Ear and Hearing

Auditory Training Effects on the Listening Skills of Children with Auditory Processing Disorder

--Manuscript Draft--

Manuscript Number:	EANDH-D-14-00310R1
Full Title:	Auditory Training Effects on the Listening Skills of Children with Auditory Processing Disorder
Article Type:	Research Article
Section/Category:	Auditory Neuroscience: Central
Keywords:	auditory training; auditory processing disorder; Speech in noise; listening skills
Corresponding Author:	Doris-Eva Bamiou, M.D., M.Sc., Ph.D. University College London (UCL) London, UNITED KINGDOM
Corresponding Author Secondary Information:	
Corresponding Author's Institution:	University College London (UCL)
Corresponding Author's Secondary Institution:	
First Author:	Jenny Loo
First Author Secondary Information:	
Order of Authors:	Jenny Loo
	Stuart Rosen
	Doris-Eva Bamiou, M.D., M.Sc., Ph.D.
Order of Authors Secondary Information:	
Abstract:	Objectives: Children with Auditory Processing Disorder (APD) typically present with 'listening difficulties', including problems understanding speech in noisy environments. We examined, in a group of such children, whether a 12-week computer-based auditory training programme with speech material improved the perception of speech-in-noise test performance, and functional listening skills as assessed by parental and teacher listening and communication questionnaires. We hypothesised that, after the intervention: 1) trained children would show greater improvements in speech-in-noise perception than untrained controls; 2) this improvement would correlate with improvements in observer-rated behaviours; 3) the improvement would be maintained for at least 3 months after the end of training. Design: This was a prospective randomised controlled trial of 39 children with normal nonverbal intelligence, aged 7 to 11 years, all diagnosed with APD. This diagnosis required a normal pure tone audiogram and deficits in at least two clinical auditory processing tests. The APD children were randomly assigned to: A. a control group who received only the current standard treatment for children diagnosed with APD, employing various listening/educational strategies at school (N=19), or to; B. an intervention group, who undertook a 3-month 5-days/week computer-based auditory training programme at home, consisting of a wide variety of speech-based listening tasks with competing sounds, in addition to the current standard treatment. All 39 children were assessed for language and cognitive skills at baseline and on 3 outcome measures at baseline and immediate post-intervention. Outcome measures were repeated 3 months post-intervention in the intervention group only, to assess the sustainability of treatment effects. The outcome measures were: 1) the mean speech reception threshold obtained from the 4 subtests of the Listening in Spatialised Noise (LiSN) test, that assesses sentence perception in various configurations of masking speech, and in which the target speakers and test materials

	were unrelated to the training materials; 2) the Children's Auditory Performance Scale (CHAPS) that assesses listening skills, completed by the children's teachers; 3) The CELF-4 Pragmatic Profile (PP) that assesses pragmatic language use, completed by parents. Results: All outcome measures significantly improved at immediate post-intervention in the intervention group only, with effect sizes ranging from 0.76 - 1.7. Improvements in speech-in-noise performance correlated with improved scores in the CHAPS questionnaire in the trained group only. Baseline language and cognitive assessments did not predict better training outcome. Improvements in speech-in-noise performance were sustained 3 months post-intervention. Conclusions: Broad speech-based auditory training led to improved auditory processing skills as reflected in speech-in-noise test performance and in better functional listening in real life. The observed correlation between improved functional listening with improved speech-in-noise perception in the trained group suggests that improved listening was a direct generalisation of the auditory training.

To The Editors

Ear Hearing

28th November 2014

Dear Sirs

We would like to submit our paper: Auditory Training Effects on the Listening Skills of Children with Auditory Processing Disorder , Authors: Jenny Hooi Yin Loo , Stuart Rosen , and Doris-Eva Bamiou to be considered for publication in Ear Hearing.

Notification of Ethical Adherence

Ethical approval was obtained by the National Healthcare Group Singapore. Informed written consent from parents and assent from children were obtained. The Clinical Trial was registered with ClinicalTrials.gov (registration number: NCT02111343).

Statement of Authorship:

Jenny Loo: Dr. Loo helped design the study, conducted testing and auditory training, conducted the initial analysis, drafted the initial manuscript, and approved the final manuscript as submitted.

Stuart Rosen: Dr Rosen designed the study, designed the training intervention, reviewed and finalised the analyses, reviewed and revised the manuscript, and approved the final manuscript as submitted.

Doris-Eva Bamiou: Dr Bamiou conceptualised and designed the study, conducted the initial analysis, critically reviewed the initial draft and finalised the manuscript, and approved the final manuscript as submitted.

With many thanks

On behalf of the authors

Doris-Eva Bamiou, MD, MSc, PhD, FRCP

Thank you for the helpful comments. The manuscript has been revised accordingly. In particular:
Section Editor Comments:

" I share the concern of the reviewers about the study's lack of an active control group."
This has been acknowledged in the discussion.

"I also wonder about some aspects of the results that were contrary to expectation (e.g. 'Other analyses' p. 14), but somewhat underplayed in the Discussion and Abstract**." We have sought to address this, by conducting additional analysis (as per the reviewers' comments) and editing abstract and discussion.**

"Finally, I am concerned especially about Fig. 3. In contrast to the LiSN-S, where most of the change in the AT group was negative, suggesting improved performance, the abundance of CHAPS change scores for the AT group were around zero, with just a few outliers, three of whom got poorer scores following training, defining the significant non-parametric correlation. Yet the most dramatic conclusions of the study (ramped up to very bold statements in the Abstract) appear to rest on this figure."

In fact, 12 of 16 of the trained listeners (4 did not have CHAPS scores) obtained an improvement in the CHAPS measure. Our use of the bootstrap technique was meant to avoid any kind of informal selection and exclusion of data.

"Given this issue, and Reviewer 2's question of the use of box plots, I'd also like to see individual data for the PP Scores in Fig. 2."
Bee swarm boxplots have now been provided for Fig 2

Minor issues:
"* Find another way for naming the test times. The fact that training lasted 3 months and the post-training time was a further 3 months leads to confusion."

We now use the labels 'Baseline', 'Immediate Post-intervention' and '3-months Post-intervention'

"* Fig. 1 A and B should be combined. I notice that the Control group had better thresholds at baseline than the AT group, so a combined figure will provide a somewhat more informative perspective. "

Done.

“* On p.12 there are two t-scores of 4.3 at least one of which is certainly incorrect”

Thank you, this has now been corrected.

“* On p14, "Figure 3 depicts the relationship between these two" Which two? “
The two are the change in mean LiSN-S and the change in mean CHAPS. This has been clarified in the text

“Given that the Control group correlation in Fig. 3 was n.s., why was it combined with the AT for fitting the regression line?”
Correlations in separate groups cannot be compared through significance levels, and we combined the two groups because, as mentioned, ‘separate robust analysis shows the slopes and intercepts of the two groups being indistinguishable’.

Reviewer #1:

Specific comments
“Lines 59 - 73: This valuable summary of desirable characteristics did not include an obvious one: a placebo or sham treatment for the control group, which has been used in at least one study (Cameron & Dillon, 2012). Of course, inclusion of this desirable characteristic does mean that the field is still in the position of no single study incorporating all the characteristics desirable in a study of this type, even including the current study.”
This limitation is fully acknowledged in the Discussion.

“Lines 89 - 98: The children in this study would be much better characterised if we were told how many children failed each of the tests in the battery. This could be accomplished with a simple addition to Table 2. Such characterisation is important given the heterogeneity of children diagnosed with APD, so that the applicability of the findings to other groups can be assessed. For several reasons, it is not possible to infer this information from the ranges given in Table 2.”
Done.

