The Effect of Practice on Low-Level Auditory Discrimination, Phonological Skills, and Spelling in Dyslexia

Tina Schäffler, Juliane Sonntag, Klaus Hartnegg and Burkhart Fischer*

Brain Research Group, Hansastr. 9, D-79104 Freiburg, Germany

Phonological awareness is believed to play a major role in the auditory contribution to spelling skills. The previous paper reports low-level auditory deficits in five different subdomains in 33–70% of the dyslexics. The first study of this paper reports the results of an attempt to improve low-level auditory skills by systematic daily practice of those tasks that had not been passed in previous diagnostic sessions. The data of 140 dyslexics indicate that the average number of unsolved tasks can be reduced from 3 of 5 to 1 of 5. The success rates have values of 70–80% for intensity and frequency discrimination and for gap detection, but reach only 36% for time-order judgement and 6% for side-order judgement. The second study reports that successful low-level auditory training transfers completely to language-related phonological skills and also to spelling with the largest profit in spelling errors due to poor auditory analysis. Control groups (waiting and placebo) did not exhibit significant improvements. It is concluded that low-level auditory deficits should be considered and improved by practice in order to give the dyslexics more phonological help when trying to transfer what they hear to spelling. Copyright © 2004 John Wiley & Sons, Ltd.

Keywords: auditory processing; development; dyslexia; training

INTRODUCTION

The importance of phonological deficits in dyslexia has been a matter of considerable debate (Witton et al., 1998; Stein & Talcott, 1999; Mc Anally & Stein, 1996; Tallal, 1980). There is little doubt that deficits in phonological
awareness do contribute to the problems of a dyslexic child, but these are related to language processing and language comprehension (Wright, Lombardino, King, Leonard, & Merzenich, 1997; Schulte-Körne, Deimel, Bartling, & Remschmidt, 1998). It has also been reported that low-level auditory discrimination may be impaired indicating that dyslexics may suffer from different kinds of deficits within the auditory system, specifically in the temporal auditory domain (Mc Anally & Stein, 1996) as reviewed earlier (Tallal, Merzenich, Miller, & Jenkins, 1998). Other authors, however, failed to find auditory deficits in dyslexic children (Watson & Miller, 1993). However, in a new study with a large group of dyslexics (see previous paper) it has been shown that one or more of five different independent subfunctions of low-level auditory processing may be impaired in a dyslexic child (Fischer & Hartnegg, 2004).

Among other non-auditory problems, dyslexics may also suffer from visual and optomotor deficits such as deficits in dynamic vision (Fischer, Hartnegg, & Mokler, 2000; Demp, Boynton, & Heeger, 1998; Eden et al., 1998; Livingstone, Rosen, Drislane, & Galaburda, 1991; Lovegrove, 1993; Talcott, Hansen, Assoku, & Stein, 2002a; Talcott et al., 2002b ) and/or in saccade control (Biscaldi, Fischer, & Hartnegg, 2000; Eden, Stein, Wood, & Wood, 1994) and/or in binocular stability (Fowler, 1992). There have been successful attempts to improve those aspects of the visuomotor process by daily practice by giving the children visual tasks designed to specifically challenge specific aspects of the visual and/or optomotor process (Fischer et al., 2000); (Stein & Fowler, 1985).

It has been shown that children with language-learning impairments profit from auditory training by improvements of the ability to recognize brief and fast sequences of speech and non-speech stimuli (Merzenich et al., 1996). It was concluded that slow auditory processing contributes to the deficits in phonological awareness, which in turn is a determining factor in speech comprehension (Tallal, 1980).

This study takes up the idea of daily practice to improve deficits in the low-level auditory domain: the dyslexics were diagnosed by a battery of five tasks. Each task that was performed below the 16-percentile of age matched controls was given to the child for daily practice in a adaptive and controlled training version. The aim of the training was to improve the auditory skills under consideration and hopefully also to reduce the spelling errors in dyslexia, when they try to write what they have heard (dictation).

The results of this training procedure are reported. It is shown that many but not all children could improve their capacity of low-level auditory discrimination in a specific way. In a second study the transfer of the auditory training effects to language-related phonological skills and to spelling have been investigated in a placebo-controlled study.

While language comprehension in adult subjects does not require perfect auditory perception, the acquisition of speech and especially of spelling uses the auditory signals and is facilitated by high-quality low-level auditory discrimination.

