Original Article

Trends in Hearing
Volume 22: 1–14

DOI: 10.1177/2331216518790902
journals.sagepub.com/home/tia
SAGE

Long-Term Training-Induced Gains of an Auditory Skill in School-Age Children As Compared With Adults

Y. Zaltz[1], D. Ari-Even Roth[1], A. Karni[2,3], and L. Kishon-Rabin[1]

Abstract

The few studies that compared auditory skill learning between children and adults found variable results, with only some children reaching *adult-like* thresholds following training. The present study aimed to assess auditory skill learning in children as compared with adults during single- and multisession training. It was of interest to ascertain whether children who do not reach *adult-like* performance following a single training session simply require additional training, or whether different mechanisms underlying skill learning need to reach maturity in order to become *adult-like* performers. Forty children (7–9 years) and 45 young adults (18–35 years) trained in a single session. Of them, 20 children and 24 adults continued training for eight additional sessions. Each session included six frequency discrimination thresholds at 1000 Hz using adaptive forced-choice procedure. Retention of the learning-gains was tested 6 to 8 months posttraining. Results showed that (a) over half of the children presented similar performance and time course of learning as the adults. These children had better nonverbal reasoning and working memory abilities than their *non-adult-like* peers. (b) The best predicting factor for the outcomes of multisession training was a child's performance following one training session. (c) Performance gains were retained for all children with the *non-adult-like* children further improving, 6 to 8 months posttraining. Results suggest that mature auditory skill learning can emerge before puberty, provided that task-related cognitive mechanisms and task-specific sensory processing are already mature. Short-term training is sufficient, however, to reflect the maturity of these mechanisms, allowing the prediction of the efficiency of a prolonged training for a given child.

Keywords

mature auditory learning, auditory training, adult-like learning, frequency discrimination, difference limen for frequency

Date received: 10 November 2017; revised: 24 June 2018; accepted: 29 June 2018

Introduction

The ability of a child to gain from training using an auditory task is of special interest because the auditory modality is crucial for his or her development and well-being from infancy through adulthood. It is the primary modality by which a child monitors the environment (e.g., Kishon-Rabin & Boothroyd, 2018) and is considered vital for his or her early cognitive, language, speech, and social development (e.g., Boothroyd, 1997). In school years, efficient auditory processing also contributes to the development of literacy and plays an important role in academic achievements (e.g., Boothroyd & Boothroyd-Turner, 2002). Therefore, assessing whether a child can reach efficient learning of an auditory skill is important for theoretical and practical reasons. Theoretically, it may reflect the maturity of the different underlying mechanisms for auditory learning. Practically, estimating when a child is most susceptible to the positive outcomes of auditory training may help design efficient auditory training protocols.

---

[1]Department of Communication Disorders, Sackler Faculty of Medicine, Tel Aviv University, Israel
[2]Department of Human Biology, Faculty of Natural Sciences and Education, The Edmond J. Safra Brain Research Center for the Study of Learning Disabilities, University of Haifa, Israel
[3]Division of Diagnostic Radiology, The Chaim Sheba Medical Center, Ramat Gan, Israel

Corresponding Author:
L. Kishon-Rabin, Department of Communication Disorders, Steyer School of Health Professions, Sackler Faculty of Medicine, Tel Aviv University, Israel.
Email: lrabin@tauex.tau.ac.il

Several mechanisms are assumed to contribute to skill learning, including top-down (cognitive) and bottom-up (sensory) processing mechanisms. Multiple top-down abilities such as working memory and executive functions including attention and internal control mechanisms were suggested to be involved in the tuning and adaptation processes that take place in the initial phase of learning (Hauptmann & Karni, 2002; Hauptmann, Reinhart, Brandt, & Karni, 2005; Karni et al., 1998). These assist in the formation of effective task solution strategies and in reducing response bias (e.g., Ahissar & Hochstein, 1997; Jones, Moore, Shub, & Amitay, 2015; Karni & Sagi, 1993; Vakil, Hassin-Baer, & Karni, 2014). Bottom-up statistical learning processes were suggested to be activated with repeated exposure to the trained stimuli. These processes are induced by internal (neural) feedback from the updating of *synaptic weights* at the sensory or motor level, based on the statistical distribution of the trained stimuli (e.g., Janacsek, Fiser, & Nemeth, 2012). Statistical learning is considered mostly implicit, without awareness of what has been learned. Auditory sequence learning in infants, for example, has been attributed to statistical learning (e.g., Saffran, Johnson, Aslin, & Newport, 1999). Finally, neuronal memory consolidation processes were suggested to occur over time (e.g., Ari-Even Roth, Kishon-Rabin, Hildesheimer, & Karni, 2005; Brashers-Krug, Shadmehr, & Bizzi, 1996; Karni, 1996). During consolidation, neural representations of the trained task are established, and presumably, structural synaptic changes are completed. These allow for long-term retention of the learned skills (e.g., Dudai, 2012). The fact that some of the mechanisms that underlie skill learning involve executive functions that may not be fully developed in childhood (e.g., Flavell, Miller, & Miller, 1993; Jones et al., 2015; Luna, Garver, Urban, Lazar, & Sweeney, 2004; Toga, Thompson, & Sowell, 2006) raises the possibility that children will be at a disadvantage in acquiring new auditory skills compared with adults.

While many studies so far examined auditory skill learning in children (e.g., Edwards, Giaschi, & Low, 2005; Halliday, Taylor, Millward, & Moore, 2012; Millward, Hall, Ferguson, & Moore, 2011; Moore, Rosenberg, & Coleman, 2005; Soderquist & Moore, 1970; Tomblin & Quinn, 1983), only a few compared the learning in children to that of adults within the same study (e.g., Halliday, Taylor, Edmondson-Jones, & Moore, 2008; Huyck & Wright, 2011, 2013; Zaltz, Ari-Even Roth, & Kishon-Rabin, 2017). Of these, two studies compared the two age groups following a *single* session of training (Halliday et al., 2008; Zaltz et al., 2017) and two studies following *multiple* sessions of training (Huyck & Wright, 2011, 2013). The results, however, were inconclusive. Halliday et al. (2008), for example, reported that following a single session of training on a frequency discrimination task, 6- to 11-year-old children were divided to one of three subgroups based on their pattern of results. The first subgroup showed poorer thresholds than the adults (with no evidence of within-session learning) and were termed *non-adult-like*. The second subgroup showed *adult-like* thresholds at the first measurement but no evidence of continued learning. The third subgroup showed *non-adult-like* thresholds at the first measurement but demonstrated significant within-session learning, reaching *adult-like* thresholds at some point during training. Age, nonverbal IQ, and attention skills were associated with performance, with lower IQ and younger age for the *non-adult-like* subgroup. Similar findings were also reported in a more recent study (Zaltz et al., 2017). In contrast, Huyck and Wright (2011), for example, reported that not even one of their 11 year olds who trained on a temporal-interval discrimination over 10 sessions showed significant training induced gains compared with half of the 14-year-old adolescents and all the adults who showed significant and efficient learning. The authors concluded that mature perceptual learning does not emerge until late in adolescence. Similar findings were reported in a later study of Huyck and Wright (2013), who trained adolescents in a backward masking task for 10 sessions.

The poor learning reported for the adolescents in the studies of Huyck and Wright (2011, 2013) may be explained by the tasks that were used for training. It is possible that these tasks (temporal-interval discrimination and backward masking) required processing abilities that are not fully matured in childhood (e.g., Buss, Shuman, Grose, & Hall, 2013; Hartley, Wright, Hogan, & Moore, 2000). Thus, different results may emerge if the task chosen for assessing training-induced gains following multisession training in children would be a task that already has shown to induce significant learning for some children within a single training session. Using such a task would allow us to examine whether multiple training sessions can be as beneficial for children as they are for adults, leading to an *expert* performance on the trained task. It may also allow determining whether a child who presents *non-adult-like* performance within a single training session simply requires more experience with the trained task in order to reach *adult-like* performance or other processes are required to reach target performance. These insights will help unravel the developmental trajectories of the different cognitive or sensory mechanisms that underlie mature auditory skill learning. The purpose of the present study was, therefore, to assess auditory skill learning in children as compared with adults during single- and multisession training using a frequency discrimination task.

