Generalization of temporal order detection skill learning: two experimental studies of children with dyslexia

C.F.B. Murphy and E. Schochat
Departamento de Fonoaudiologia, Fisioterapia e Terapia Ocupacional, Faculdade de Medicina, Universidade de São Paulo, São Paulo, SP, Brasil

Abstract
The objective of this study was to investigate the phenomenon of learning generalization of a specific skill of auditory temporal processing (temporal order detection) in children with dyslexia. The frequency order discrimination task was applied to children with dyslexia and its effect after training was analyzed in the same trained task and in a different task (duration order discrimination) involving the temporal order discrimination too. During study 1, one group of subjects with dyslexia (N = 12; mean age = 10.9 ± 1.4 years) was trained and compared to a group of untrained dyslexic children (N = 28; mean age = 10.4 ± 2.1 years). In study 2, the performance of a trained dyslexic group (N = 18; mean age = 10.1 ± 2.1 years) was compared at three different times: 2 months before training, at the beginning of training, and at the end of training. Training was carried out for 2 months using a computer program responsible for training frequency ordering skill. In study 1, the trained group showed significant improvement after training only for frequency ordering task compared to the untrained group (P < 0.001). In study 2, the children showed improvement in the last interval in both frequency ordering (P < 0.001) and duration ordering (P = 0.01) tasks. These results showed differences regarding the presence of learning generalization of temporal order detection, since there was generalization of learning in only one of the studies. The presence of methodological differences between the studies, as well as the relationship between trained task and evaluated tasks, are discussed.

Key words: Temporal order detection; Learning; Dyslexia; Children

Introduction
It has been shown that a deficit in temporal processing of auditory stimuli can be the basis for the reading problems of children with dyslexia (1-4). More specifically, researchers have demonstrated a significantly worse performance in this group of children in a specific skill of auditory temporal processing, i.e., temporal order detection (4-7). Temporal order detection skill refers to the processing of multiple auditory stimuli in their exact order of occurrence. Thanks to this skill, individuals are able to identify the correct sequence of sounds (8). The relationship between this task and written language skills is associated with the identification of the sequence of phonemes in words such as "mats" vs "mast" (9).

A number of studies have investigated the effect of training auditory temporal processing on reading skills on the basis of this hypothesis. Results have shown that this is a controversial topic; some studies have found improvement of verbal skills as a consequence of auditory training (10-13), whereas others have found improvement of the auditory ability but not of verbal skills (14-17). One of the hypotheses considered in this last case is related to the type of learning taking place during auditory training. According to Robinson and Summerfield (18), the improvement of performance resulting from training seems to be associated with the improvement regarding the procedure carried out ("procedure or task learning") and it might be related to perceptual discrimination ("perceptual learning"). Therefore, it has been suggested...
that the post-training improvement found in the auditory tests might demonstrate the learning of the procedure used (for instance: specific task to discriminate the temporal order of stimuli having different frequencies) instead of demonstrating the learning of the processes involving perceptual skills as a whole, regardless of the characteristics of the stimulus used. Thus, learning restricted to the task could explain the absence of improvement of verbal skills.

Lakshminarayanan and Tallal (19) claim that it is not easy to investigate procedural learning and perceptual learning as separate processes because, in most studies, the tasks used during training are the same as those used in tests applied before and after training. An alternative manner to investigate this issue is through the phenomenon of learning generalization, which shows the perceptual learning of a trained skill in tasks different from those used during training.

Two studies have been conducted to investigate the occurrence of learning generalization of temporal order detection skill. According to Shinn (20), Frequency and Duration Pattern tests are the most widely used tests to investigate order detection, each one of them involving different tasks. According to Mossbridge et al. (9), in a frequency discrimination task, the listener distinguishes the order of sounds having different frequencies, whereas, in a duration discrimination task, the listener distinguishes the order of sounds having different durations.

In study 1, the performance of the experimental group ("trained group with dyslexia") in the above mentioned tests was compared to that of a control group ("untrained group with dyslexia") before and after training. In study 2, a group with dyslexia was compared at three different time intervals: 2 months before auditory training and at the beginning and at the end of training. During the 2 months preceding auditory training, the participants only attended speech therapy sessions, which were considered to be a "control" intervention. During the last 2 months, the participants underwent auditory training in combination with speech therapy. Hence, it was possible to compare the effect of both types of intervention on the same group. In addition, the effect of differences that might be present between the profiles of each group in study 1 could be controlled, such as intellectual quotient, gender, motivation to participate, or presence of another intrinsic characteristic of each child.