“Line 115: It is more common to have a space between numbers and their units - e.g. 250 Hz.”
Done

“Lines 161-162: Does the statement about a standard score below 85 indicating "marked difficulties" have any relevance to this study? Each of four tests in this section (TONI, CELF, TAPS-R and PhAB) conclude with a statement about what is considered normal and abnormal. These sentences are all expressed differently, they all refer to arbitrary cut-offs, and for three of them, a cut-off of one SD (i.e. a standard score of 85 or percentile of 16) is selected, and for the fourth one a cut-off of two SD is selected. Why these different criteria about what is normal, and do any of them have any relevance to the study? The inclusion criteria given in lines 92 to 97 mention only non-verbal IQ (presumably based on the TONI scores) as an inclusion criterion, and the cut-off score of 85 was given in that section. It would be less confusing, if the descriptions just made it clear that the results of each are expressed as standard scores (i.e. with a population mean of 100 and SD of 15), because the actual range of results (or distribution statistics) for each test are given in Table 2."

We have made it clear that the results of each are expressed as standard scores, and provided numbers of children failing each test in table 2. We kept the criterion for language impairment (LI) that we provided since this is the criterion used in Singapore.

"Line 178: Reads more easily if "is" is inserted before "adaptively"."
Done.

"Lines 213 - 224: The training games were intentionally varied, and some may have been more effective than others at achieving the outcomes reported in this paper, but we have no way of even knowing which games were actually used, let alone their individual effectiveness. If the children were free to choose which game was used on any occasion, then does the data log referred to indicate the total training time per game, or number of occasions each game was used? If so, this could usefully be indicated."
The children were not free to choose which game was used. As mentioned in the text 'A daily AT timetable was issued for 12 weeks with two different listening games to perform for 30 minutes per session, 5 sessions per week.' We have added a table (table 4) with an indicative weekly schedule.

"Line 223: Unclear who (parents or clinician) crosschecked the training logbook with the datalog, when this occurred, what the purpose was, and what happened if they differed. No need to necessarily add lots of details as it's not critical to the paper; just make it clear what this sentence means."
The training logbook and datalog were crosschecked by Dr Jenny Loo at the end of the study. This was done in order to establish the amount of training each child received.

"Line 240: An earlier version of this paper that I (positively) reviewed reported 4 or 5 times greater training hours than this. I presume the earlier version was in error and this one is correct. "
Yes, we erroneously reported number of sessions (of 15 min each) as number of hours completed by each child in the previous report. This has now been corrected.

"Lines 250 to 253: There is of course no contradiction in having an insignificant 3rd order interaction despite the second order interaction (time by condition) being significant for the trained group but not for the untrained group. The lack of the 3-way interaction means that one cannot conclude that the two 2-way interactions are in any way different, despite one being significant and one not. I would therefore reverse the order of the two parts of the sentence to make it clearer how this finding should be interpreted.”

We have excised the mention of the different significance values in the two groups, and simply talked about the effect being numerically larger in the AT group.

“Lines 253 - 256: This is not an accurate summary, as the improvement for the SV0 condition is just as large as for the DV90 condition. Suggest deletion.”
As the following table shows, our statement is correct.

Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation
DV90change	20	-6.20	.60	-2.1200	1.66658
DV0change	20	-4.20	.80	-.9850	1.42949
SV0change	20	-2.90	1.00	-1.2650	.99909
SV90change	20	-3.90	.90	-1.5200	1.53712
Valid N (listwise)	20				

“Line 310: The logic does not seem right here. I think the authors intend the "presumably because" statement to be an explanation of why the ANOVA had a significant time by group interaction, rather than an explanation of why the t-tests were not significant. Rearrange the para.”
Yes. Thank you and corrected.

“Lines 336 - 339 and lines 379-380: On several counts I don't think it is possible to draw any conclusions from the lack of correlation between hours of training and change in each of LiSN-S, CHAPS and PP. First, the variance of 13% mentioned corresponds to a correlation coefficient of 0.36. With an n of only 15 an r-value of 0.36 has a 95% confidence interval from -0.20 to 0.74. Values near the top of this range would indicate an extremely strong relationship between the measures and so cannot be ruled out by this data. Second, the hours of training had an upper limit of 30, but a median of 27, so for half the group there was almost no variation in training duration. This restricted range severely limits the ability to detect a correlation in this data, and the correlation observed was probably determined almost totally by those few children with much less than median training hours. Third, if one includes the control group, there is obviously an intrinsic effect of training time present, as one group had zero training time, and the two groups had different results. For these reasons, the conclusion (lines 379-380) that training had produced an asymptotic degree of benefit is unsupportable from these data and should be deleted.”

Agreed. This statement has been deleted, and we have noted that: ‘As more than half the number of participants from whom data was available completed all the specified training sessions, we did not try to relate total hours of AT to changes in any of the outcome measures..

“Line 373: typo "may is".”
Done.

“Lines 342 - 343: This positive conclusion, which is justified, should include the words "on average" or "nearly all", as it is clear from Figure 3 that the training had no beneficial effect (as assessed from CHAPS and LiSN-S) for two out of 16 children.”
Done.

“Lines 354 - 380: Another difference between experiments that has not been considered here, which is probably the most important, is that the median 27 hours of training in this experiment was distributed across four tasks, only one of which (the CCRM task) used a speech target spatially separated from the noise. This study consequently had much shorter training time using spatialized sounds than was used in Cameron & Dillon (2011), Cameron et al (2012) and Cameron et al (2014), that doubtless contributed to the lack of improvement in spatial advantage found here.”
We added this sentence: Training with a speech target that was spatially separated from the noise was only done during one of the four AT games the children played, thus reducing total training time for this task.

“Lines 373 - 375: If "SPD" here is actually intended to mean "improvement in spatial advantage" then perhaps I understand this sentence, but if not I have no idea what it means. “
Changed to read:

‘In contrast to the findings reported by Cameron and Dillon (2011), Cameron et al. (2012) and Cameron et al. (2014), performance in the four LiSN-S conditions showed comparable improvements, rather than a specific improvement in conditions with spatial separation.’

“Lines 399 - 402: This sentence is confusing. I think the intended meaning of the first part is that auditory training may actually improve general cognitive skills and that these improved cognitive skills will improve scores on language measures. What exactly is being hypothesized to be affected by the "combined use of intensive auditory and visual stimulation" in the second part of the sentence? Best to break this sentence into separate topics and explain each in its own (re-written) sentence. “

This has been rephrased.

“Line 411: Delete "other" as the children in this study were not diagnosed with SPD.”

Done.

“Lines 412 - 415: This is too imprecisely expressed. Do "AT outcomes" refer to changes in scores as a result of training or the absolute scores after training is completed? What types of scores are meant: language, parental/teacher report, or auditory processing ability? This paper reports the correlations between changes in various scores and the baseline language scores. I doubt that this is what Watson et al (2003) reported.”

Modified to read ‘We also found that neither baseline language nor cognitive abilities predict the degree of improvement with AT, in agreement with other studies (Sharma et al. 2012).’

“Line 433: Period missing after "English)".”
Inserted.

Reviewer #2: This is in general a well written manuscript describing a well designed study of auditory training (AT) in children with APD. However, two major conceptual issues, which have strong implications on the evaluation and interpretation of the findings were not addressed. I also find the Results section (both text and graphics) confusing and hard to follow (see below).

Major conceptual issues:
“First, although from a clinical stand point documenting the outcomes of auditory training programs for children with APD is important, the manuscript is sparse on the theoretical rationale for training. Although it is stated that training is expected to be helpful due to "neural plasticity", no model of APD is provided to justify the specifics of the training program and the outcome measures. For example, why is it relevant to train on speech in noise? Is it expected to ameliorate a deficit specific children with APD? Alternatively, since aspects of speech in noise performance are immature in even typically developing children at this age range, perhaps training is targeted at speeding up a general process, and is thus beneficial for APD as well as for typically developing children which were not targeted here. This point is especially troublesome because no normative data is reported in the manuscript (from this or from earlier studies), and thus one can not be sure of the relative standing of the APD participants that formed the current study group. On the other hand, no support is provided for the claim that speech in noise deficits are paramount in APD. In fact, it is not clear whether the studies cited for this point (L48) actually support it. Ferguson

et al., 2011 state that "Speech intelligibility in both noise and quiet was UNIMPAIRED in the SLI and APD groups". Although difficulties in noise were part of the APD inclusion criteria in the Iliadou & Bamiou 2012 study, the lower scores of their APD group on the CHAPPS were not specific to the noise subscale, whereas the non-APD clinical group also scored lower on the CHAPPS noise subscale. **“**

These are valid comments. We have modified the introduction substantially to take account of the comments made here. We have summarised some key issues currently debated regarding APD, and provided a rationale for AT and for speech based AT. The references for the “speech in noise” claim have been changed.