Materials and Methods

Participants

Forty-five young adults 18 to 30 years (mean age = 23.34, $SD = 0.45$) and 40 children 7 to 9 years (mean age = 8.1, $SD = 0.08$) took part in the training groups of the present study.[1] The age of the children was selected based on a previous study that demonstrated that children more than 7 years old are capable of performing a difference limen for frequency (DLF) task using the same paradigm that is used for young adults (Zaltz et al., 2017). A control group of 13 children 7 to 9 years old (mean age = 7.78, $SD = 0.14$) was tested for the effect of maturation over time. All participants met the following criteria: (a) normal hearing sensitivity in both ears (pure-tone air conduction thresholds $\leqslant 15$ dBHL at octave frequencies of 500–4000 Hz; Acoustical Society of America, 1996), (b) no history of language or learning disorders, (c) no known attention deficit disorders, (d) minimal or no musical training (less than 1 year), and (e) no previous experience in psychoacoustic testing. Criteria 2 to 5 were based on self-reporting for the adults and on parents' reporting for the children. The adults were paid for their participation, and the children were given stickers and prizes during and at the end of each training session. All participants were from mid-high socioeconomic status based on parent's education and residential area of living. Participants were recruited through fliers and social media. Informed consent was obtained from the parents of all the tested children and from the adult participants. The study was approved by the institutional review board of Tel Aviv University and by the human experimentation ethics committee of the Sheba Medical Center.

Stimuli

Stimuli consisted of a 1000 Hz reference pure-tone and 200 different comparison tones (Zaltz et al., 2017). The comparison tones varied from 1001 Hz to 1200 Hz in 1 Hz steps. All stimuli lasted 300 ms and were gated with rise or fall time cosine ramps of 25 ms. The interstimulus interval was 500 ms. Stimuli were 16-bit digitally generated at a sampling rate of 22050 Hz using Sound-Forge 7.0 software. They were delivered from an IBM compatible personal computer, via an external sound card and a GSI-61 audiometer and were presented monaurally via THD-50 headphones at a comfortable level (55 dBSL above individual thresholds at 1000 Hz).

DLF Threshold Measurements

The frequency discrimination task that was chosen for the present study is considered to reflect a basic psychoacoustic skill that is necessary for efficient language processing in adults and in infants (e.g., Kishon-Rabin, Segal, & Algom, 2009; Muller, Friederici, & Mannel, 2012) and therefore has ecological validity. A three-interval, two-alternative, forced-choice adaptive procedure (3I2AFC) was used for the DLF threshold measurements similar to a paradigm described earlier (Zaltz, Ari-Even Roth, & Kishon-Rabin, 2011, Zaltz et al., 2017). Each trial consisted of three stimuli: two reference tones and one target tone. The first stimulus in each trial was the reference tone, and the target tone was presented randomly as either the second or the third in a sequence. A two-down, one-up tracking procedure was used in order to estimate the frequency difference corresponding to the 70.7% correct point on the psychometric function (Levitt, 1971). The initial step size (40 Hz) was cut by half every turn-point until reaching a minimal step size of 1 Hz. Thresholds were calculated as the geometric mean of the DLFs at six turn-points with the minimal step size. Each DLF threshold measurement comprised approximately 35 to 50 trials that included approximately 105 to 150 stimuli overall. The presented stimuli were also indicated by visual lights on the computer monitor. Participants responded by clicking the computer's mouse on the light that corresponded to the stimulus that was different. Visual corrective feedback was provided immediately after each response, indicating the actual variant stimulus. There was no time limit for the response. The next trial was presented only after the participant keyed in his or her response.

Cognitive Assessment

Nonverbal reasoning, auditory memory capacity, and auditory working memory abilities were assessed for all the trained children using the "Raven's standard Progressive Matrices" test (Raven & Court, 1998), a forward digit span subtest of the "Wechsler intelligence scale for children," and a backward digit span subtest of the "Wechsler intelligence scale for children" (Wechsler, 1991), respectively. The Raven's standard progressive matrices test consists of 60 visual patterns with a missing piece, which are divided into five sets of 12. The children were required to select one of six or eight patterns in order to complete correctly the visual display. In each set, the completion task is initially easy and becomes progressively more difficult. Each child received a Raven score based on the relative percentage of his or her correct completed patterns. In the forward and backward digit span subtests of the Wechsler intelligence scale, the children heard sequences of numbers (e.g., 2, 5, 4, and 1) and were asked to repeat them in the same or in the reverse order, respectively. The passing criterion to the next longer sequence was two successful repetitions of sequences of similar length. Each child received forward and backward digit span scores based on the number of the correctly repeated sequences.

Experiment Design

Single-session training. All the trained participants took part in a single training session and a follow-up testing session. Prior to testing, a short familiarization phase was provided in which the listener's task was to discriminate a 1500 Hz target from the 1000 Hz reference. This was performed until 10 successively correct responses were made, ensuring that the listener understood the requirements of the task. Overall, the training session included six DLF measurements and lasted about 30 to 45 min. A short break of 5 to 8 min was given between the first three and the last three DLF measurements. In the follow-up session, spaced 1 to 3 days apart (testing session), another three DLF measurements were conducted. Overall, the participants performed nine DLF measurements while listening to approximately 315 to 450 trials (i.e., approximately 945–1350 stimuli in total).

Multisession training. Twenty-four adults (mean age = 22.89, $SD = 0.45$) and 20 children (mean age = 8.05, $SD = 0.11$) continued training for eight additional training sessions. These participants underwent nine training sessions with successive sessions spaced 1 to 3 days apart. All training sessions were identical to the first one. In addition, a testing session that consisted of only three DLF measurements was conducted 1 to 3 days following the end of training (at the 10th session). Overall, the participants who underwent multisession training listened to approximately 2,065 to 2,850 training trials (i.e., approximately 6,195–8,550 stimuli in total) throughout the entire training period. Half of the participants were trained in the right ear and half in the left ear. All training sessions were conducted in a sound-treated room (background noise $\leqslant 45$ dB SPL).

The ability to generalize the learning gains to untrained conditions was tested following training and will be reported in a follow-up study.

Long-Term Retention

Six to eight months after the termination of the multi- and single-session training, a retention session was performed which included three DLF measurements in the trained task. All the trained participants except for three adults (one from the single-session trained group and two from the multisession trained group) and one child (from the single-session trained group) participated in this session. To control for the influence of the children's natural cognitive development during the 8 to 10 months (between the first training session and the retention session), an additional control group of children was tested in two sessions, 8 to 10 months apart. Each testing session comprised three DLF measurements at 1000 Hz. No training was given between the two sessions.

Data Analysis

Two dependent variables were used: Frequency discrimination thresholds presented in relDLF% ($\Delta f/f \times 100$) and within-measurement variance (i.e., the variance between all the turn-points within an adaptive DLF measurement). For the statistical analyses, data were log-transformed in order to normalize the distribution of the relDLF% and variance (Kolmogorov–Smirnov test: $p > .05$) and to allow parametric statistics. Because there were missing measurements (about 4% of the children's DLF measurements were not completed due to fatigue or other interferences), missing values were imputed by calculating the mean of the nearest two measurements in order to use repeated-measures analyses of variance (ANOVAs). All post hoc and simple effects analyses were conducted following Bonferroni corrections. Latent growth curve (LGC) analysis was used to model change in the trajectory in the repeated measures data (e.g., McArdle, 2012). When the variability of the trajectory parameters was found significant, we proceeded to latent class growth analysis (LCA) which allowed to identify homogeneous subgroups of participants characterized by similar patterns of change over time (e.g., Jung & Wickrama, 2008).

Results

Single-Session Training

The time course of learning. The DLF thresholds during the training and testing sessions for the children are shown in Figure 1. It can be seen that although, on average, the children performed worse than the adults, both groups showed some improvement in performance following training. A two-way repeated measures ANOVA was conducted to assess *within-session* learning with measurement (1–6) as the within-subject variable and group (children, adults) as the between-subject variable. The results revealed a significant effect of group, $F(1, 84) = 50.17$, $p < .001$, $\eta^2 = 0.381$, with the adults showing better thresholds (mean relDLF% $= 1.02 \pm 1.00$) compared with the children (mean relDLF% $= 4.36 \pm 4.66$) and a significant effect of measurement, $F(5, 84) = 29.6$, $p < .001$, $\eta^2 = 0.260$, with linear ($p < .001$) and quadratic ($p < .001$) effects reflecting significant learning. A trend for significant Group × Measurement interaction was shown, $F(5, 84) = 1.84$, $p = .078$, $\eta^2 = 0.062$, with significant Group × Measurement linear effect ($p = .013$) reflecting different time course of learning for the children and the adults.