If there is learning generalization, which is demonstrated by improvement in both tests, then we may assume that there is perceptual learning of the trained skill. If there is no learning generalization, which might be evidenced by the improvement only in the Frequency Pattern test, we may assume that there was only procedural learning. The implications of the possible results will be discussed regarding reading changes.

**Material and Methods**

The study was approved by the Ethics Committee for Analysis of Research Projects (CAPPesq) of Hospital das Clínicas, Medical School, Universidade de São Paulo (FMUSP) on August 9, 2006, Research Protocol #551/06.

**Participants and design of both studies**

Data were collected at the Laboratory of Speech and Language Investigation in Auditory Processing of the Speech and Language Therapy Course, Medical School, University of São Paulo, from January through December 2007.

The inclusion criteria were: to have a diagnosis of dyslexia established according to the Brazilian Association of Dyslexia (ABD), including the specific classification of moderate or severe level. This diagnosis was based on the following criteria: average or above average intelligence according to the Wechsler Intelligence Scale for Children-III (90 or above in verbal and non-verbal intelligence tests), 2-year delay in reading skills and phonological awareness compared with chronological age, anamnesis with parents to investigate any alternative problem that could interfere with reading, such as education or teaching method; to be a native speaker of Brazilian Portuguese; to be between 7 and 14 years old; to attend private schools in the same municipality to make sure that the participants had similar socioeconomic and cultural levels; to achieve normal results on the Basic Audiologic Assessment (audiometry, speech audiometry, and impedance audiometry); no evidence of cognitive, psychological, neurologic, and ophthalmologic abnormalities or delayed oral language acquisition. The subjects should also have no history of otitis media or any musical knowledge.

All participants were undergoing speech language therapy for the treatment of dyslexia when the study was started. According to the ABD, the rehabilitation method most widely used by professionals responsible for the treatment of dyslexia is the Phonic Method. Phonic literacy consists of a combination of exercises for the development of phonological awareness and teaching of the correspondence between graphemes and phonemes (21).

In study 1, the participants of each group were distributed so that the means of the groups regarding the variables age, reading skills, and phonological awareness were similar (Table 1).

The trained group consisted of 12 children with dyslexia (9 boys and 3 girls, mean age 10.9 ± 1.4 years) and the untrained group comprised 28 children with dyslexia (19 boys and 9 girls, mean age 10.4 ± 2.1 years). Based on the Student t-test, statistics indicated that the age distribution was similar for both groups (P = 0.329). In addition, the analysis showed that there were no differences between the groups regarding reading skills, phonological awareness, or auditory
processing (Table 1). Furthermore, considering the fact that both groups consisted of individuals of varied ages, i.e., between 7 and 14 years, the existence of a correlation between the performances on each test before training and the participants’ age was investigated using the Spearman correlation coefficient (22).

There were positive significant correlations between age and the Word Reading and Text Reading tests, and a negative significant correlation between age and the Frequency test. Thus, analysis of group performance on each test was carried out considering the division into the following age groups: 7-10 and 11-14 years.

In study 2, the untrained group of study 1 was submitted to auditory training after participating in the first study and a third evaluation was performed following the first study, containing the same tests. As a result of dropouts between the two studies, of the 28 children in the untrained group of study 1, 18 remained in study 2 (12 boys and 6 girls, mean age 10.1 ± 2.1 years).

The tests applied before and after training are described below.

**Frequency Pattern test**

The Frequency Pattern test was designed by Ptacek and Pinheiro (23). In this test, the participants listen to 20 trials including three stimuli each. The participants are advised to listen carefully to the three stimuli and to discriminate the order of appearance, providing oral answers. The stimuli consist of 880 Hz, which should be classified as "low (L)" or 1122 Hz, classified as "high (H)". Therefore, there were 8 possible oral answers: "HHH", "LLL", "HHL", "HLL", "LLH", "LHH", "LHL", and "HLH". The stimuli had a constant duration of 150 ms and the time interval between the three stimuli (interstimulus interval, ISI) was 200 ms and the time interval between trials was 10 s. Each stimulus presents a rise/fall of 10 ms. The test was applied in an acoustic booth at a 40-dB SL (decibel sensation level) intensity.