“Second, the size of the treatment effect and performance of the trained and untrained APD groups re- typically developing children are hard to evaluate given the lack of information on comparative data for the three outcome measures. So yes, a mean improvement of 1 dB on the LiSN test is significant and the reported effect size is quite large, but what does it mean in terms of the performance of trained APD children relative to the "normal range"? There are no normative values in Singapore for LiSN-S. “

There are no LiSN-S norms as yet in Singapore, and we cannot address this comment.

Results and Figures:
“1. Figure 1 makes it hard to compare the trained and untrained groups. It would have been easier to see the potential effects of the intervention program if trained and untrained groups were presented on the same panel (with perhaps different panels for the different LiSN subtests).”

Figure 1 has been improved (we hope!) through being redrawn.

“2. Figure 1 (and the discrepancies between the medians shown on Figure 1 and the total average reported in Table 4) also makes it hard to determine whether data meets the requirements of the ANOVAs that form the major statistical analysis in the results section and none of the supporting data is reported (e.g., sphericity, homogeneity of variance, approximation to normal distribution etc').”

Hyunh-Feldt corrections are now applied in all the repeated measures ANOVAs when necessary.

“3. The presentation of the 2x2x2 ANOVA on the LiSN SRTs is very confusing. Why start with the higher order interactions before establishing a significant group x time interaction which is most critical in terms of the goals of the paper? “

We prefer this approach because lower-order interactions are not readily interpretable in the presence of higher order ones.

Minor issues:
"1) There are a few typos throughout the manuscript (for example on P1, L5 and on P3, L71), and in general the manuscript could benefit from some language editing."

Many changes have been made throughout the manuscript. We have done our best!

"2) I find the discussion quite lengthy given the paucity of "theory" in the introduction and the extent of the findings."

We have shortened this somewhat but feel that what remains is important.

Short Summary

What is known on the subject: Auditory training improves speech-in-noise test results in children and adults, although the extent of generalisation to real life situations is largely unknown.

What this paper adds: Auditory training improved speech in noise perception and functional listening/communication skills in children with auditory processing disorder (APD). Correlation of improved functional listening to improved speech-in-noise perception suggests that improved listening was a direct generalisation effect of the auditory training.

Auditory Training Effects on the Listening Skills of Children with Auditory Processing Disorder

Jenny Hooi Yin Loo [1, 2], Stuart Rosen [3], and Doris-Eva Bamiou [2, 4]

Correspondence to: Jenny Loo, National University Hospital, 5 Lower Kent Ridge Road, Level 3, ENT – Head & Neck Surgery Clinic, Singapore 119074.

Email address: Jenny_LOO@nuhs.edu.sg

Sources of funding: Internal funding of the Centre for Hearing Intervention and Language Development (CHILD), Singapore. Dr Loo's salary was funded by the Singapore Research Council. The authors have no financial relationships relevant to this article to disclose.

Conflicts of Interest: The authors have no conflicts of interest relevant to this article to disclose.

---

[1] National University Singapore, Department of Otolaryngology, Head & Neck Surgery, Singapore

[2] University College London (UCL) Ear Institute, London, UK

[3] UCL Speech, Hearing, and Phonetic Sciences, London, UK

[4] Neuro-otology Department National Hospital for Neurology Neurosurgery, London, UK

Abstract

Objectives: Children with Auditory Processing Disorder (APD) typically present with 'listening difficulties', including problems understanding speech in noisy environments. We examined, in a group of such children, whether a 12-week computer-based auditory training programme with speech material improved the perception of speech-in-noise test performance, and functional listening skills as assessed by parental and teacher listening and communication questionnaires. We hypothesised that, after the intervention: 1) trained children would show greater improvements in speech-in-noise perception than untrained controls; 2) this improvement would correlate with improvements in observer-rated behaviours; 3) the improvement would be maintained for at least 3 months after the end of training.

Design: This was a prospective randomised controlled trial of 39 children with normal nonverbal intelligence, aged 7 to 11 years, all diagnosed with APD. This diagnosis required a normal pure tone audiogram and deficits in at least two clinical auditory processing tests. The APD children were randomly assigned to:

A. a control group who received only the current standard treatment for children diagnosed with APD, employing various listening/educational strategies at school (N=19), or to;

B. an intervention group, who undertook a 3-month 5-days/week computer-based auditory training programme at home, consisting of a wide variety of speech-based listening tasks with competing sounds, in addition to the current standard treatment.

All 39 children were assessed for language and cognitive skills at baseline and on 3 outcome measures at baseline and immediate post-intervention. Outcome measures were repeated 3 months post-intervention in the intervention group only, to assess the sustainability of treatment effects.

The outcome measures were:

1) the mean speech reception threshold obtained from the 4 subtests of the Listening in Spatialised Noise (LiSN) test, that assesses sentence perception in various configurations of masking speech, and in which the target speakers and test materials were unrelated to the training materials;

2) the Children’s Auditory Performance Scale (CHAPS) that assesses listening skills, completed by the children’s teachers;

3) The CELF-4 Pragmatic Profile (PP) that assesses pragmatic language use, completed by parents.

Results: All outcome measures significantly improved at immediate post-intervention in the intervention group only, with effect sizes ranging from 0.76 – 1.7. Improvements in speech-in-noise performance correlated with improved scores in the CHAPS questionnaire in the trained group only. Baseline language and cognitive assessments did not predict better training outcome. Improvements in speech-in-noise performance were sustained 3 months post-intervention.

Conclusions: Broad speech-based auditory training led to improved auditory processing skills as reflected in speech-in-noise test performance and in better functional listening in real life. The observed correlation between improved functional listening with improved speech-in-noise perception in the trained group suggests that improved listening was a direct generalisation of the auditory training.

Introduction

The nature of Auditory Processing Disorder (APD, H93.25 in ICD-10) remains a matter of intense debate. This is also of crucial clinical importance, since the theoretical framework adopted for APD determines the diagnostic and management process (ASHA 2005; AAA 2010; BSA 2011). The clinical presentation in children is characterized by ‘listening difficulties’ despite normal pure tone thresholds, with a hallmark symptom of excessive difficulty in understanding speech in the presence of background noise (Chermak et al, 2002; Iliadou & Bamiou 2012; Dillon et al., 2012). However, families seek help because of difficulties in language development or educational attainment rather than the speech in noise symptoms (Tomlin, 2014; Heine and Slone 2008; Myklebust 1954).

Much theorising about APD has centred on the importance of low-level sensory deficits, but it is becoming increasingly clear that higher order factors, for example attention and memory, appear to be crucial in accounting for the clinical presentation (Moore et al, 2010). Even AP tasks with a higher degree of complexity (that require feature extraction and categorization) show only weak correlation with language processing after controlling for cognitive factors (Grube et al, 2012). More recently developed AP tests thus attempt to disentangle auditory processing from cognitive or language influences by calculating difference scores between e.g., speech-based measures, in which the degree of difficulty is varied by manipulation of a specific situation (Cameron & Dillon, 2011). However, while the effects of higher order factors may be minimized, they are unlikely to be eliminated. Diagnostic assessment thus requires multidisciplinary input (AAA 2010; BSA 2011).

These new findings have thus led to the notion that APD results from impaired neural function within the central auditory nervous system (ASHA 2005; AAA 2010) and beyond the auditory cortex (Moore, 2013), since higher-level cognitive processing heavily subserves listening skills (Moore, 2013; Ahmmed et al, 2014). It is therefore natural for APD management strategies to attempt to capitalise on the brain's ability for structural and functional reorganisation in response to sensory input across the life span. This brain "plasticity" may involve the activation of inactive neuronal connections and/or the formation of more efficient synaptic connections within the brain (Chermak, Bellis, & Musiek, 2007).