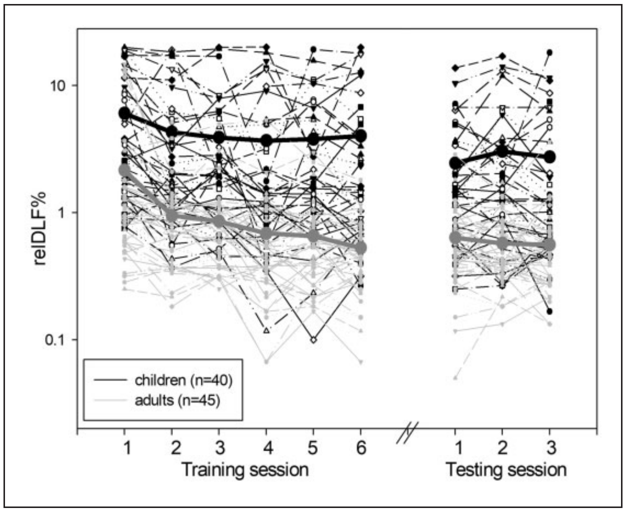

Figure 1. DLF performance during and following a single training session for the single- and multisession trained children and adults. Individual frequency discrimination thresholds (thin lines) and mean groups (thick lines) in relDLF% ($\Delta f/f \times 100$) in the six DLF measurements at the training session and the three DLF measurements at the testing session (1–3 days later) for adults and children. RelDLF% = relative difference for frequency in percentage.

LGC analysis was conducted on the six measurements of the training session. The intercept was not significantly different from zero ($p = .55$), while the slope was equal to $-.07$ ($p < .001$). The age-group was found as a significant predictor of the individual intercepts ($B = 0.48$, $p < .001$) and of slopes ($B = 0.04$, $p = .004$). These results indicated that the mean DLF thresholds were higher (worse) and the slopes were shallower (slower learning) for the children, as compared with adults. The residual variance of the trajectory parameters in this model was not significant ($p = .20$), and therefore, the LCA analysis was not conducted.

To assess *between-sessions* learning, two-way ANOVA with repeated measures was conducted with measurement (last DLF measurement in the training session, first DLF measurement at the testing session) as the within-subject variable and group (children, adults) as the between-subjects variable. Results showed significant effect of group, $F(1, 84) = 50.562$, $p < .001$, $\eta^2 = 0.373$, with no significant effect of measurement, $F(1, 84) = 1.974$, $p = .164$, but with significant Group × Measurement interaction, $F(1, 84) = 13.256$, $p < .001$, $\eta^2 = 0.143$. Simple effects analysis revealed significant improvement between the measurements *only* for the children ($p = .001$), reflecting delayed gains in performance for this group (with an improvement of 1.59 [39%] in their mean relDLF%).

LGC analysis was also conducted on the testing session (measurements 7–9) to assess the effect of the training session on performance and its change over time. In this analysis, the intercept and the slope were not significantly different from zero ($p > .10$) and neither was the variance of the slope ($p = .76$). However, the variance of the intercept was significant ($p < .001$), and it remained significant after age-group was added to the analysis as a predictor of the intercept. This result allowed us to conduct LCA analysis in order to test whether the sample could be divided into homogeneous subgroups based on their mean DLF. Results showed that the participants in the testing session were grouped to one of two subgroups: The first subgroup included 44 adults (44/45 or 97.8%) and 22 children (22/40 or 55%), and the second subgroup included 18 children (45%) and a single adult. A significant difference ($p < .05$) was shown in the variance of the intercept between these two distinct subgroups, with the first subgroup showing lower (better) DLF thresholds ($0.13 \leqslant$ relDLF% $\leqslant 1.43$) compared with the second subgroup ($1.42 \leqslant$ relDLF% $\leqslant 13.89$). In other words, following a single training session, more than half of the children were found to perform as well as the adults. These children were termed *adult-like*, whereas those that belonged to the second group were termed *non-adult-like*.

Results of the single training session were also assessed for the *adult-like* children, *non-adult-like* children, and adults with respect to their within-*measurement* variance at the testing session. One-way ANOVA that was conducted on this mean within-measurement variance showed a significant effect of subgroup, $F(2, 43) = 27.176$, $p < .001$. Post hoc analysis revealed larger within-measurement variance for the *non-adult-like* children (mean variance = 194.20 ± 195.43), as compared with the *adult-like* children (mean = 57 ± 54.10) and the adults (mean = 40.35 ± 26.79; $p < .001$), with no significant difference between the two latter subgroups ($p = .810$).

Individual performance. Figure 2 (left) shows the individual thresholds at the testing session (mean measurements 7–9) of the children that belonged to the *adult-like* subgroup and those that belonged to the *non-adult-like* subgroup based on the LGC analysis. Also shown for comparison are the data of the adults with their mean thresholds (short solid line) and 1.5 *SD* above their mean (horizontal dashed line). It can be seen that all the *adult-like* children (with one exception) reached thresholds that were within 1.5 *SD* of the adult's mean thresholds following a single training session. Thus, for the remainder of this article, an *adult-like* performance was defined as performance that was within 1.5 *SD* of the adult's mean performance. Note that there were also two adults who exceeded the range of mean+1.5 *SD*. Their data, however, were not removed from the group mean of the adults.

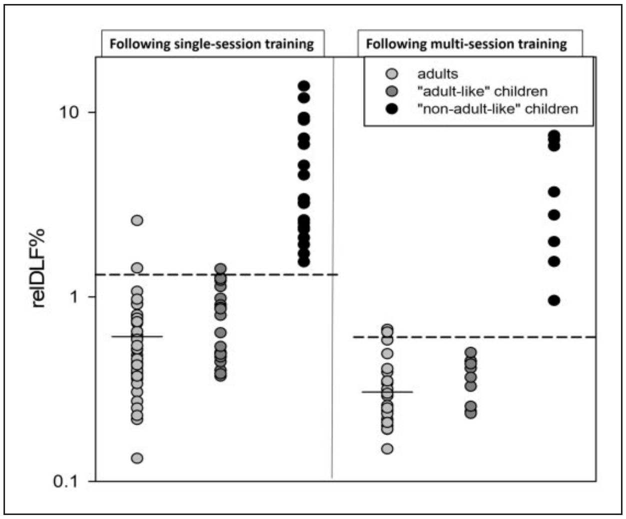

Figure 2. Mean individual thresholds following single-session training (left: mean measurements 7–9 at the second session) and following multisession training (right: mean measurements 54–57 at the 10th session) for the *adult-like* children, *non-adult-like* children, and adults. The division to *adult-like* and *non-adult-like* subgroups was based on LGC statistical analysis of performance following the first training session. The short horizontal solid line shows the mean of the adult's performance. The horizontal dashed line shows 1.5 *SD* above the adult's mean. RelDLF% = relative difference for frequency in percentage.

Predicting factors for performance. Several analyses were conducted in order to assess whether any of the background factors can explain the differences between the two subgroups of children (*adult-like* and *non-adult-like*). Independent sample *T* tests comparing age and cognitive abilities showed that the two subgroups of children differed in nonverbal reasoning (Raven scores of 67.08 ± 1.81 vs. 52.67 ± 3.39; $t = 3.818$, $p < .001$) and in working memory (backward digit span scores of 5.65 ± 0.41 vs. 4.30 ± 0.26; $t = 2.762$, $p = .009$), with the *non-adult-like* children showing poorer scores. No significant age differences were found between the *adult-like* and *non-adult-like* children ($p > .05$). A two-way repeated measures ANOVA assessing differences in pure tone thresholds between the two subgroups of children showed no significant difference between the two subgroups, $F(1,\ 37) = 0.128$, $p = .723$, and no significant Subgroup × Frequency interaction, $F(7,\ 37) = 0.739$, $p = .636$, suggesting similar hearing sensitivity for both subgroups. Pearson correlation analysis showed no significant correlations between the average hearing sensitivity in the right or left ears (in dB HL) and the relDLF% thresholds following training ($r = 0.60$, $p = .716$, $r = 0.114$, $p = .460$, respectively).