**Duration Pattern test**

The Duration Pattern test was proposed by Musiek et al. (24). In this test, the participants listen to 20 trials containing three stimuli each. The participants are advised to listen carefully to the three stimuli and to discriminate the order of appearance, providing an oral answer. The stimuli had a constant frequency of 1000 Hz and a duration of 250 ms for stimuli considered short (S) or 500 ms for stimuli considered long (L). Therefore, there were eight possible answers: "SSL", "SLL", "LSS", "SSS", "LLL", "LSS", "LSL", and "SLS". The time interval between the three stimuli was 300 ms and the time interval between trials was 10 s. Each stimulus presents a rise/fall of 10 ms. The test was applied in an acoustic booth at a 40-dB SL intensity.

At the end of both tests, the percentage of correct answers was calculated, and this value was later compared to the value obtained before training.

**Characteristics of training**

Auditory training was performed using a software program specifically designed for the training of temporal order detection.

Training consisted of phases with different levels of difficulty. In all phases, the auditory stimuli comprise multiple frequencies and are classified as rising (consisting of frequencies ranging from low to high) and falling (consisting of frequencies ranging from high to low). These stimuli were developed similarly to the Portuguese phonemes regarding duration, frequency, and the variation of octaves per second present in the phonemes (25). According to Murphy et al. (25), in Brazilian Portuguese the duration of the consonants varied from 31 ms to 170.5 ms and the values of the second format frequency varied from 2697.0 to 4916.0 Hz. According to F2 initial and final average values, the mean value of octaves per second was 6.8 considering all consonants. The stimuli of the computer program have a final or initial frequency of 500, 1000, or 2000 H, a duration of 40 to 200 ms, and an ISI of 20 and 500 ms. Stimulus duration and ISIs vary automatically according to the level of difficulty of the game (Figure 1).

In the first phase, the participants listen to two auditory stimuli and see the symbols “equal to” and “different from” on the screen; if both stimuli are the same, i.e., if both are rising or falling, the participants must click on the equal to symbol; if the stimuli are different, i.e., one is rising and the other is falling or vice-versa, the participants must click on the different from symbol (Figure 2A). In phase 2, the participants listen to two auditory stimuli and see arrows on the screen; these arrows represent each stimulus: ↑ for rising or ↓ for falling. The participants must associate each stimulus by clicking on the corresponding symbol (Figure 2B). In phase 3, the participants listen to 3 auditory stimuli that must also be associated with the symbols on the screen, similarly to what happens in phase 2.

The game started with the stimuli lasting 200 ms and separated by 500-ms ISIs. Automatic reduction of values occurred after 70% of right answers in each step, which was composed of 12 trials. These values were reduced in six steps for stimulus duration (200/150/100/80/60/40 ms) and in 13 steps for ISIs (500/450/400/350/300/250/200/150/100/80/60/40/20 ms). If less than 70% of answers were right, the phrase “GAME OVER” appeared on the screen and the
A CD containing the games and all the guidelines and rules was provided to the subjects before performing auditory training. The rules were read and explained to each parent. In both studies, the participants trained in their own house. After installing the software, they were instructed to play for 2 months, five times a week, 20 min a day. In addition, the participants were advised to send their scores to the researcher by e-mail using an Internet link provided with the software. Thus, it was possible to confirm if the participants were following the training rules, as previously arranged, and to check the evolution of their performance in each game.

**Results**

Analysis of variance for repeated measures (26), non-parametric analysis of variance of repeated ordinal data (27), and Bonferroni's correction were used to compare groups and to evaluate the effect of treatment. The level of significance was set at 0.05. This value indicates that the study had sufficient statistical power to detect significant differences and a high level of confidence. When the level of significance was between 0.05 and 0.1 the lack of statistical power of the test was attributed to the small sample size (27).

**Study 1**

Table 2 shows the mean number of days played by the trained group, which corresponded to approximately 30 days, and the last stage achieved (stage 6), in which stimuli lasted 150 ms and were separated by 300-ms ISIs. In addition, the mean number of correct answers for both groups in each test is also shown. Data were analyzed statistically by analysis of variance with repeated ordinal data.