Therefore, one possible avenue of remediation for APD is auditory training (AT), which is to say listening exercises that aim to improve auditory system function (Loo et al. 2010). The desired effect of AT is to achieve successful auditory learning, i.e. a relatively permanent improvement of perception and behaviour (Fahle & Poggio, 2002). AT studies in normal adults indicate that learning is better when the task is difficult enough to tax attention (Amitay et al, 2006). Learning appears to be driven by attention focusing on the specific stimulus dimension that is of relevance to the training task (Halliday et al, 2011). In addition, studies in normal children indicate that learning may not generalise to untrained tasks or stimuli (e.g., Halliday et al, 2012), so it may be important to train with a wide variety of material and situations. Studies of computer based auditory training programmes with a variety of simple and complex auditory tasks report post- training improvements on a range of auditory and non- auditory measures for a broad range of paediatric populations with disorders that overlap APD (Loo et al. 2010). However it remains unknown whether such improvements generalise to real-life listening situations.

Current AT studies are hampered by significant limitations. Studies on normal adults or typically learning children may not be directly applicable to children with developmental disorders, particularly as learning mechanisms may change during maturation (Halliday et al, 2012). Only a handful of studies have assessed auditory training outcomes explicitly in children with APD as defined by appropriate diagnostic criteria (Wilson et al. 2013; Cameron & Dillon 2011; Sharma et al. 2012; Cameron et al. 2012; Cameron et al. 2014). There is also a paucity of studies that used an untrained control group to estimate practice or maturational effects (e.g. Sharma et al. 2012). Assessments of listening behaviours outside the laboratory, such as by means of questionnaires to investigate whether listening in real life improves after training and whether improvements correlate with improved performance in the auditory skills the AT purports to address are similarly underemployed (Cameron & Dillon 2011; Cameron et al. 2012; Cameron et al. 2014). A true treatment effect remains uncertain, as training materials are sometimes too similar to what is employed in outcome measures, e.g. by using the same talker for training and testing (Cameron and Dillon 2011; Cameron et al. 2012; Cameron et al. 2014). Although sustainability of AT benefits after intervention has been assessed in some studies (Gillam et al. 2008; Strehlow et al. 2008), these are rare for specifically APD populations (Cameron & Dillon 2011). Finally, no single study meets all of the above criteria.

Here we examine the effectiveness of a computer-based auditory training intervention for children with APD using a broad range of AT 'games' with ecologically valid speech stimuli, diagnosed as per explicit criteria (AAA 2010; BSA 2011). We expected that the complex nature of the sound stimuli and the task demands would mean that such training would be more likely to generalise to untrained behaviours. AT was aimed at improving speech-in-noise listening performance, because speech-in-noise test deficits reportedly correlate with other listening and communication indices (e.g,, Moore et al, 2010). In order to assess the effects of AT on children's real life behaviours, we compared the changes in speech-in-noise perception and in observer-rated listening/communication behaviours between trained children and untrained controls immediately post-intervention. We hypothesised that, after intervention, children from the AT group would show a greater improvement in speech-in-noise perception than untrained controls and this improvement would correlate with improvements in observer-rated behaviours. We evaluated speech-in-noise performance of the trained group again at 3 months post-intervention to determine whether AT improvements would be sustained for that period. Finally, we examined whether the training outcomes are predictable from baseline assessments of language or cognitive skills.

Materials and Methods

Participants

Fifty-five consecutive cases of newly diagnosed children with APD, who fulfilled the inclusion criteria below, were identified by clinical staff at the Centre for Hearing Intervention and Language Development (CHILD) in Singapore and invited to participate. Inclusion criteria were: (1) being in mainstream school (2) referred for evaluation of listening difficulties, (3) normal peripheral hearing assessment in both ears (see below), (4) failure in both ears (2 SD criterion) in two or more but not all behavioural tests of a test battery used to assess auditory processing (see below and in Table 1) (5) normal nonverbal intelligence quotient (IQ) score of more than 85 (6) absence of autism and (7) absence of frank neurological conditions such as brain tumour or head injury. Thirty-nine cases consented and were enrolled in the study.

Ethics

Ethical approval was obtained by the National Healthcare Group Singapore. Informed written consent from parents and assent from children were obtained. The Clinical Trial was registered with ClinicalTrials.gov (registration number: NCT02111343).

Setting

Children had baseline assessments conducted and were recruited to the study by clinical staff at CHILD between 2009 and 2011. Outcome measures were also conducted at CHILD by author JL, while the intervention was conducted at the participants' homes.

Study Design and Protocol

APD children were semi-randomly assigned in a sequential method for the two sexes to an intervention auditory training (AT) group (n= 20) or an untrained control group (n = 19) by author JL, who was blinded to the children's baseline assessments.

Baseline Test Procedures

Children were referred for APD assessment after an earlier clinical appointment had confirmed normal peripheral hearing sensitivity with: (1) pure tone thresholds of 20 dB HL or better at octave frequencies from 250 Hz to 8 kHz; (2) normal middle ear function with Type-A tympanograms (Jerger 1970); (3) an ipsilateral acoustic reflex present at 1 kHz with a threshold less than 100 dB HL; (4) speech discrimination scores in quiet (NU6 word list) of 80% or better in both ears presented at 50 dB HL. All recruited children had the following assessments for the study purposes, conducted within a 3.5 hour test session with short intervals between tests to avoid fatigue and to reduce the effect of inattention on test performance:

- Auditory processing tests (see Table 1 for test details and versions used). These were selected as per the American Speech Language and Hearing association (ASHA,

2005) and American Academy of Audiology (2010) recommendations and our previous study on this multilingual population (Loo et al. 2012). These included: two temporal sequencing tests – the frequency pattern test (FPT) and duration pattern test (DPT); a temporal resolution task, the random gap detection test (RGDT); a binaural processing task, the masking level difference (MLD); a dichotic speech test, the dichotic digits test (DDT). Tests were administered using an Orbiter 922 clinical audiometer (Madsen Electronics, Canada) with calibrated TDH-39 headphones (Telephonics, Farmingdale NY). All the test materials were presented using a Sony DVD player (Sony, Tokyo, Japan).

- The Test of Nonverbal Intelligence – 3rd Edition (TONI-3; Brown, Sherbenou & Johnsen 1982) assesses participants' cognitive skills in abstract/figural problem solving. The TONI-3 is a norm-referenced, language free measure that can be used in individuals ages 6;0 through 89;11.

- The Clinical Evaluation of Language Fundamentals – Fourth UK Edition (CELF-4 UK; Semel, Wiig & Secord 2006) assesses core language skill. The CELF-4 consists of the following subtests depending on the child's chronological age: Concepts and following directions (5 to 12 years old), Word structure (5 to 8 years old), Recalling sentences (5 to 12 years old), Formulated sentences (5 to 12 years old), Word classes 2 (receptive, expressive, and total) (9 to 12 years old).

- The Test of Auditory Perceptual Skills-Revised (TAPS-R; Gardner 1996) assesses short-term auditory memory. The TAPS-R has four subtests: Auditory Number Forward Memory (ANFM, otherwise known as digit span), Auditory Number

Backward Memory (ANBM, otherwise known as backwards digit span), Auditory Word Memory (AWM) and Auditory Sentence Memory (ASM).

- The Phonological Assessment Battery (PhAB; Frederickson, Frith & Reason 1997) assesses a wide range of phonological skills. Alliteration, Rhyme, Spoonerisms and Non-word Reading subtests were conducted.

The raw scores of TONI-3, CELF-4, TAPS-R and PhAB were converted into standard scores (i.e. with a population mean of 100 and SD of 15), with scores of 85 and below considered as abnormally low. In addition, a child with normal nonverbal intelligence (NVIQ score > 85, based on TONI-3) and a standard score of 70 and below (2SD's below the UK-referenced norm mean) was considered as having language impairment (LI) in this study.

The group results of these baseline assessments are summarised in table 2. Nine of the AT and twelve of the control group failed at least one subtest of TAPS-R. Three of the AT and eight of the control group failed at least 1 subtest of the PhAB. Three children in the AT and four children in the control group would be classified as having a language impairment.