Multisession Training

Because statistical analysis following single-session training identified two subgroups of children (*adult-like* and *non-adult-like*), all further analyses were conducted with respect to these two subgroups.

The time course of learning. The time course of learning over the entire nine training sessions for the participants who continued training is shown in Figure 3. It can be seen that all three groups continued to improve across sessions. However, the performance of the *adult-like* children shadowed that of the adults throughout all sessions. In contrast, the gap in performance between the *non-adult-like* children and the adults continued to increase throughout the training sessions. Differences can also be observed in the within-session learning. While an improvement or a plateau in thresholds was shown for the adults and *adult-like* children within each training session, the *non-adult-like* children showed worsening of thresholds in most of the sessions (indicated in Figure 3 by pointing arrows).

A three-way repeated measures ANOVA was conducted only on the DLF thresholds from the additional training sessions with session (3–9) and measurement (1–6) as the within-subject variables and subgroup (adults, *adult-like*, *non-adult-like*) as the between-subject variable. Results showed a significant effect of subgroup, $F(2,\ 41) = 93.599$, $p < .001$, $\eta^2 = 0.820$. Post hoc analysis revealed that all three subgroups differed from each other significantly: The *non-adult-like* children (mean relDLF% = 4.60 ± 2.19) had poorer performance than the adults and the *adult-like* children (mean relDLF% = 0.37 ± 0.19 and 0.54 ± 0.19, respectively; $p < .001$), and the *adult-like* children had poorer performance than the adults ($p = .030$). There was a significant effect of session, $F(6,\ 41) = 3.945$, $p = .002$, $\eta^2 = 0.088$, with significant linear effect ($p = .001$), but with no significant Session × Subgroup interaction, $F(12,\ 41) = 0.906$, $p = .532$, indicating similar learning rate across sessions for all three subgroups. There was a significant effect of measurement, $F(5,\ 41) = 3.389$, $p = .006$, $\eta^2 = 0.076$, and significant Measurement × Subgroup interaction, $F(10,\ 41) = 6.047$, $p < .001$, $\eta^2 = 0.228$. A linear Measurement × Subgroup effect ($p < .001$) confirmed significant differences in within-session learning between the subgroups, coinciding with the trends of the measurements for each subgroup as shown in Figure 3. Three repeated-measures ANOVA that were conducted separately for each subgroup revealed a within-session improvement for the adults, $F(5,\ 23) = 4.281$, $p = .002$, $\eta^2 = 0.157$, with a significant linear effect, $F(1,\ 23) = 14.551$, $p = .001$, $\eta^2 = 0.388$, a within-session worsening for the *non-adult-like* children,

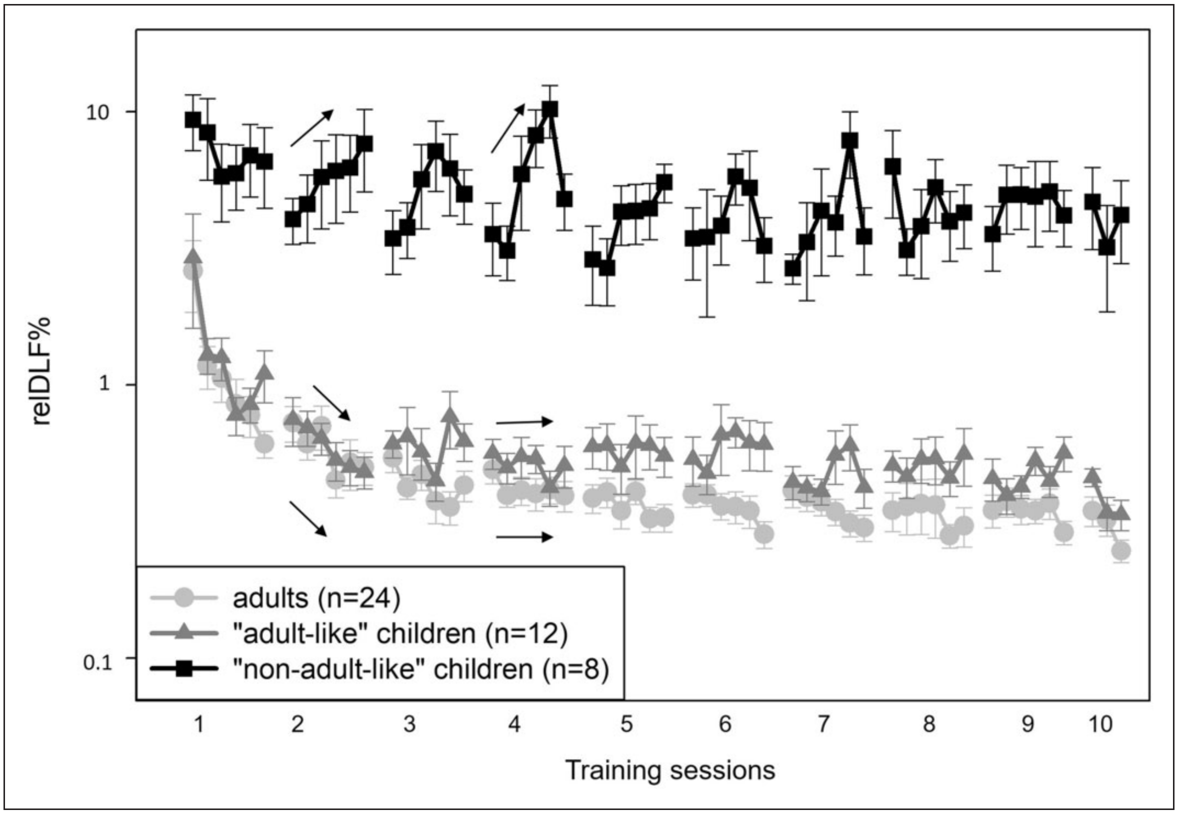

Figure 3. Mean relDLF% performance (±*SE*) during and following multisession training for *adult-like* children, *non-adult-like* children, and adults. Sessions 1 to 9 (training) included six DLF measurements each, whereas the 10th session included three DLF measurements. RelDLF% = relative difference limen for frequency in percentile; *SE* = standard error.

$F(5, 7) = 5.221$, $p = .009$, $\eta^2 = 0.427$, with a significant linear effect, $F(1, 7) = 27.949$, $p = .001$, $\eta^2 = 0.800$, and no within-session change in performance was shown for the *adult-like* children, $F(5, 11) = 1.485$, $p = .214$.

Further support for the different time course of within-session learning of the *non-adult-like* children compared with the *adult-like* children and the adults, stems from the within-*measurement* variance. Figure 4 displays the average within-measurement variance across the training sessions for each of the subgroups. It can be seen that the average variance was considerably greater for the *non-adult-like* children compared with the adults and *adult-like* children throughout the training sessions. Two-way repeated measures ANOVA was conducted on the within-measurement variance in the additional training sessions with session (3–9) as the within-subject variable and subgroup (*adult-like*, *non-adult-like*, adults) as the between-subject variable. This analysis showed a significant effect of subgroup, $F(1, 41) = 116.994$, $p < .001$, $\eta^2 = 0.851$, with post hoc analysis revealing larger within-measurement variance for

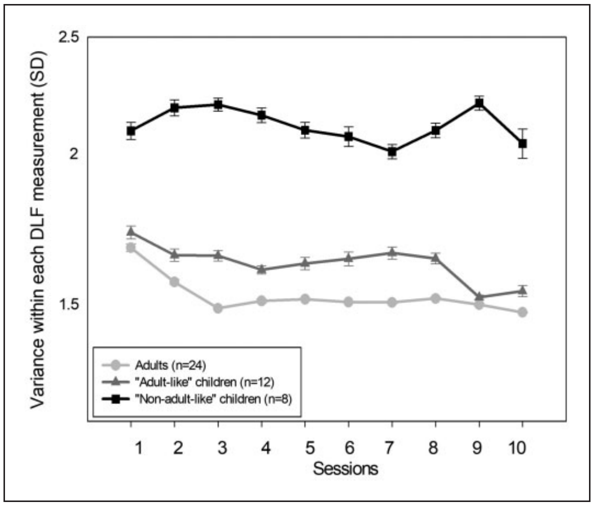

Figure 4. Mean within-measurement variance (±*SE*) throughout the multisession training for the adults, *adult-like*, and *non-adult-like* children. Sessions 1 to 9 (training) included six DLF measurements each, whereas the 10th session included three DLF measurements. DLF = difference limen for frequency; *SE* = standard error.

the *non-adult-like* children (mean variance $= 150.50 \pm 76.42$) compared with the adults (mean $= 32.78 \pm 8.53$) and the *adult-like* children (mean $= 47.09 \pm 31.78$; $p < .001$) and no significant difference between the *adult-like* children and the adults ($p > .05$). There was no significant effect of session and no significant Session × Subgroup interaction ($p > .05$).