METHODS

The general methods of auditory testing have been described earlier in great detail in the previous paper (Fischer & Hartnegg, 2004). Here, we present only a short account.
Control subjects. The data of 413 control subjects were used for determining the percentile of each dyslexic child within an age-matched control group. The data of 5 age groups were available for this purpose: 7–8 (115), 9–10 (126), 11–13 (117), 14–17 (33), 18–22 (22) years; the numbers of control subjects in each group are given in brackets.

Study I. Effects of Daily Practice on Low-Level Auditory Discrimination

Dyslexic subjects: For the first study a total of 140 dyslexic subjects were recruited from the optomotor laboratory of the University of Freiburg to participate in the training. These children passed the classical pure tone auditory test of absolute thresholds over a large frequency range. Standardized German spelling and reading tests and an intelligence test (K-ABC) were used in our laboratory. If the spelling and/or reading performance was below the 25th percentile and if this performance level was at least 1 standard deviation below the IQ, a subject was classified as dyslexic (discrepancy criterion). While all these measures were available for 52 subjects, another 88 children were tested in other institutions and we did not have the exact values of the literacy diagnosis. In addition to the diagnostic tests, the 16 percentiles of each of the five auditory tasks were determined: to be included in the test group at least 1 of the 5 percentiles must have been below 16. Other details have been described in detail earlier (Biscaldi et al., 2000).

Test subjects of this main group (N = 140) were 7–21 years old. Male subjects formed 3/4 of the whole group. As in the previous paper, the test subjects were divided into four age groups containing 3, 44, 65 and 28 subjects, respectively. The oldest group contained subjects from 14 to 22 years of age. We also have the data of another 350 dyslexics, who together with the main group built an extended group of 490 subjects, whose data will be presented at the end of study I. The number of subjects in the four age groups were 110, 184, 157 and 39, respectively.

Design: The discrimination thresholds for (1) intensity, (2) frequency and (3) temporal gap in a broadband noise were determined. In the fourth and fifth task the time-order judgement for monaural (4) and the side order for binaural (5) stimulation were determined. Each task was based on a two alternative forced choice procedure. The difficulty of the test tasks was increased by decreasing the difference between the reference and the test stimulus in 30 to 40 steps of decreasing size. For each task the threshold value was defined by the last correct response preceding the first of three errors in a sequence of seven consecutive trials.

This task design was also used for the training sessions, but the training started by using only the six easiest levels of difficulty presented in random order. The difficulty was increased by one step not before the child reached a level above 75% correct responses. The first sessions at the next day always began with the difficulty level reached at the preceding day.

The physical parameters of the tasks are given below:

1. Intensity discrimination: white noise tones 300 ms in duration, reference signal 55 dB(A).
2. Frequency discrimination: reference tone 1 kHz, 300 ms in duration, 63 dB(A).

Copyright © 2004 John Wiley & Sons, Ltd.

3. Gap detection: 60 dB(A), 300 ms, white noise.
4. Time-order judgement: 1 and 1.12 kHz, 200 ms in duration, 64 dB(A).
5. Side-order judgement: clicks (10 ms) with 55 dB(A) were delivered one to the right one to the left ear in random order.

The training was carried out at home using a custom-made hand-held device with a built-in response key pad. A small LCD screen provided feedback after each trial during the training sessions. Stimuli were applied through headphones (Sony MDR-P70). Each child performed the training of a given task for 10 days with a daily session lasting about 10–15 min. The tasks begin in an easy version at day one. The difficulty was increased as the child’s performance increased. Only one task per day was trained. The tasks were always trained in the order as described above. The details of the training procedures were stored by the instruments and allowed a later analysis of the training. The corresponding training protocol was analysed and sent to the parents of all participating children. It allowed to control for compliance. If the training was not successful because of poor compliance or only small or unstable training effects the training of the corresponding task was continued another 10 days. Three criteria were used: (i) percentile of 20 was not yet reached (low-level criterion), (ii) task performance in one of the last two sessions was below that of the previous session (stability criterion), (iii) performance level was increasing slowly and did not reach a plateau to control for stability (completeness criterion). If compliance was correct and the training was still not successful, then the training was not repeated.

Data analysis: The percentiles before (pre) and after (post) the training were determined for each subject and the respective mean values were calculated. The percentages of subjects who performed above the 20th percentile after the training were determined as success rates and their respective mean percentile ranks were also calculated.