Outcome Measures

Outcome measures included an objective measure of performance as well as two questionnaires related to real-life function skills. These included:

(a) The Listening in Spatialised Noise – Sentence test (LiSN-S) (Cameron & Dillon 2007, 2008) produces a three-dimensional auditory environment under headphones and assesses the ability of children to repeat back simple sentences in the background of two other talkers which can either be the same or different to the target talker. By manipulating the location and vocal quality of the masking talkers (the target is always perceived as straight ahead), four listening conditions are created: different voices at ± 90° azimuth (DV90°), same voice at ± 90° azimuth (SV90°), different voices at 0° azimuth (DV0°), and same voice at 0° azimuth (SV0°). Responses are scored manually by keyword by the assessor on a computer and the stimulus presentation level is adaptively adjusted depending on participant response. A maximum of 30 sentences are presented in each of the four listening conditions. The outcome measure in each condition was the signal-to-noise ratio (SNR) in decibels (dB) necessary for the correct reporting of 50% of the key words in the sentences, known as the *speech reception threshold* (SRT). Lower SRT values indicate better performance. The LiSN-S outcome measures typically involve differences between selected conditions as a way to 'subtract out' the effect of various cognitive skills in test performance, like attention and linguistic closure. The advantage measures represent the benefit in dB gained when either vocal (DV0°), spatial (SV90°), or both vocal and spatial cues (DV90°) are incorporated in the maskers, compared to the baseline (SV0°) condition where fewer cues are present in the maskers (Cameron & Dillon, 2007). Because our interests are primarily in how measures change over time in the same listener, we used the individual SRTs, the overall LiSN-S performance calculated as the average of the 4 LiSN-S conditions, as well as the derived measures of voice, spatial and total advantage for analysis.

(b) Questionnaires of listening/communication skills

(1) The CELF-4 Pragmatic Profile (PP) (Semel et al. 2006) has 52 items and aims to identify verbal and nonverbal pragmatic deficits that may negatively impact on communication skills. Each item is scored from 1 = never to 4 = always, based on the frequency of occurrence of each skill. The PP was completed by parents who, by the nature of the experimental design, could not be blind to whether or not their child had received the intervention.

(2) The Children's Auditory Performance Scale –CHAPS (Smoski et al. 1992) has 36 questions evaluating listening skills in 6 different areas (noise, quiet, ideal, multiple inputs, auditory memory sequencing, and auditory attention span) scored from +1 (less difficulty) to – 5 (cannot do at all). Raters are asked to compare the child with his/her peers. A total score is calculated from the 6 subscore averages. The CHAPS was completed by participants' teachers, who were blinded to intervention status.

Both groups had all outcome measures at baseline and at the end of the training period. The AT group then underwent a no-intervention 3 months phase, after which LiSN-S was repeated.

Intervention

The AT group were issued a 3-month computer based AT programme to conduct at home under parental supervision (see Table 3 for details of the training games). Three different listening games were used for speech-in-noise training, aiming to improve speech understanding, discrimination of fine phonetic detail, and keyword extraction in the presence of various types of background noises. Dichotic speech listening training with directed attention to one ear was incorporated in a fourth game. All games were presented in a child-friendly visual format with visual feedback provided after each response. A daily AT timetable was issued for 12 weeks with two different listening games to perform for 30 minutes per session, 5 sessions per week (see table 4 for an indicative weekly schedule). Children were rewarded upon completion of each training session with a token or fun activity to promote compliance. Parents kept a training logbook that was crosschecked with the datalog stored in the computer at the end of the training (containing the dates and times of training).

The untrained control group received no auditory training. All participants were requested to not engage with any other auditory-based interventions, except from regular school attendance and educational activities. All participants were receiving the standard current treatments for management of APD which, at the time of the study, were employment of listening strategies (such as preferential sitting) and other educational strategies (such as provision of lecture notes or pre-teaching of new concepts/vocabulary) at school and/or at home.

Results

Subject characteristics

There were no significant differences in any baseline measure between the two groups (Table 2). All 39 children in both groups completed the study. Data on the amount of training undergone was incomplete or missing for 5 of the 20 children in the AT group because of technical failures (e.g., having to replace a faulty computer). The remaining 15 children trained for a median of 27 hours (with a range of 9-30 hours). On average, each participant completed more than 80% of the targeted training sessions for each listening game.

Changes in speech-in-noise performance

All SRTs obtained from the LiSN-S are shown in Figure 1. A repeated measures ANOVA was conducted to assess differences in the 4 LiSN-S scores between the two groups before and at the end of the intervention period (2 times x 2 groups x 4 LiSN-S conditions). Hyunh-Feldt epsilon corrections were applied to all F tests involving LiSN-S condition because Mauchley's Test of Sphericity indicated a violation for this factor (p=0.04). The highest 3rd order interaction (p=0.14), and 2nd order interaction of LiSN-S condition by group (p=0.91) were not significant. However, there was a significant interaction of LiSN-S condition by time [F(3,111)=3.7; p=0.014; partial eta squared = .09], indicating that listeners' SRTs changed by different amounts in the different conditions. Although the lack of the 3rd order interaction implies that this effect was not different between the two groups, the changes were numerically greater in the AT group, with a tendency for more improvement in the two conditions with spatial separation between target and maskers (by about 1 dB).

Most importantly, the time by group interaction was highly significant [F(1, 35)= 27.0; p<0.001; partial eta squared = .43] indicating that the trained group improved its performance by more (≈ 1.5 dB) than the untrained group (≈ 0.1 dB). The effect size (Cohen's d) for the difference in averaged LiSN-S SRTs was large at 1.7.

The main effects of time and condition were also highly significant (both p<0.001). The time by group interaction reflects the training effect which has also influenced the main effect of time, and the large effect of condition is well known and expected (Cameron and Dillon, 2007).

The sustainability of this improvement in SRTs was evaluated in the trained group only using a repeated measures ANOVA (3 test times x 4 LiSN-S conditions). Again, Hyunh-Feldt epsilon corrections were applied to all F tests involving LiSN-S condition because Mauchley's Test of Sphericity indicated a violation for this factor (p<0.005).There was no time by condition interaction (p=0.13), but there were highly significant main effects of condition [F (2.4, 44.8) = 184.0; p < .001; partial eta squared = .91] and time [F (2, 38) = 23.8; p < .001; partial eta squared = .56] effects. Helmert contrasts showed a significant difference between baseline and subsequent testing points [F (1, 19) = 93.4; p < .001; partial eta squared = .83], but no significant difference between the SRTs immediately and 3-months post-intervention, [F (1, 19) = .49; p = .49] indicating sustainability of improvement. In fact, the mean SRT after 3 months was slightly lower than immediately after the intervention, by about 0.25 dB.

Changes in derived measures

Because of their use in clinical applications, we also applied a repeated measures ANOVA on the 3 derived LiSN-S advantage scores, comparing the two groups before and at the end of the intervention period (2 times x 2 groups x 3 LiSN-S measures). Crucially, no interaction term involving group was significant, meaning that the intervention had no effect on changes on these outcome measures, which is not surprising given that all four base measures improved in the trained group by roughly the same amount. Only one of the four interaction terms were significant, that of time by advantage score [F(2,74)=3.7; p=0.029; partial eta squared = .09], meaning that advantage scores changed by different amounts. Paired t-tests comparing the advantage scores at the two times showed a significant improvement (p=0.006) only for the Total Advantage (the difference between the SRTs for SV0° and DV90°). Although significant, even this change was small with only a 0.7 dB improvement over the two times. All these findings are consistent with the analyses on the four individual SRT measures above.