Individual performance. To examine individual learning throughout the entire training and testing period for the multisession trained participants (Sessions 1–10), linear regression functions were fitted to the DLF thresholds for measurements 1 through 57 ($df = 56$) for each participant. Participants who had significant linear regression ($p < .05$) with a negative slope were defined as *learners*. An individual analysis revealed a similar proportion of learners for the *adult-like* children (11/12, 92%) and the adults (22/24, 92%; Fisher exact test: $p > .99$). A smaller proportion of *learners* was shown for the *non-adult-like* children (4/8, 50%) as compared with the other two subgroups (Fisher exact test: $p = .023$). By the end of training (at the testing session), all the *adult-like* children and none of the *non-adult-like* children reached *adult-like* performance, as measured at that point in time (Figure 2, right).

Predicting factors for performance. To assess the contribution of different factors to the performance of the children following multisession training (i.e., tested at the 10th session), several Pearson coefficient correlations were conducted as detailed in Table 1. Of the cognitive factors, scores of nonverbal reasoning (Raven test) were the only factor associated with DLF performance following training as well as with the variance within and between measurements. These associations were negative, that is, higher (better) scores on the Raven test were associated with lower (better) relDLF thresholds and smaller within- and between-measurement variances. The strongest association was found between the DLF measurements following a single training session (measurements 7–9) and the DLF performance following multisession training ($r = 0.923$, $p < .001$). Specifically, lower (better) relDLF thresholds following a single training session were associated with better relDLF thresholds following multisession training, suggesting that a child's performance *following* the first training session best predicted how he or she would perform following multiple training sessions.

Retention of the Training-Induced Gains

Six to 8 months posttraining, no deterioration in thresholds was shown in the trained conditions for both subgroups of children and for the adults, as shown in Figure 5. Furthermore, the *non-adult-like* children not only retained their learning-induced gains but also showed improvements over time, even though no further training was provided. An improvement in thresholds was also shown for the control group of children, who did not undergo training. A two-way repeated measures ANOVA was conducted on the relDLF% with session (end of training, retention) as the within-subject variable

Table 1. Pearson Coefficient Correlations for the Children Who Underwent Multisession Training.

	Raven	age	Wechsler forward	Wechsler backward	S1DLFf	S1DLFl	S2DLF
Correlations							
S10DLF							
Pearson correlation	−.636**	−.043	−.035	−.398	.831**	.852**	.923**
Significance (two tailed)	.003	.858	.884	.082	.000	.000	.000
Within-measurement variance							
Pearson correlation	−.530*	−.075	−.090	−.504*	.790**	.881**	.908**
Significance (two tailed)	.016	.754	.706	.024	.000	.000	.000
Between-measurement variance							
Pearson correlation	−.553*	.033	.081	−.159	.735**	.595**	.589**
Significance (two tailed)	.011	.892	.733	.504	.000	.006	.006

Note. Pearson coefficient correlations for the children who underwent multisession training with cognitive factors (Raven score, Wechsler Forwards and Backward Digit Span Scores), age, mean first three DLFs (naïve performance), mean last three DLFs in the first session (Measurements 4–6), and mean first three DLFs at the second day (following a single training session) as independent variables and results of training, that is, mean three DLF thresholds, within-measurement variance, and between-measurement variance at the 10th session, as dependent variables. Mean first three DLFs = S1DLFf; mean last three DLFs in the first session = S1DLFl; mean first three DLFs at the second session = S2DLF; mean last three DLFs = S10DLF.

*Correlation is significant at the .05 level (two tailed).

**Correlation is significant at the .01 level (two tailed).

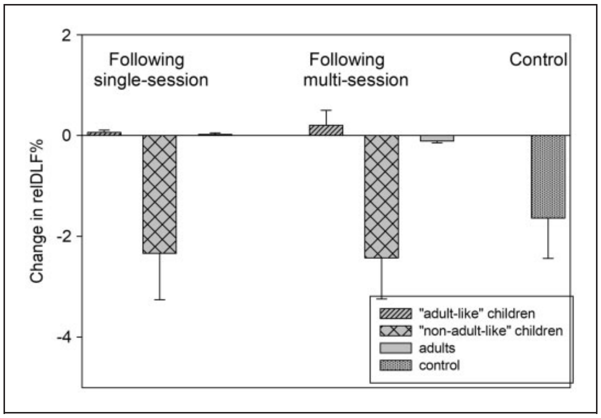

Figure 5. Mean change in relDLF% performance (±1 *SE*) between the end of training (three measurements at the testing session which was the second for the single-session trained participants and 10th for the multisession trained participants) and retention session. Results are shown for the adults, *adult-like* children, and *non-adult-like* children, separately for the single- and multisession trained participants. Also shown are the mean change in relDLF% (±1 *SE*) of the control group of children, who received no training and was tested at 1 kHz 8 to 10 months apart. Note that a negative value reflects an improvement in thresholds between the first and second session. RelDLF% = relative difference limen for frequency in percentile; *SE* = standard error.

and training duration (single session, multisession) and subgroup (*adult-like* children, *non-adult-like* children, adults) as the between-subjects variables. Results showed significant effect of training duration, $F(1, 72) = 12.355$, $p = .001$, $\eta^2 = 0.146$, with the multisession trained participants having better thresholds (mean relDLF% = 1.2 ± 1.6) as compared with the single-session trained participants (mean relDLF% = 1.66 ± 1.76). There was a significant effect of subgroup, $F(2, 72) = 99.834$, $p < .001$, $\eta^2 = 0.735$, and of session, $F(1, 72) = 15.145$, $p < .001$, $\eta^2 = 0.174$, with significant Subgroup × Session interaction, $F(2, 72) = 14.934$, $p < .001$, $\eta^2 = 0.293$. Post hoc analysis revealed that only the *non-adult-like* children improved their thresholds at the retention session (from 4.43 ± 0.92 to 2.24 ± 1; $p < .001$), with no continued improvement at the retention session for the *adult-like* children ($p = .650$) and the adults ($p = .706$). No other main effects or interactions were found significant. Individual data of the children revealed that 11.1% (2/18) of the *non-adult-like* children reached *adult-like* performance (in the range of + 1.5 *SD* of the adult's mean) at the retention testing.

To examine whether the improvement that was shown for the *non-adult-like* children could be attributed to natural development, their performance was compared with that of the control group of children, who did not undergo training. A two-way ANOVA with repeated measures was conducted with session (first three measurements in the first session and retention for the *non-adult-like* children and the two sessions for the control group) as the within-subject variable and subgroup (*non-adult-like* single-session trained, *non-adult-like* multisession trained, control) as the between-subjects variable. Results showed no significant effect of subgroup, $F(2, 29) = 0.131$, $p = .878$, with significant effect of session, $F(1, 29) = 66.419$, $p < .001$, $\eta^2 = 0.696$, and significant Subgroup × Session interaction, $F(2, 29) = 5.861$, $p = .007$, $\eta^2 = 0.288$. Significant linear effect for session ($p < .001$) and for Session × Subgroup ($p = .007$) reflected differences in improvement size (between the first and second sessions) between the subgroups. Post hoc analyses showed significant improvements for all three subgroups ($p < .017$) with no significant differences between the subgroups in either the first or second sessions ($p > .05$).