**Frequency pattern.** There was an effect of age (P = 0.005) and group (P < 0.001). There was no difference between mean scores in either period for the untrained group in the 7- to 10-year age range and there was a small effect size according to Cohen's $d$ inspection ($d = 0.3$). However, the post-training mean score for the trained group was higher than the pre-training score ($P < 0.001$) with a large effect size ($d = 2.51$). Conclusions were similar for the 11- to 14-year age group, with a higher mean post-training score than pre-training score for the trained group ($P = 0.050$), with a moderate effect size ($d = 0.5$), but there was no difference between means in either interval for the untrained group, with a small effect size ($d = 0.07$).

**Duration pattern.** There was a marginal effect of age on mean scores ($P = 0.060$) and there was no difference between means in either period, regardless of group. Inspection of Cohen's $d$ indicated a minimum within-group effect size for the pre- and post-treatment periods within the trained group for 7- to 10-year-old ($d = 0$) and 11- to 14-year-old children ($d = 0.157$).

**Study 2**

Table 3 shows the mean number of days played by the group, which corresponded to approximately 31 days and the last stage achieved (stage 5), in which stimuli lasted 150 ms and were separated by 400-ms ISIs. In addition, the mean number of correct answers in each test for both periods is also shown. The data for both tests were analyzed statistically by non-parametric analysis.

**Frequency pattern.** There was no effect of age on mean scores. There was no significant difference between means in pre-training 2 and 1, but the mean post-training score was higher than the pre-training 2 score ($P < 0.001$). Inspection of Cohen's $d$ indicated a minimum within-group effect size for pre-training 1 and 2 for 7- to 10-year-old children ($d = 0.174$) and a moderate effect size for 11- to 14-year-old children ($d = 0.651$). For the pre-training 2 and post-training periods, there was a small within-group effect size for 7- to 10-year-old children ($d = 0.329$) and a moderate effect size for 11- to 14-year-old children ($d = 0.428$).

**Duration pattern.** There was no effect of age group on mean scores. There was no significant difference between means in pre-training 2 and 1, and the post-training mean was higher than the pre-training 2 mean ($P = 0.010$). Inspection of Cohen's $d$ indicated a small within-group effect size for pre-training 1 and 2 for 7- to 10-year-old children ($d = 0.299$) and a minimum effect size for 11- to 14-year-old children ($d = 0$). For the pre-training 2 and post-training periods, there was a moderate within-group effect size for 7- to 10-year-old children ($d = 0.654$) and a large effect size for 11- to 14-year-old children ($d = 1.717$).

**Discussion**

The results of both studies showed significant improvement after training for the Frequency Pattern test, confirming the efficacy of training for learning the trained task. The studies yielded different results for the Duration Pattern test, i.e., a significant improvement only in the post-training period in study 2, and the absence of significant differences between
pre- and post-training results for both groups in study 1. While the results of study 2 suggest learning of the trained skill based on learning generalization, the results of study 1 suggest only task learning. In addition, based on the results of study 2, it is also possible to state that the post-training improvement was actually a consequence of auditory training since the training considered as “control” did not influence the results. Tables 1 and 2 also show that the mean number of days played and the last stage reached during the game were similar in both studies (study 1: 30 days and stage 6; study 2: 31 days and stage 5). These data confirm that the trained groups had the same level of difficulty and the same training intensity in both studies.

Based on these results, learning generalization is controversial, since the studies showed different results in terms of learning generalization. These differences might be explained by the methodological differences between studies. Considering the number of participants in each study, in study 1, the trained group consisted of a smaller number of participants compared to study 2 (12 in study 1 and 18 in study 2). This factor might have had a negative influence on the observation of a significant change in the group for the Duration Pattern Test. Another reason may be related to the differences between pre-training performances in the Duration Pattern Test in both studies. In study 2, the group investigated showed worse performance compared to the trained group in study 1 during pre-training, which might have facilitated the observation of a more intense post-training improvement. Other investigators also observed such behavior (17,28). Both studies showed that individuals with worse performance and more difficulty in performing the auditory training had more post-training gains.

On the basis of these considerations, it is possible to assume that there was learning generalization, which could not be observed in study 1 for methodological reasons. Therefore, the possible occurrence of learning generalization found in study 2 suggests learning of perceptual skills. This hypothesis could explain the results of the studies in which there was improvement of verbal skills after nonverbal auditory training. With improvement of the perceptual skill, resulting from the presence of generalization, the origin of reading problems would be resolved, with a consequent improvement of verbal skills.