Changes in functional listening skills following training

The total pragmatic profile and CHAPS scores at baseline and at 3 months (post intervention) are shown in table 4. Six of the PP questionnaires (2 from the AT group; 4 from the control group) were incomplete with more than one question rated as “not applicable” and analysis was thus conducted on 33 PP questionnaires only. Similarly, 4 CHAPS questionnaires from the AT group and 2 from the control group were excluded from the following analysis, as some of the questions were unrated and scores could not be tabulated.

a) *Pragmatic Profile (PP)*

A repeated measures ANOVA revealed a significant time by group interaction [F(1,31) = 8.0, p= .008, partial eta squared = .205], showing that improvements in the trained group were larger than those in the untrained group (Fig. 2). Separate paired t-tests for the two groups show a highly significant change in the trained group (t(17)=4.3, p=0.001) and none in the untrained group (t(14)=4.3, p=0.3), which confirms the omnibus test was not overly sensitive to the differences in variability in PP scores between the two groups at baseline. Cohen’s d calculated from difference scores across the groups was 1.0.

b) *CHAPS*

A repeated measures ANOVA showed a significant time by group interaction [F(1,31) = 4.9, p= .035, partial eta square = .136], indicating that the trained group again improved more than the untrained group. This effect was weaker than for the Pragmatic Profile, in that separate paired t-tests for the two groups show no significant change for either group. The significant time by group interaction was presumably found because the untrained group's scores worsened slightly over time. Cohen's d as calculated from the difference scores across groups was 0.76.

Correlation between changes in AP skills and changes in functional listening abilities of children with APD

Due to outliers (in particular, one trained listener who improved the most by far on the LiSN-S and the CHAPS), robust methods were used to explore the relationships among the changes in the three outcome measures (Wilcox 2012). First, a boot strap method was used to evaluate Pearson correlations among changes in the two functional measures of listening and performance for speech in noise (by subtracting the baseline value from the post-intervention value for each individual). One-tailed tests were used because of the predicted direction of the correlation. The correlation between the two functional measures was relatively weak ($\rho=0.31$, $p=.046$), and would not survive a Bonferroni correction, so this was considered no more.

Much stronger relationships were found between changes in mean LiSN-S and the two functional measures (PP: $\rho=-0.44$, $p=0.007$; CHAPS: $\rho=-0.64$, $p<0.001$), so these were investigated more thoroughly. Of primary interest is the extent to which changes in the outcome measures are correlated within groups, especially for the trained group (even though these tests have less power because of the splitting of participant numbers into two groups). Again, these were examined with a bootstrap method. Neither of the two correlations were significant in the untrained group ($p>0.18$ for both), as would be expected from the narrow range of changes in LiSN-S in this group. In the trained group, changes in LiSN-S were not a significant predictor of changes in PP ($\rho=-0.28$, $p=0.15$), but they were for CHAPS, with the correlation of similar magnitude to that obtained in the whole group ($\rho=-0.66$, $p=0.003$).

Figure 3 depicts the relationship between the change in mean LiSN-S and the change in CHAPS, with a single fitted line because a separate robust analysis shows the slopes and intercepts of the two groups being indistinguishable ($p>0.5$ using ols1way() in Wilcox & Clark 2015).

Other analyses

From the baseline measures, neither language and phonological skills, nor nonverbal IQ and auditory memory correlated with the changes in the overall LiSN-S performance. As more than half the number of participants from whom data was available completed all the specified training sessions, we did not try to relate total hours of AT to changes in any of the outcome measures.

Discussion

We found that children with APD who had undergone a 12-week intensive speech-based auditory training programme showed on average improved speech-in-noise test performance. These improvements also correlated with improvements in observer-rated communication behaviours, as assessed by questionnaires, indicating that this training led to real life benefits. No such improvements were found in untrained control children with APD. These results are to some extent consistent with two case series studies (Cameron and Dillon 2011; Cameron et al. 2014), and a small randomised controlled trial (N=10) (Cameron et al. 2012). These authors reported significant improvements in children with APD who trained with the LiSN & Learn programme on individual low cue (Cameron et al. 2014) and high cue SRTs (Cameron and Dillon 2011; Cameron et al., 2012; Cameron et al. 2014) as well as on questionnaires that assess real life listening (Cameron and Dillon 2011; Cameron et al. 2012;

Cameron et al. 2014). Taken together, these results provide further evidence for the benefits of such training. However, there are some inconsistencies between the different studies. In contrast to the findings reported by Cameron and Dillon (2011), Cameron et al. (2012) and Cameron et al. (2014), performance in the four LiSN-S conditions showed comparable improvements, rather than a specific improvement in conditions with spatial separation. This may be because the previous studies recruited and trained children not on the basis of a general diagnosis of APD as we did, but on the basis of a spatial deficit on the LiSN-S, a so-called spatial processing disorder (SPD). SPD is present in 6% in a population with high incidence of chronic otitis media (Cameron et al. 2014) and up to 15% in children referred for speech in noise difficulties (Cameron and Dillon 2011) and may have been present in very few of our study children. We could not test for SPD due to the lack of norms for the Singaporean population. The difference in recruited populations may thus account for the difference in the observed results. Cameron et al (2012) have similarly reported no benefits of Earobics training on the LiSN-S scores for 5 children with SPD, arguing that AT intervention for APD needs to be deficit specific. However, this lack of improvement for the LiSN-S derived measures that was observed in our study may be due to other protocol differences between the studies. Training with a speech target that was spatially separated from the noise was only done during one of the four AT games the children played, thus reducing total training time for this task. We employed outcome measures using test material and talkers that had not been used for training purposes, while the previous studies used the same female voice as target in both training and outcome measures. Listeners, however, perform better with a familiar talker than an unfamiliar one (Nygaard et al. 1994). While the subtractive procedure is argued to eliminate the effect of talker familiarity, it is still possible that the improvement in spatial advantage could be greater for a trained talker than an untrained one. Dosage effects may also need to be considered, as those who complete less than a “threshold” number of AT sessions show significantly poorer outcomes versus those who complete more sessions (Chisolm et al. 2013), and a strong correlation has been reported between LiSN-S benefit and the number of LiSN & Learn sessions accomplished (Cameron et al. 2014).

Real life communication skills as reflected on the PP questionnaires improved in the AT group only. The PP was filled in by parents, who were not blinded to the intervention, and a potential bias, due to a tendency of the parent to provide a pleasing response to the researchers cannot be excluded (Lam & Bengo 2003). However, we also found a correlation between the LiSN-S and CHAPS improvements in the trained group only. This suggests that benefits were not due to a simple halo effect of the intervention, but was directly caused by the change in SRTs. The CHAPS was filled in by the teachers who were blinded to the intervention, while in several cases, the baseline CHAPS and the 3 month CHAPS were filled in by different teachers. Thus auditory training benefits appeared to generalise to better listening in the classroom environment, as rated by the teachers.

It would be tempting to attribute these benefits to improved auditory processing per se. Benefits in laboratory tests after AT with noise have been reported previously in typically developing young adult listeners, claimed to be underpinned by the enhancement of cues to pitch as measured in the frequency following response (Song et al. 2012). However, while auditory processing test performance improves after different types of auditory training, this improvement does not necessarily correlate with and thus account for the broader functional improvement of the child as reflected on language measures (Gillam et al. 2008). The language improvements may thus be related to improvements in general cognitive skills by the auditory tasks per se, by the combined use of intensive auditory and visual stimulation, or by the task cognitive requirements of the computer games. Attention/memory processes are important for speech-in-noise perception (Schmithorst et al. 2012) and have been reported to improve- to some extent, and in terms of some subcomponents- in several studies of children with language related disorders or APD (Sharma et al. 2012; Stevens et al. 2008). The relative effect of auditory processing vs. cognitive type improvement vs the interaction of both improvements on the observed improved communication of children following AT remains an intriguing and debatable question.

The observed improvements in our AT group were sustained for at least 3 months in speech-in-noise test performance, similar to reports for children with SPD (Cameron & Dillon, 2011) indicating that speech based training may lead to sustainable improvements. We also found that neither baseline language nor cognitive abilities predict the degree of improvement with AT, in agreement with other studies (Sharma et al. 2012).