Within-measurement variance was calculated for each subgroup at the retention session, and a univariate analysis was conducted testing main effects of subgroup (*adult-like*, *non-adult-like*, adults) and training duration (single session and multisession). Results showed a significant effect of subgroup, $F(2, 75) = 13.264$, $p < .001$, $\eta^2 = 0.275$, with no significant effect of training duration, $F(1, 75) = 2.866$, $p = .095$, but a significant Subgroup × Training Duration interaction, $F(2, 75) = 3.866$, $p = .051$, $\eta^2 = 0.162$. Post hoc analysis showed that at retention testing, the within-measurement variance of the *non-adult-like* children was larger than that of the *adult-like* and adults but only for the subgroups that underwent a single session of training ($p < .001$). No significant differences between the within-measurement variance was found for the subgroups that underwent multisession training ($p > .05$).

Discussion

The results of the present study support several major findings: (a) More than half of the 7- to 9-year-old children showed similar performance and time course of learning as the adults during both single- and multisession training. (b) The best predicting factor for the outcomes of the multisession training was performance *following* a single session of training. (c) Nonverbal reasoning and working memory abilities were associated with DLF performance following training. (d) Six to eight months after the cessation of training, the children retained their learning-gains, with further improvement for the *non-adult-like* children.

What Allows a Child to Become an Adult-Like Performer Following Training?

Several mechanisms may be needed for a child to reach *adult-like* performance following the DLF training.

These include cognitive abilities, bottom-up statistical learning processes, efficient sensory coding of the trained stimuli, and consolidation processes.

To form an optimal task solution strategy, cognitive abilities such as attention and memory are needed to be engaged (Ahissar & Hochstein, 1997; Jones et al., 2015; Vakil et al., 2014). Mature attention mechanism, for example, is necessary in order to stay focused on the task for 3 to 5 min at a time (the duration of a DLF measurement) through each of the six threshold measurements. Lack of attention can result in inconsistent responses, large within-measurement variance, and subsequently, poor DLF performance (Moore, Ferguson, Halliday, & Riley, 2008). This can explain, at least in part, the poor performance of the *non-adult-like* subgroup of children in the present study. Working memory abilities may have also contributed to the learning of the DLF task. To perform the trained task, the children had to store consistently three sounds in memory long enough to reach a decision regarding which stimulus had the different pitch. Immature memory for pitch was recently proposed to be the main cause for poor frequency discrimination in children when tested using AXB procedure (Buss, Taylor, & Leibold, 2014). Immature attention or memory may have caused, therefore, failure in producing consistent perceptual anchoring with repeated exposure to the task, impeding the setting of an effective task solution routine for the DLF task (Banai & Yifat, 2011). Our findings that the *adult-like* children showed better nonverbal reasoning and working memory than the *non-adult-like* children support this explanation.

Statistical learning processes may also be important for reaching *adult-like* performance following training. These processes depend on efficient bottom-up sensory processing of the stimuli to produce consistent internal (neural) feedback, which updates *synaptic weights* at the sensory level (e.g., Janacsek et al., 2012). While both *adult-like* and *non-adult-like* children showed similar hearing sensitivity that was within the normative range, it is possible that they differed in specific bottom-up stimuli coding mechanism they used for processing the DLF task. It was previously suggested that in order to discriminate between high frequencies (above approximately 1500 Hz), an adult listener extracts and compares the patterns of excitation evoked successively in the basilar membrane (excitation-pattern model; Moore, 2003). For lower frequencies (<1500 Hz), the listener compares the coding of time intervals between peaks in the fine structure of the stimulus waveform (neural phase-locking model; Moore, 2003). To date, most studies (with one exception: Buss et al., 2014) have shown children to reach mature (*adult-like*) DLF thresholds at high frequencies before they reach mature DLF thresholds at low frequencies (e.g., Jensen & Neff, 1993; Maxon & Hochberg, 1982; Thompson, Cranford, & Hoyer, 1999) supporting the notion that place coding matures before temporal coding. Thus, it is possible that in the present study, children who failed to reach *adult-like* performance relied mainly on place coding, which is considered less beneficial for 1000 Hz stimuli (Moore, 2003). Future studies training children with high frequency stimuli (2000 or 4000 Hz) may provide further insight to this hypothesis.

Efficient sensory processing of the trained task may also require low levels of internal noise (Buss, Hall, & Grose, 2006, 2009). Internal noise is defined as uncertainty in the internal response to a sensory input which is generated by sources intrinsic to the observer, such as stochastic neural encoding and transmission, or physiological maskers such as heartbeats or blood flow (Jones, Moore, Amitay, & Shub, 2013).It is possible, that elevated internal noise levels may have caused poor and inconsistent signal-to-noise ratios for some of the children, making it difficult for them to focus on the pitch differences between the tones. This may have prevented the updating of *synaptic weights* based on the statistical distribution of the stimuli (Hinton & Sejnowski, 1999; McClelland, Thomas, McCandliss, & Fiez, 1999), resulting in large within- and between-measurement variance as well as poor DLF thresholds throughout training. The notion that children may produce inefficient bottom-up, neural feedback following frequency discrimination training is supported by a recent study showing no improvement in thresholds in children who were trained without external feedback (Zaltz et al., 2017).

Finally, efficient consolidation processes may be needed for reaching *adult-like* performance following training. These were reported to include the establishment of the neural representations of the trained stimuli and the completion of structural synaptic changes that were induced by training (e.g., Dudai, 2012; Karni, 1996). In the present study, we found, however, that all the children (good and poor performers) showed improvement in thresholds between the first and second sessions, presumably reflecting efficient consolidation mechanisms.

The present findings are in keeping with those studies that showed developed auditory skill learning in some school-age children (Halliday et al., 2008; Zaltz et al., 2017). They are different, however, from those that failed to show developed auditory skill learning in adolescents (Huyck & Wright, 2011, 2013). This difference may stem from the different tasks used for training. In Huyck and Wright's (2013) study, for example, training was conducted using a backward masking task. This task is thought to involve central auditory temporal processes (the *temporal windows* theory: Moore, Glasberg, Plack, & Biswas, 1998) that continue to develop into the second decade of life (e.g., Ari-even Roth, Kishon-Rabin, &

Hildesheimer, 2002; Hartley et al., 2000; Moore, 2003). In contrast, our study used a frequency discrimination task that elicited similar thresholds to those of adults in some of the children (e.g., Halliday et al., 2008; Zaltz et al., 2017). It is possible that this task allowed for better use of bottom-up processing for statistical learning, resulting in improved outcomes of training in children. It is also possible that the difference between our findings and Huyck and Wright's findings is related to the training protocols. Although the present training task did not include a computer software that was devised specifically for children (Halliday et al., 2008 and Moore et al., 2008), it included a near-by tester who set behind the participants (both children and adults) and provided verbal reinforcements at the end of each DLF measurement. Such reinforcement may have helped the children to maintain high motivation and attention throughout the training sessions. This explanation is in favor of the notion that motivation and attention highly influence the outcomes of training in children (e.g., Amitay, Halliday, Taylor, Sohoglu, & Moore, 2010; Halliday et al., 2008; Moore et al., 2005, 2008).

A Single Training Session Predicts Outcomes of Multisession Training

The outcomes of the multisession training were predicted from the performance following a single session of training. Specifically, those children who performed *like adults* and were grouped with the adults following the statistical analysis remained *adult-like* performers throughout the training sessions, whereas those children who were grouped as *non-adult-like* remained so after multiple training sessions. Furthermore, only one of the eight *non-adult-like* children ended his multisession training within 1.5 *SD* of the adult's mean starting performance (following the first training session). These findings are in keeping with recent findings from the motor modality, showing that only children who succeeded in a mirror-drawing task during the first training session improved their performance following subsequent training (Julius & Adi-Japha, 2016). The maturity of the aforementioned mechanisms, including task-relevant cognitive abilities and sensory processing may further explain why additional training, beyond a single session did not change the basic ability of the children to cope with the trained task, nor did it change their learning characteristics.

Retention of Performance

The ability of both the *adult-like* and *non-adult-like* children to retain their learning-gains similarly to the adults may provide further support to the notion that all the children had effective consolidation into long-term memory of the training-induced (perhaps structural) neural modifications (e.g., Dorfberger, Adi-Japha, & Karni, 2012; Karni & Sagi, 1993; Meulemans, Van der Linden, & Perruchet, 1998; Moore et al., 2005). This explanation supports the suggestion that some procedural learning processes are already mature in children as young as 7 years of age (e.g., Perez, Peynircioglu, & Blaxton, 1998; Thomas et al., 2004). Similar findings have been reported for other perceptual and motor tasks in children following single- or multisession training (e.g., Dorfberger et al., 2012; Karni & Sagi, 1993; Meulemans et al., 1998).