Study II: Transfer of the Training Effects to Phonological Skills and Spelling

Subjects: In the second study 41 dyslexic children participated. Because of the lack of correlations with general intelligence (see previous paper) the participants did not have to pass a standard intelligence test. Rather, we used the school grades for reading and spelling in comparison with those of other grades, e.g. mathematics. Participants were divided into three groups: the experimental group consisted of 25 subjects, the waiting group of 6 subjects and the placebo group of 11 subjects.

Procedure: While the members of the experimental group performed their individual auditory training, the members of the waiting group had no further support and—as all others—continued to visit the school. The members of the placebo group were given a dynamic visual orientation discrimination task for everyday training. This visual training had many things in common with the auditory training: the hand-held training instruments looked the same with the exception of the earphones. The daily time consumption was about the same. Both training procedures were adaptive starting with very easy versions of the training task and increasing the difficulty in parallel with an improvement in
the performance. Feedback was given in both cases. The three criteria for the continuation of the training described above were also used in this study.

Pre- and post tasks: In addition to the low-level auditory test, each participant was also tested for phonological skills using the language-related ‘Heidelberger Lautdiskrimination Test’ (H-LAD). In this test the subject has to discriminate two similarly sounding German words played from an audio compact disc player. The results obtained before and after the training (or waiting) period were transformed into percentiles.

Similarly, a standardized German spelling test, Diagnostischer Rechtschreibtest (DRT) was used before and after the training period. The results were given in numbers of spelling errors and also in percentiles. The spelling errors of the DRT were classified into three types: Type I errors rely on auditory misperception, for example: ‘treckig’ versus ‘dreckig’. Type II errors have to do with misspelling of lengthening like hit versus heat. Type III errors occur because of a poor knowledge of language rules and grammar. Type III errors, by definition, cannot be prevented by auditory analysis, for example: ‘nent’ versus ‘nennt’. The three error types are standard classifications of the German DRT.

Data analysis: Pre- and post percentile values were determined for each subject. For statistical analysis ANOVAs and t-tests for repeated measurements were applied. The specificity of the spelling improvements were assessed by analyzing the number of errors for the different types of errors.

RESULTS

Study I: Effects of Daily Practice on Low-Level Auditory Discrimination

Relation to general intelligence and spelling scores: The mean value of the IQ was $103 \pm 13$, with all single values above 85. As in the previous paper we calculated the correlation coefficients between IQ values and each of the five task variables. No significant correlations were obtained. All values were below 0.3 and none of them reached a significance level of 1%. The sizes of the training effects in this group were estimated by the percentage of reduction of unsolved tasks (percentile < 16). The reduction was 43–56% depending on age with an average of 47%. There was no correlation between training effects and IQ values ($r = 0.18; p = 0.17$).

Because of this general independence of low-level auditory task performance and its improvement from intelligence we will present the analysis of the data without referring to IQ values.

Correlation coefficients between the percentiles reached in the five auditory tasks and those reached in the spelling tests were calculated from 58 subjects. All values were below $r = 0.3$ and none of them reached a significance level of 1%. One of the reasons is that too many subjects reached percentile ranks of zero in the auditory tasks producing a floor effect. Therefore, we calculated the correlation between the spelling scores and the number of unsolved tasks. As a result we obtained $r = -0.30 (p = 0.020)$, indicating that there was a tendency of lower spelling skills corresponding to difficulties in more domains of low-level auditory discrimination.
General analysis: Each of the 140 dyslexic subjects performed the five tasks during a diagnostic test session. On average 2.94 (median=3) tasks per subject were unsolved (below 16th percentile). After the training this number was reduced to an average of 1.41 (median=1) tasks per subject. The overall training effect was estimated as a 52% reduction of unsolved tasks.

Figure 1 shows the distributions of the number of unsolved tasks before (white bars) and after the training (black bars). By definition, none of these subjects could solve all tasks before the training, because otherwise they would not have been included in the test group of this study. After the training 23% of the subjects succeeded in solving all tasks (zero unsolved tasks). Before the training 63% of the subjects failed in three or more tasks, after the training this number was reduced to 20%. Fifteen subjects performed below the 16th percentile in all five tasks. After the training only one subject was left with this kind of a complete failure.