There are some limitations in this study. Firstly, due to the lack of normative data for a speech-in-noise test for the Singaporean population, we were unable to determine whether the children with APD actually had any speech-in-noise or spatial processing deficits. The training incorporated several different speech-in-competition tasks and it is unclear if the improvement in the AT group was driven by a specific exercise. Further studies would need to consider separating the different types of training to examine their effectiveness, and whether this depends upon the individual. The study was unblinded, and we did not include an active control group to assess for other intervention-related effects, which could have been related, e.g., to the parent engaging with the child or to the child conducting activities designed to promote thinking and problem solving (Gillam et al. 2008). However, whilst acknowledging the ways in which the study design could be improved, it also had many strengths. Inclusion of a no-treatment control group helped assess to what extent the changes in the outcome measures were due to the intervention vs. maturational changes over time (Loo et al. 2010). The outcome measures included tests assessing speech in noise auditory processing together with questionnaires assessing functional listening, while test materials (in Australian English) were completely unrelated to and in a different accent than the training materials (in British English). One questionnaire was completed by teachers who were blinded to the intervention. Thus, observed post-training improvements appear to reflect a genuine learning effect. Effects are likely to generalize to other clinical populations with APD, in that children were recruited from a general audiology department, and without excluding participants with language related or other developmental disorders (with the exception of autism and low IQ).

In conclusion, a 12-week long 5-day/week training with speech stimuli ranging from single words to complex sentences in the presence of competing stimuli under different conditions of spatial separation (thus resembling real-life listening conditions), led to improved speech in noise perception in tests that was reflected in improved functional listening in children with APD. Further research is required to tailor auditory training to the individualized needs of listeners.

Acknowledgements

We are grateful to the clinical team at CHILD for their help with the study, to the team from the Biomedical Informatics and Engineering school, Temasek Polytechnic (Singapore) for their help with the development of the dichotic training game and to the children, their families and teachers for participating.

Author contribution statements

All authors contributed substantially to this work. DEB conceptualised the study whilst JL, SR and DEB finalized the design. SR designed, adapted and implemented all the computer-based interventions except for the dichotic training. JL conducted the testing and auditory training. All 3 authors participated in the statistical analyses of the results and in writing the paper.

References

Ahmmed, A.U., Ahmmed, A.A., Bath, J.R., et al. (2014). Assessment of children with suspected auditory processing disorder: a factor analysis study. *Ear Hear*, 35, 295-305.

American Academy of Audiology. (2010). Diagnosis, treatment and management of children and adults with Central Auditory Processing Disorder. Available at: www.audiology.org [accessed 17/02/2014].

American Speech-Language-Hearing Association. (2005) (Central) auditory processing disorders [Technical Report]. Available from www.asha.org/policy [accessed 17/02/2014]

Amitay, S., Irwin, A., & Moore, D. R. (2006). Discrimination learning induced by training with identical stimuli. Nature Neuroscience, 9, 1446–1448.

British Society of Audiology. (2011). Position statement: Auditory processing disorder (APD). Retrieved from www.thebsa.org.uk/images/stories/docs/BSA_APD_PositionPaper_31March11_FINAL.pdf

Amitay, S., Irwin, A., Moore, DR. (2006). Discrimination learning induced by training with identical stimuli. *Nat Neurosci*, 9, 1446-8.

Bolia, R.S., Nelson, W.T., Ericson, M.A., Simpson, B.D. (2000). A speech corpus for multitalker communications research. *J Acoust Soc Am*, 107, 1065-6.

Brown, L., Sherbenou, R.J., and Johnsen, S.K. (1982).Test of Nonverbal Intelligence, 3rd edition. Texas: Pro-Ed, Inc.

Cameron S, Dillon H. (2007). The Listening in Spatialized Noise - Sentences Test (LiSN-S): test–retest reliability study. *Int J Audiol*, 46:145–153.

Cameron, S., and Dillon, H. (2008). The Listening in Spatialised Noise – Sentences test (LISN-S): Comparison to the prototype LISN and results from children with either a suspected (central) auditory processing disorder or a confirmed language disorder. *J Am Acad Audiol*, 19, 377-391.

Cameron, S., and Dillon, H. (2011). Development and evaluation of the LISN & Learn Auditory Training Software for deficit-specific remediation of binaural processing deficits in children: Preliminary findings. *J Am Acad Audiol*, 22, 678–696.

Cameron, S., Dillon, H., Glyde, H., Kanthan, S., & Kania, A. (2014). Prevalence and remediation of spatial processing disorder (SPD) in Indigenous children in regional Australia. *Int J Audiol*, 53, 326-335.

Cameron, S., Glyde, H., Dillon, H. (2012). Efficacy of the LiSN & Learn auditory training software: randomized blinded controlled study. *Audiol Res*, 2, e15.

Chermak, G.D., Tucker, E., Seikel, J.A. (2002). Behavioral characteristics of auditory processing disorder and attention-deficit hyperactivity disorder: predominantly inattentive type. *J Am Acad Audiol*, 13, 332-8.

Chermak, G.D., Bellis, T.J., and Musiek, F.E., 2007. Neurobiology, cognitive science, and intervention. In: G.D. Chermak, F.E. Musiek, eds., 2007. Handbook of (Central) auditory processing disorder, comprehensive intervention, Vol II. San Diego: Plural Publishing, pp. 3-28.

Chisolm, T.H., Saunders, G.H., Frederick, M.T., et al. (2013). Learning to Listen Again: The Role of Compliance in Auditory Training for Adults with Hearing Loss. *Am J Audiol*, 22, 339-42.

Cohen, W., Hodson, A., O'Hare, A., et al. (2005) Effects of computer-based intervention through acoustically modified speech (Fast ForWord) in severe mixed receptive-expressive language impairment: outcomes from a randomized controlled trial. *J Speech Lang Hear Res*, 48, 715-729.

Dillon, H., Cameron, S., Glyde, H., et al. (2012). An opinion on the assessment of people who may have an auditory processing disorder. *J Am Acad Audiol*, 23, 97-105.

Fahle, M. & Poggio, T. (2002). Perceptual Learning. MIT Press, Cambridge, MA.

Ferguson, M.A., Hall, R., Riley, A. et al. (2011). Communication, listening, cognitive and speech perception skills in children with auditory processing disorder (APD) or specific language impairment (SLI). *J Speech Lang Hear Res*, 54, 211-227.

Frederickson, N., Frith, U., Reason, R. (1997). Phonological Assessment Battery. London: GL assessment.

Gardner, M.F. (1996). Test of Auditory-Perceptual Skills-Revised (TAPS-R). Novato: Academic Therapy Publications.

Gillam RB., Loeb, D.F., Hoffman, L.M., et al. (2008). The efficacy of Fast ForWord Language intervention in school age children with language impairment: a randomized controlled trial. *J Speech Lang Hear Res*, 51, 97-119.

Grube, M., Kumar, S., Cooper, F.E., et al. (2012). Auditory sequence analysis and phonological skill. *Proc Biol Sci*, 279, 4496-504.

Halliday, L.F., Taylor, J.L., Millward, K.E., et al. (2012). Lack of generalization of auditory learning in typically developing children. *J Speech Lang Hear Res*, 55, 168-81.

Halliday, L.F., Moore, D.R., Taylor, J.L., et al. (2011). Dimension-specific attention directs learning and listening on auditory training tasks. *Atten Percept Psychophys*, 73, 1329-35.

ICD-10 H93.25 Central Auditory Processing Disorder. Retrieved online from http://www.icd10data.com/ICD10CM/Codes/H60-H95/H90-H94/H93-/H93.25 on 10th April 2014.

Heine, C., Slone, M. (2008). The impact of mild central auditory processing disorder on school performance during adolescence. *J Sch Health*, 78, 405-7.

Iliadou, V., Bamiou, D.E (2012). Psychometric evaluation of children with auditory processing disorder (APD): Comparison with normal-hearing and clinical non-APD groups. *J Speech Lang Hear Res*, 55, 791-799.

Lam, T.C.M., Bengo, P. (2003). A comparison of three retrospective self-reporting methods of measuring change in instructional practice. *Am J Eval*, 24, 65–80.

Loo, J.H., Bamiou, D.E., Campbell, N., et al. (2010). Computer-based auditory training (CBAT): benefits for children with language- and reading-related learning difficulties. *Dev Med Child Neurol*, 52, 708-17.

