - control group
F0 - fundamental frequency
N - total number of subjects
Q1 - first quartile
Q3 - third quartile
To compare the means of shimmer and jitter we used the Analysis of Variance with two fixed factors and repetitive measures in a factor after the transformation of the variable response according to the box plot. To shimmer and the level of significance of 5%, we concluded that there was no evidence that the means of GTF were different from GC (p=0.191). For jitter, the test did not show differences between GC and GTF (p=0.865). Figures 2 and 3 refer to Box plots that compared GC and GTF concerning the variables of shimmer and jitter.

Finally, for intensity, we used the Analysis of Variance with two fixed factors and repetitive measures of one factor. At a level of significance of 5%, we detected evidence that the mean of intensity would differentiate GTF from GC (p=0.002). Figure 4 shows box plots comparing GC and GTF in relation to intensity.

**DISCUSSION**

The results found in this study showed that both children in the GC and in the GTF presented values of F0 similar to other studies. Thus, mean values of F0 for vowel /a/ were between 243 and 246 Hz and for vowel /i/ between 229 and 228 Hz, values close to those found by Hasek et al., in which mean fundamental frequency values found for male subjects aged 5 years were 247.5 Hz, 6 years 262.5 Hz, 7 years 234.2 Hz, 8 years 235.6 Hz, 9 years 230.4 Hz and 10 years 228.9 Hz. For female subjects, the following values were found for 5 years 257.7 Hz, 6 years 254.3 Hz, 7 years 261.7 Hz, 8 years 264 Hz, 9 years 246.7 Hz and 10 years 253.7 Hz.

The results found by Navas indicated mean values between 298.1 and 290.9 Hz for male and 299.8 and 290.9 Hz for female. Behlau observed that the value of F0 for children aged 8 to 11 years was 236 Hz and Awan and Mueller detected in spontaneous speech 243 Hz for female subjects and 240 Hz for male subjects.

Only vowel /e/ differentiated the two groups (in GC the mean value found was 237 Hz, whereas in the GTF this value was 126 Hz), a fact that may be related with the configuration of the vocal tract required for sound production. The differences in vocal tract may be resulting from the adapted movements children who have phonological disorders make to execute their productions. It is interesting to emphasize that this finding should be properly investigated to check the possibility of being used as a differential element in children with this disorder.

Even though there were no significant differences between GC and GTF for jitter and shimmer, the values found were below those reported in another study with Brazilian children, by Behlau, in which jitter was 2.3% and shimmer was 2.5 dB.

This fact could have been influenced by the use of different equipment for the acoustic analysis.

As to intensity, we observed statistically significant differences between the two groups, in which GTF presented mean intensities lower than those in the GC. This finding may be related with psychosocial aspects that interfered in the communication of children with phonological disorder.

It is important to emphasize that it was difficult to compare our acoustic findings with the findings of other studies, given that these studies focused on adults and children with affections and different communication disorders and not phonological disorders. Thus, the variation...
of methodologies and equipment for the acoustic analysis using these studies limits the possible comparison between studies.

However, the results found in the present study point to the fact that these children with phonological disorders compared to children without disorders do not present any abnormality that affects the vocal folds, either muscle or neural activity involved with phonation, either lesions that may cause increase in aperiodicity of vocal fold vibration, which reflect the increased values of jitter. As to shimmer, the study indicated that the characteristics such as reduction of glottic resistance, vocal fold mass lesions and greater noise at production, factors that could lead to affections of shimmer values, were not necessarily found.

**CONCLUSION**

The study showed that, in the comparison between GC and GTF, F0 of vowel /e/ is lower than for other vowels in GTF, whereas in GC all vowels had similar F0 values. Another detected difference between the groups was the mean intensity of vowel /a/, /e/ and /i/, which were lower in GTF.

We did not find statistically significant differences between the groups concerning means of jitter and shimmer of vowels. For the vowel effect, we observed that the mean of jitter of vowel /i/ was lower than for vowels /a/ and /e/ and the mean of shimmer of the vowel /a/ was higher than for vowels /e/ and /i/.

Similarly, F0 of vowel /e/ and intensity were acoustic measurements that led to differentiation between the groups, which should be further investigated. As pointed out by the studies by Shriberg 8,10,14, these characteristics may be used to complement the diagnosis of phonological disorders and have to be further studied.

**ACKNOWLEDGEMENT:**

To Fapesp for having sponsored the present study.

**REFERENCES**

22. Miller JF. SALT: Reference data base project. Language Analysis Laboratory Waisman Center on Mental Retardation and Human Development University of Wisconsin-Madison; 1990.
27. Miller JF. SALT: Reference data base project. Language Analysis Laboratory Waisman Center on Mental Retardation and Human Development University of Wisconsin-Madison; 1990.