The finding that the *non-adult-like* children in the present study not only retained their performance but also improved it over time may reflect a general maturation process that occurred in either auditory or cognitive task-related processes that were not necessarily related to the training experience per se. Support for this notion can be found in the control group of children who improved between two testing sessions, though no training was provided.

Nevertheless, the improvement for the *non-adult-like* children was larger than that of the control group over the months when no training was given. While it is possible that this was the result of the control group including both *adult-like* and *non-adult-like* children, it is also possible that the larger improvement of the *non-adult-like* trained groups reflected a prolonged consolidation process following their inefficient learning during the training phase. That is, these children may have needed more time to reach their full potential for frequency discrimination following training. For a few *non-adult-like* children, a period of 6 to 8 months was enough to *close the gap* and reach *adult-like* performance. For others, improvements during this period were smaller, and they failed to reach *adult-like* performance, suggesting that additional time may have been needed for them to reach maturation or complete consolidation processes. This explanation is supported by our finding that the *non-adult-like* children who underwent multisession training reduced their within-measurement variance to similar values as the *adult-like* children and adults at the retention session. It is also in accordance with a previous study from the motor modality showing that slower establishment of effective representations of a trained task in memory may lead to latent memory consolidation processes (i.e., beyond a 24-h interval; Adi-Japha, Fox, & Karni, 2011). This hypothesis can be tested in future studies where retention will be assessed at a greater interval posttraining. This will allow determining whether *non-adult-like* children can reach *adult-like* performance if longer consolidation time is provided.

Limitations of the Study

One limitation of the present study is that training was performed using only one psychoacoustic nonlinguistic task. It is possible that different auditory tasks will result in different learning characteristics, thus limiting the generalization of the outcomes to other auditory tasks. A second limitation is that a small range of children's ages all from the same socioeconomic status (SES) were tested. While this may be considered a strength of the study because it allowed controlling for confounding factors of age and SES, generalization of the outcomes to other ages from different SES needs to be assessed.

Practical Implications

The present results may have practical implications for the design of training programs for children. It is suggested that obtaining optimal outcomes from auditory training in children is not age dependent per se but rather depends on the maturity of the relevant underlying mechanisms, which can be recognized and determined following a single training session. Therefore, training may best be tailored for each child individually, depending on his or her maturation of task-related sensory processing and general cognitive abilities. Future training studies should test this model in various auditory tasks with children of different age groups.

Acknowledgments

The authors acknowledge the contribution of Yehuda Ben-Simon for the designing of the DLF testing and training software. Special thanks to Ilan Roziner and Haya Grinvald for the statistical analysis and to all the adults and children who participated in the present study.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The authors wish to thank Yairi grant (Communications Disorder Department, Tel-Aviv University) and Steyer grant (School of Health Professions, Tel-Aviv University) for their financial support.

Note

1. Note that 42 children were initially recruited for the study, but 2 were excluded from the analyses. One child was unable to discriminate between stimuli that differed at the maximum range of testing (200 Hz). The second child, who was assigned to the multisession training group, showed a decline in motivation with training and refused to complete the nine training sessions.

ORCID iD

Y Zaltz http://orcid.org/0000-0003-0927-0528

References

Adi-Japha, E., Fox, O., & Karni, A. (2011). Atypical acquisition and atypical expression of memory consolidation gains in a motor skill in young female adults with ADHD. *Research in Developmental Disabilities*, *32*, 1011–1020. doi: 10.1016/j.ridd.2011.01.048.

Ahissar, M., & Hochstein, S. (1997). Task difficulty and the specificity of perceptual learning. *Nature*, *387*, 401–406.

Amitay, S., Halliday, L., Taylor, J. L., Sohoglu, E., & Moore, D. R. (2010). Motivation and intelligence drive auditory perceptual learning. *PLoS One*, *5*, e9816. doi: 10.1371/journal.pone.0009816.

Acoustical Society of America. (1996). *American national standard specification for audiometers* (ANSI, S3.6-1996). New York, NY: Author.

Ari-Even Roth, D., Kishon-Rabin, L., & Hildesheimer, M. (2002). Auditory backward masking in normal-hearing children. *Journal of Basic Clinical Physiology and Pharmacology*, *13*(2), 105–116.

Ari-Even Roth, D., Kishon-Rabin, L., Hildesheimer, M., & Karni, A. (2005). A latent consolidation phase in auditory identification learning: Time in the awake state is sufficient. *Learning & Memory*, *12*(2), 159–164.

Banai, K., & Yifat, R. (2011). Perceptual anchoring in preschool children: Not adult-like, but there. *PLoS One*, *6*(5), e19769. doi:10.1371/journal.pone.0019769

Boothroyd, A. (1997). Auditory development of the hearing child. *Scandinavian Audiology*, *46*, 9–16.

Boothroyd, A., & Boothroyd-Turner, D. (2002). Post implantation audition and educational attainment in children with prelingually acquired profound deafness. *Annals of Otology, Rhinology & Laryngology*, *189*, 79–84.

Brashers-Krug, T., Shadmehr, R., & Bizzi, E. (1996). Consolidation in human motor memory. *Nature*, *382*, 252–255.

Buss, E., Hall, J. W. III., & Grose, J. H. (2006). Development and the role of internal noise in detection and discrimination thresholds with narrow band stimuli. *Journal of the Acoustical Society of America*, *120*, 2777–2787.

Buss, E., Hall, J. W. III., & Grose, J. H. (2009). Psychometric functions for pure tone intensity discrimination: Slope differences in school aged children and adults. *Journal of the Acoustical Society of America*, *125*, 1050–1058.

Buss, E., Shuman, H., Grose, J. H., & Hall, J. W. III. (2013). The monaural temporal window based on masking period pattern data in school-aged children and adults. *Journal of the Acoustical Society of America*, *133*, 1586–1597. doi: 10.1121/1.4788983.

Buss, E., Taylor, C. N., & Leibold, L. J. (2014). Factors affecting sensitivity to frequency change in school-age children and adults. *Journal of Speech Language and Hearing Research*, *57*, 1972–1982. doi: 10.1044/2014_JSLHR-H-13-0254.

Dorfberger, S., Adi-Japha, E., & Karni, A. (2012). Sequence specific motor performance gains after memory consolidation in children and adolescents. *PLoS One*, *7*(1), e28673. doi:10.1371/journal.pone.0028673

Dudai, Y. (2012). The restless engram: Consolidations never end. *Annual Review of Neuroscience*, *35*, 227–247. doi: 10.1146/annurev-neuro-062111-150500.

Edwards, V. T., Giaschi, D. E., & Low, P. (2005). Sensory and nonsensory influences on children's performance of dichotic pitch perception tasks. *Journal of the Acoustical Society of America*, *117*, 3157–3164.

Flavell, J. H., Miller, P. H., & Miller, S. A. (1993). *Cognitive development* (3rd ed.). Englewood Cliffs, NJ: Prentice-Hall.

Halliday, L. F., Taylor, J. L., Edmondson-Jones, M., & Moore, D. R. (2008). Frequency discrimination learning in children. *Journal of the Acoustical Society of America*, *123*, 4393–4402. doi: 10.1121/1.2890749.

Halliday, L. F., Taylor, J. L., Millward, K. E., & Moore, D. R. (2012). Lack of generalization of auditory learning in typically developing children. *Journal of Speech, Language, and Hearing Research*, *55*, 168–181. doi: 10.1044/1092-4388(2011/09-0213).

Hartley, D. E., Wright, B. A., Hogan, S. C., & Moore, D. R. (2000). Age-related improvements in auditory backward and simultaneous masking in 6- to 10-year-old children. *Journal of Speech Language and Hearing Research*, *43*(6), 1402–1415.

Hauptmann, B., & Karni, A. (2002). From primed to learn: The saturation of repetition priming and the induction of long-term memory. *Brain Research. Cognitive Brain Research*, *13*, 313–322.