Effect of age: Because many variables describing sensory processing and optomotor control depend on age, we also analysed the training effects with respect to age. Figure 2 shows the age curves obtained from the main group of 140 subjects (left side) and from the extended group (N = 490, right side). The two pairs of curves are very similar indicating that group size did not play a significant role. Little if any age dependence can be seen in either group for the three younger groups, neither before nor after the training. The percentage of reduction of unsolved tasks varied between 45 and 68% with the two largest values belonging to the oldest groups.

Task specific analysis: Table 1 presents the numbers and percent numbers of subjects failing or succeeding in the different tasks. The percentages of subjects failing in the diagnostic sessions are given in the first row of Table 1 (compare previous paper (Fischer & Hartnegg, 2004)). The analysis of the extended group reveals the same results. The second row shows the reduction of these percentages due to the training. The success rates (percentile above 20) are given in row 3.
Figure 3 presents the results graphically. The height of each column indicates the mean value of the individual percentiles. For each task three columns are shown: the first column (pre, black) shows the mean percentile before the training. Of course, all mean values are below 10, because only subjects with performance below the 16th percentile were included in the training group for this task. The second column (post, white) shows the mean percentile of the trained group. The third column (post, black) shows the mean percentile of the successful subjects only, to indicate that these subjects as a group did not improve just a little above the 20th percentile, but performed very well and are hardly different from the control group. Above each triplet of columns the success rate is given in percent (see also Table 1, third row).

The success rates were different for the five tasks. While the tasks 1–3 were learned by 70–80% of the subjects, task 4 was learned by only 36%, but these successful subjects reached a mean percentile of about 70. Task 5 was hardly learned at all. Only 6% (3/47) were successful.

<table>
<thead>
<tr>
<th>Intensity</th>
<th>Frequency</th>
<th>Gap</th>
<th>Time order</th>
<th>Side order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Below 16 p</td>
<td>67/140</td>
<td>101/140</td>
<td>49/140</td>
<td>90/140</td>
</tr>
<tr>
<td>pre-training</td>
<td>47.9%</td>
<td>72.1%</td>
<td>35.0%</td>
<td>64.3%</td>
</tr>
<tr>
<td>Below 16 p</td>
<td>17/140</td>
<td>18/140</td>
<td>10/140</td>
<td>51/140</td>
</tr>
<tr>
<td>post-training</td>
<td>12.1%</td>
<td>12.9%</td>
<td>7.1%</td>
<td>36.4%</td>
</tr>
<tr>
<td>Above 20 p</td>
<td>49/67</td>
<td>81/101</td>
<td>38/49</td>
<td>32/90</td>
</tr>
<tr>
<td>post-training</td>
<td>73.1%</td>
<td>80.2%</td>
<td>77.6%</td>
<td>35.6%</td>
</tr>
</tbody>
</table>

Figure 2. Age curves of the number of unsolved tasks before and after the training. The left and the right sides show the results obtained from the main group and the extended group, respectively.
From the analysis of the training protocols it was found that most children improved their performance day by day. Once arriving at the 16th percentile many subjects improved their performance further during the subsequent days and most of them reached stable values considerably better than the age-matched controls.

Substantial fractions of all of subjects with deficits remained only in the time-order (36.4%) and side-order (27.1%) domains, while the other three domains exhibit only numbers below 16%, i.e. within the range of the controls (second row of Table 1).

Since both the frequency discrimination task and the time-order task require frequency discrimination, one could argue that both tasks challenge the subjects in the same auditory domain. By contrast, 41 subjects passed the frequency task before or after the training but did not pass the time-order task after the training. They were unable to use their capacity of frequency discrimination to learn the time-order judgement. Twenty-nine subjects failed both tasks before the training but succeeded only in the frequency task and one subject failed both tasks before the training but succeeded only in the time-order task.

These numbers support the notion that the time-order judgement challenges a subsystem of auditory discrimination different from frequency discrimination.

A factor analysis was also carried out using the pretraining percentiles. As a result, four factors explain 85% of the variance. A factor analysis was also carried out for the control subjects using the original threshold values of 9-year-old children. To explain 84% of the variance three factors were needed. This indicates that there is a slight interdependence between the variables, but it is far too loose in order to predict the result of one task from the result of another in a given subject.