Such generalization was also found by Agnew et al. (16). These researchers trained a group of children with poor academic performance using the American “Fast ForWord” software (responsible for training discrimination skill and frequency ordering, among other auditory skills) and investigated whether this training influenced a Duration Pattern test applied before and after intervention. The researchers found improvement in performance after training for this test, suggesting the existence of learning generalization. Other studies also reported learning generalization, but in different situations such as untrained frequencies (29), untrained phonemes (13,30) and stimuli with different durations (31).

Another methodological issue is related to administration of the test. One of the hypotheses is that the improvement found in study 2 could be related to the number of times the tests were applied. Unlike study 1, in which each test was applied twice, in study 2, the same test was applied three times, which might have had a positive influence on test performance, perhaps because the situation was more familiar or due to task learning itself. According to Wright and Sabin (32), there is no consensus about how much training is enough to achieve learning, but several investigators have reported improvement in performance resulting from a single training session, which might be attributed to the learning of the procedure and of the task or only to processing of the stimulus used during the session (29). Thus, the number of times the test was applied in this study might have been enough to generate some change in performance. On this basis, it is also possible to assume that, if there was improvement resulting from administration of the test, the number of times the game was played (30 days) by both groups should also have been enough to cause improvement as a consequence of training.

If the hypothesis related to test administration is true, we may conclude that there was no learning generalization since the improvement observed in study 2 was probably related to other factors. It should also be considered that there probably was no learning of the auditory skills, but only learning of the trained task. This hypothesis could explain the result of the studies conducted on children with dyslexia who did not show improvement of verbal skills (written language) after nonverbal auditory training in spite of improvement regarding the auditory tests applied (14-17). This is due to the fact that there cannot be improvement of the verbal skills if there is no improvement of the perceptual skills related to auditory temporal processing, since this is considered to be the origin of reading problems detected in children with dyslexia (1-4).

Another hypothesis that may be considered is that the tests selected might have had an influence on the observation of a possible generalization. Perhaps the tests applied might require skills other than temporal order judgment. According to Mossbridge et al. (33), learning generalization from a trained situation to an untrained situation can only occur if the neural circuits modified during training also influence the performance of the untrained situation. Musiek et al. (24), for example, claimed that the Duration Pattern test is able to detect brain injuries that the Frequency Pattern test cannot detect and vice-versa. Baldeweg et al. (34) found abnormal mismatched negativity responses to changes in stimulus frequency, but not to changes in duration. These studies showed that the brain processes involved in the two tests might
be different, possibly influencing the analysis of generalization. According to Lakshminarayanan and Tallal (19), neural mechanisms are also important to determine the extension of learning transference. Perhaps these differences between tests are related to the presence of spectral cues in the Frequency Pattern test. Although this test involves analysis of temporal ordering, stimuli must be spectrally discriminated first to be ordered later. Therefore, these differences between tests might have had a negative influence on the analysis of a possible learning generalization.

Our studies showed different results regarding the presence of learning generalization of temporal order detection. Study 2 showed a possible occurrence of generalization, evidencing that after auditory training there might be improvement of verbal skills since there is learning of the perceptual skills. Study 1 showed the opposite, questioning the occurrence of learning of auditory skills and the use of auditory training as a method of effective intervention in children with dyslexia. The presence of methodological differences between the studies, as well as the relation between trained task and evaluated tasks are discussed. Further studies are necessary for a better understanding of this phenomenon.

Acknowledgments

Research supported by FAPESP (#06/507264).

References

25. Murphy CFB, Pagan-Neves LO, Wertzner HF, Schochat E. Análise acústica de características temporais de consoantes no
Table 1. Characteristics of the members of the trained and untrained groups in study 1.