Hauptmann, B., Reinhart, E., Brandt, S. A., & Karni, A. (2005). The predictive value of the leveling off within-session performance for procedural memory consolidation. *Brain Research. Cognitive Brain Research*, *24*, 181–189.

Hinton, G. E., & Sejnowski, T. J. (1999). *Unsupervised learning: Foundations of neural computation*. Cambridge, MA: MIT Press.

Huyck, J. J., & Wright, B. A. (2011). Late maturation of auditory perceptual learning. *Developmental Science*, *14*, 614–621. doi: 10.1111/j.1467-7687.2010.01009.x.

Huyck, J. J., & Wright, B. A. (2013). Learning, worsening, and generalization in response to auditory perceptual training during adolescence. *Journal of the Acoustical Society of America*, *134*, 1172–1182. doi: 10.1121/1.4812258.

Janacsek, K., Fiser, J., & Nemeth, D. (2012). The best time to acquire new skills: Age-related differences in implicit sequence learning across the human lifespan. *Developmental Science*, *15*(4), 496–505. doi:10.1111/j.1467-7687

Jensen, J. K., & Neff, D. L. (1993). Development of basic auditory discrimination in preschool children. *Psychological Science*, *4*, 104–107.

Jones, P. R., Moore, D. R., Amitay, S., & Shub, D. E. (2013). Reduction of internal noise in auditory perceptual learning. *Journal of the Acoustical Society of America*, *133*, 970–981. doi: 10.1121/1.4773864.

Jones, P. R., Moore, D. R., Shub, D. E., & Amitay, S. (2015). The role of response bias in perceptual learning. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *41*(5), 1456–1470. doi:10.1037/xlm0000111

Jung, T., & Wickrama, K. A. S. (2008). An introduction to latent class growth analysis and growth mixture modeling. *Social and Personality Psychology Compass*, *2*(1), 302–317.

Julius, M. S., & Adi-Japha, E. (2016). A developmental perspective in learning the mirror-drawing task. *Frontiers in Human Neuroscience*, *10*, 83. doi:10.3389/fnhum.2016.00083

Karni, A. (1996). The acquisition of perceptual and motor skills: A memory system in the adult human cortex. *Cognitive Brain Research*, *5*, 39–48.

Karni, A., Meyer, G., Rey-Hipolito, C., Jezzard, P., Adams, M. M., Turner, R., & Ungerlieder, L. G. (1998). The acquisition of skilled motor performance: Fast and slow experience-driven changes in primary motor cortex. *Proceedings of the National Academy of Science USA*, *95*, 861–868.

Karni, A., & Sagi, D. (1993). The time course of learning a visual skill. *Nature*, *365*, 250–252.

Kishon-Rabin, L., & Boothroyd, A. (2018). The role of hearing in speech and language acquisition and processing. In Ravid D and Baron A (Eds), *Handbook of communication disorders: Theoretical, empirical, and applied linguistic perspectives* (pp. 19–41). Berlin: Mouton de gruyter, Inc..

Kishon-Rabin, L., Segal, O., & Algom, D. (2009). Associations and dissociations between psychoacoustic abilities and speech perception in adolescents with severe-to-profound hearing loss. *Journal of Speech Language and Hearing Research*, *52*, 956–972. doi: 10.1044/1092-4388(2008/07-0072).

Levitt, H. (1971). Transformed up-down methods in psychoacoustics. *Journal of the Acoustical Society of America*, *49*, 467–477.

Luna, B., Garver, K. E., Urban, T. A., Lazar, N. A., & Sweeney, J. A. (2004). Maturation of cognitive processes from late childhood to adulthood. *Child Development*, *75*, 1357–1372.

Maxon, A. B., & Hochberg, I. (1982). Development of psychoacoustic behavior: Sensitivity and discrimination. *Ear and Hearing*, *3*, 301–308.

McArdle, J. J. (2012). Latent curve modeling of longitudinal growth data. In R. H. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 547–571). New York, NY: The Guilford Press.

McClelland, J. L., Thomas, A., McCandliss, B. D., & Fiez, J. A. (1999). Understanding failures of learning: Hebbian learning, competition for representational space, and some preliminary experimental data. In J. Reggia, E. Ruppin, & D. Glanzman (Eds.), *Brain, behavioral, and cognitive disorders: The neurocomputational perspective* (pp. 75–80). Oxford, UK: Elsevier.

Meulemans, T., Van der Linden, M., & Perruchet, P. (1998). Implicit sequence learning in children. *Journal of Experimental Child Psychology*, *69*, 199–221.

Millward, K. E., Hall, R. L., Ferguson, M. A., & Moore, D. R. (2011). Training speech-in-noise perception in mainstream school children. *International Journal of Pediatric Otorhinolaryngology*, *75*, 1408–1417. doi: 10.1016/j.ijporl.2011.08.003.

Moore, B. C. J. (2003). *An introduction to the psychology of hearing* (5th ed.). London, England: Academic press Inc.

Moore, B. C. J., Glasberg, B. R., Plack, C. J., & Biswas, A. K. (1998). The shape of the ear's temporal window. *The Journal of the Acoustical Society of America*, *83*, 1102–1116.

Moore, D. R., Ferguson, M. A., Halliday, L. F., & Riley, A. (2008). Frequency discrimination in children: Perception, learning and attention. *Hearing Research*, *238*, 147–154. doi: 10.1016/j.heares.2007.11.013.

Moore, D. R., Rosenberg, J. F., & Coleman, J. S. (2005). Discrimination training of phonemic contrasts enhances phonological processing in mainstream schoolchildren. *Brain and Language*, *94*, 72–85.

Muller, J. L., Friederici, A. D., & Mannel, C. (2012). Auditory perception at the root of language learning. *Proceedings of the National Academy of Science USA*, *109*, 15953–15958.

Perez, L. A., Peynircioglu, Z. F., & Blaxton, T. A. (1998). Developmental differences in implicit and explicit memory performance. *Journal of Experimental Child Psychology*, *70*, 167–185.

Raven, J. C., & Court, J. H. (1998). *Raven manual, Section 1. Standard Progressive Matrices*. Oxford, England: Oxford Psychologist Press Ltd.

Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1999). Statistical learning of tone sequences by human infants and adults. *Cognition*, *70*(1), 27–52.

Soderquist, D. R., & Moore, M. J. (1970). Effect of training on frequency discrimination in primary school children. *Journal of Audiology Research*, *10*, 185–192.

Thomas, K. M., Hunt, R. H., Vizueta, N., Sommer, T., Durston, S., Yang, Y., & Wordern, M. S. (2004). Evidence of developmental differences in implicit sequence learning: An fMRI study of children and adults. *Journal of Cognitive Neuroscience*, *16*, 1339–1351.

Thompson, N. C., Cranford, J. L., & Hoyer, E. (1999). Brief-tone frequency discrimination by children. *Journal of Speech Language and Hearing Research*, *42*, 1061–1068.

Toga, A. W., Thompson, P. M., & Sowell, E. R. (2006). Mapping brain maturation. *Trends in Neuroscience*, *29*, 148–159.

Tomblin, B. J., & Quinn, M. A. (1983). The contribution of perceptual learning to performance on the repetition task. *Journal of Speech and Hearing Research*, *26*, 369–372.

Vakil, E., Hassin-Baer, S., & Karni, A. (2014). A deficit in optimizing task solution but robust and well-retained speed and accuracy gains in complex skill acquisition in Parkinson's disease: Multi-session training on the Tower of Hanoi Puzzle. *Neuropsychologia*, *57*, 12–19. doi:10.1016/j.neuropsychologia.2014.02.005

Wechsler, D. (1991). *Wechsler Intelligence Scale for Children—III*. San Antonio, TX: The Psychological Corporation.

Zaltz, Y., Ari-Even Roth, D., & Kishon-Rabin, L. (2011). How specific is the learning in an auditory frequency discrimination task? *Journal of Basic Clinical Physiology and Pharmacology*, *22*(3), 69–73. doi:10.1515/jbcpp.2011.013

Zaltz, Y., Ari-Even Roth, D., & Kishon-Rabin, L. (2017). Is the role of external feedback in auditory skill learning age dependent? *Journal of Speech Language and Hearing Research*, *60*(12), 3656–3666. doi:10.1044/2017_JSLHR-H-16-0446