Figure 3. The figure shows triplets of columns one for each of the five tasks. The first (pre, black) column shows the mean percentile of the subjects before the training. The second (post, white) column shows the mean percentile of the trained group after the training. The third (post, black) column shows the mean percentile of the successful subjects after the training. This last column is presented to show that the successful subjects were performing about as well or even better than the age-matched controls.
Study II: Transfer of the Training Effects to Phonological Skills and Spelling

First of all the analysis shows that subjects with auditory deficits in the five non-linguistic tasks also failed on average the H-LAD, which is a language-related test. Figure 4 shows this result in the left column of the left part of the figure.

To assess the training transfer effects a non-parametric \( t \)-test (Wilcoxon) was performed for the H-LAD data of each of the three groups: the experimental group reached a \( p \)-value below 0.000, but for the waiting and placebo groups the \( p \)-values were 0.17 and 0.83, respectively (s. Figure 4). In other words, the transfer effect from the non-linguistic auditory training to language-related phonological skills was highly significant and specific for the training of the experimental group with no significant effects in the waiting or placebo group.

A similar analysis by non-parametric \( t \)-tests was performed with the results of the spelling tests (Figure 4 right side). The profit in spelling of the experimental group was clearly smaller than the profit of phonological skills, but larger (and significant) in the experimental group than in the two control groups both failing significance.

A transfer to spelling can hardly be expected for all the three types of errors, because it is impossible to know the spelling of a word by listening, when the correct spelling depends only on the language rules. Therefore, three separate ANOVAs were performed for the three errors types. The result was that the interaction term group \( x \) time reached significance (\( p = 0.008; F = 5.5 \)) only for type I errors, while type II and type III errors failed significance with \( p \)-values of 0.995 (\( F = 0.005 \)) and 0.97 (\( F = 0.000 \)), respectively.

This result leads to a final analysis of the data of the experimental group with respect to the error types. Figure 5 shows the diagram. \( T \)-tests were performed for the three types of errors. The \( p \)-values reached significance for type I (\( p = 0.000, F = 24 \)) and type II errors (\( p = 0.010, F = 8 \)) and failed significance (\( p = 0.028, F = 5 \)) for type III errors. In other words, the experimental group profits most by reducing the type I errors (due to poor auditory analysis), to some extent also by

![Figure 4](image_url)

**Figure 4.** The diagrams show the mean percentile reached before and after the training of the three groups. The experimental group exhibits the largest profit both in the phonological task (H-LAD) and in spelling. Vertical bars indicated the confidence intervals, \( p \)-values of non-parametric \( t \)-tests (Wilcoxon) are provided for each pre-post comparison.
reducing the type II errors and very little in reducing the type III errors (due to neglecting language rules).

DISCUSSION

The present study has shown that the daily practice of specific low-level auditory discrimination tasks increases the performance of dyslexics, who failed the 16-percentile of age-matched controls. However, not all subjects reached the performance level of age-matched controls and the success rates were clearly different for the different tasks. This fact indicates that little or no placebo effects are responsible for the success of the auditory training.

One could argue that the side-order judgement was always the last task to be trained and the subjects were no longer motivated. On the other hand, most children had already learned one or the other task before and they should profit from their experience when finally trying to learn the side-order task.

Frequency discrimination was learned by 80%, but time-order judgement was learned by only 36% of the subjects. This indicates that the two tasks challenge separate subsystems of auditory processing as concluded already from the analysis presented in the previous paper.

The second study shows that there exists a significant transfer of improvements of low-level auditory discrimination to language related phonological skills and to spelling. This finding supports the notion that low-level auditory problems play an important role in dyslexia. In fact, one could argue that low-level problems contribute to the poor phonological awareness as discussed earlier (Tallal, 1980).

It is still unknown whether the new auditory skills persist for longer periods of time. It might be important to use the new auditory capacities in everyday life, e.g. in school or at home during further training of spelling in order not to lose what has been gained in a relatively short period of time.
In addition to the several different auditory problems, a dyslexic subject may suffer from dynamic visual deficits (Fischer et al., 2000); (Talcott et al., 2002a, b) and/or from optomotor control problems, especially in the frontal lobe component of saccade control (Biscaldi et al., 2000). A corresponding multifactorial diagnosis is necessary as a basis for specific help, such as saccade and dynamic visual training, that have already been described and used successfully (Fischer & Hartnegg, 2000).

Unless almost all the several sensory problems are solved in a dyslexic child one can hardly expect that regular teaching of spelling and reading helps. This study contributes to increase the number of diagnostic and therapeutic tools available in dyslexia as well as in the context of other learning processes depending on almost perfect auditory, visual and optomotor functions.

References