<table>
<thead>
<tr>
<th></th>
<th>Trained group (N = 12)</th>
<th>Untrained group (N = 28)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gender (boys/girls)</td>
<td>9/3</td>
<td>19/9</td>
</tr>
<tr>
<td>Age (years)</td>
<td>10.9 ± 1.4</td>
<td>10.4 ± 2.1</td>
</tr>
<tr>
<td>Syllable task</td>
<td>14.7 ± 1.2</td>
<td>13.4 ± 3.2</td>
</tr>
<tr>
<td>Phonemic task (total = 16)</td>
<td>5.1 ± 4.0</td>
<td>4.8 ± 3.4</td>
</tr>
<tr>
<td>Rhyme/alliteration (total = 8)</td>
<td>6.8 ± 1.1</td>
<td>5.6 ± 2.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>7-10 years old</th>
<th>11-14 years old</th>
<th>7-10 years old</th>
<th>11-14 years old</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading words (total = 30)</td>
<td>20.4 ± 8.6</td>
<td>22.9 ± 5.7</td>
<td>17.3 ± 6.1</td>
<td>26.2 ± 3.1</td>
</tr>
<tr>
<td>Text reading (total = 10)</td>
<td>3.8 ± 3.5</td>
<td>5.6 ± 3.6</td>
<td>4.3 ± 4.2</td>
<td>8.0 ± 1.6</td>
</tr>
</tbody>
</table>

Data are reported as means ± SD. There were no statistical differences between the trained and untrained groups (Student t-test).

Table 2. Scores for each test applied before and after training to both groups in study 1.

<table>
<thead>
<tr>
<th></th>
<th>Trained group</th>
<th>Untrained group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7-10 years old</td>
<td>11-14 years old</td>
</tr>
<tr>
<td>Frequency Pattern</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>6.0 ± 3.1</td>
<td>10.9 ± 4.2</td>
</tr>
<tr>
<td>Post</td>
<td>13.2 ± 2.6*</td>
<td>13.7 ± 5.3*</td>
</tr>
<tr>
<td>Duration Pattern</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre</td>
<td>8.0 ± 3.8</td>
<td>9.4 ± 3.9</td>
</tr>
<tr>
<td>Post</td>
<td>8.0 ± 3.7</td>
<td>10.0 ± 3.7</td>
</tr>
</tbody>
</table>

Data are reported as means ± SD for a total of 20 trials. The trained group trained for 30 ± 4 days and the last stage achieved by the trained group was 6 ± 5 (duration of the stimulus: 150 ms/ISIs 300 ms; frequency: 2000 to 4925 Hz). *P < 0.05 post-training score compared to pre-training score in frequency test in trained group (analysis of variance with repeated ordinal data).
Table 3. Mean correct answers to each test applied at three different times in study 2.

<table>
<thead>
<tr>
<th>Gender (boys/girls)</th>
<th>12/6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>10.1 ± 2.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>7-10 years old</th>
<th>11-14 years old</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Frequency Pattern</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre 1</td>
<td>6.9 ± 5.3</td>
<td>6.6 ± 4.0</td>
</tr>
<tr>
<td>pre 2</td>
<td>8.5 ± 5.4</td>
<td>6.6 ± 2.8</td>
</tr>
<tr>
<td>post</td>
<td>12.2 ± 5.9</td>
<td>11.0 ± 2.3*</td>
</tr>
<tr>
<td><strong>Duration Pattern</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pre 1</td>
<td>6.7 ± 4.1</td>
<td>7.0 ± 4.9</td>
</tr>
<tr>
<td>pre 2</td>
<td>7.5 ± 5.0</td>
<td>9.8 ± 3.6</td>
</tr>
<tr>
<td>post</td>
<td>9.3 ± 5.9</td>
<td>11.2 ± 2.9*</td>
</tr>
</tbody>
</table>

Data are reported as means ± SD. The trained group trained for 31 ± 7.1 days and the last stage achieved by the trained group was 5.5 ± 4.9 (duration of the stimulus = 150 ms/ISIs = 400 ms; frequency = 1000 to 1965 Hz). *P < 0.05 post compared to pre 2 (non-parametric analysis).

Figure 1. Representation of the stimuli duration, stimulus frequency and inter-stimulus interval (ISI). The first stimulus represents a "low-high stimulus" with frequency changing from 500 to 1252 Hz and duration of 200 ms. The second stimulus represents a "high-low" stimulus with frequency changing from 1252 to 500 Hz and duration of 200 ms.
Figure 2. Screen design (A and B). A represents the stimulus discrimination phase, and B represents the ordering phase. On both screens, there is a setting (forest) and a main character (monkey) that shows a goal (to get the bananas using a bush rope). Players are instructed to listen to two or three nonverbal stimuli and to associate them with the sign displayed on the screen (Figure 2A = or ≠ for discrimination; Figure 2B ↑ or ↓ for ordering, where ↑ represents the rising stimulus and ↓ represents the falling stimulus).