Loo, J.H., Bamiou, D.E., Rosen, S. (2013). The Impacts of Language Background and Language-Related Disorders in Auditory Processing Assessment. *J Speech Lang Hear Res*, 56, 1-12.

Messaoud-Galusi, S., Hazan, V., Rosen, S. (2011). Investigating speech perception in children with dyslexia: is there evidence of a consistent deficit in individuals? *J Speech Lang Hear Res*, 54, 1682-701.

Moore, D.R., Ferguson, M.A., Edmondson-Jones, A.M., et al. (2010). Nature of auditory processing disorder in children. *Pediatrics*, 126, e382-90.

Moore, D.R., Rosen, S., Bamiou, D.E., et al. (2013). Evolving concepts of developmental auditory processing disorder (APD): A British Society of Audiology APD Special Interest Group white paper. *Int J Audiol*, 52, 3-13.

Myklebust, H. (1954). Auditory disorders in children. New York: Grune & Stratton.
Moore, D.R., Rosenberg, J.F., Coleman, J.S. (2005). Discrimination training of phonemic contrasts enhances phonological processing in mainstream school children. *Brain Lang*, 94, 72-85.

Nygaard, L.C., Sommers, M.S., Pisoni, D.B. (1994). Speech-Perception as a Talker-Contingent Process. *Psychol Science*, 5, 42-46.

Schmithorst, V.J., Holland, S.K., Plante, E. (2011). Diffusion tensor imaging reveals white matter microstructure correlations with auditory processing ability. *Ear Hear*, 32, 156-67.

Semel, E., Wiig, E.H., Secord, W. (2006). Clinical Evaluation of Language Fundamentals, 4th edition. London: Harcourt Assessments, Inc.

Sharma, M., Purdy, S.C., Kelly, A.S. (2012). A randomized control trial of interventions in school-aged children with auditory processing disorders. *Int J Audiol*, 51, 506-18.

Smoski, W.J., Brunt, M.A., Tannahill, J.C. (1992). Listening characteristics of children with central auditory processing disorders. *Lang Speech Hear Ser*, 23, 145–152.

Song, J.H., Skoe, E., Banai, K., et al. (2012). Training to improve hearing speech-in-noise: Biological mechanism. *Cereb Cortex*, 22, 1180-90.

Stevens, C., Fanning, J., Coch, D., et al. (2008). Neural mechanisms of selective auditory attention are enhanced by computerized training: Electrophysiological evidence from language-impaired and typically developing children. *Brain Res*, 1205, 55-69.

Strehlow, U., Haffner, J., Bischof, J., et al. (2008). Does successful training of temporal processing of sound and phoneme stimuli improve reading and spelling? *Eur Child Adol Psychiatry*, 15, 19-29.

Tomlin, D. (2014). Auditory processing functional outcomes and cognition in children, PhD Thesis, Melbourne: University of Melbourne.

Wilcox, R.R. (2012). Introduction to Robust Estimation and Hypothesis Testing, 3rd Edition.

Oxford UK: Academic Press.

Wilcox, R.R., Florence Clark, F. Heteroscedastic Global Tests that the Regression Parameters for Two or More Independent Groups are Identical. Communications in Statistics - Simulation and Computation 44(3): 773-786 (2015).

Wilson, W.J., Arnott, W., Henning, C. (2013). A systematic review of electrophysiological outcomes following auditory training in school-age children with auditory processing deficits. *Int J Audiol*, 52, 721-30.

List of Figures

Figure 1: Boxplots of the LiSN scores. Each panel presents the performance across time of either the control or AT group, grouped by LiSN-S condition.

Figure 2. Boxplots of the Pragmatic profile (PP) raw scores as obtained from the parents of the AT and control groups at baseline and immediate post-intervention. Each plotted point represents the score obtained from a particular participant and time point.

Figure 3. Change in the CHAPS questionnaire as a function of the change in overall LiSN-S performance. The regression line was estimated using a robust technique (tsreg() in Wilcox, 2012).

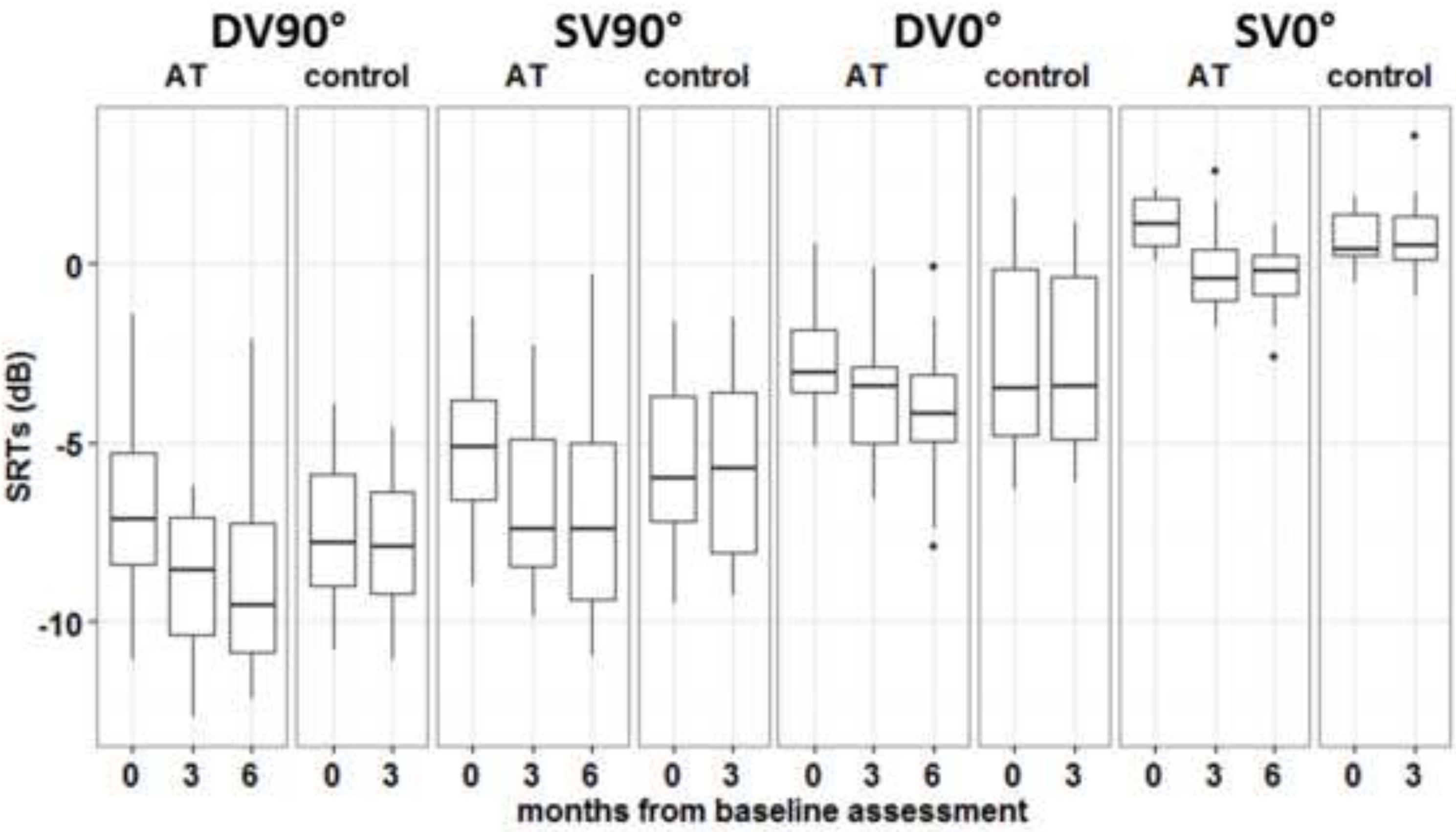
DV90°
SV90°
DV0°
SV0°
AT
control
AT
control
AT
control
AT
control
SRTs (dB)
0
-5
-10
0 3 6
0 3
0 3 6
0 3
0 3 6
0 3
0 3 6
0 3
months from baseline assessment

Figure 1

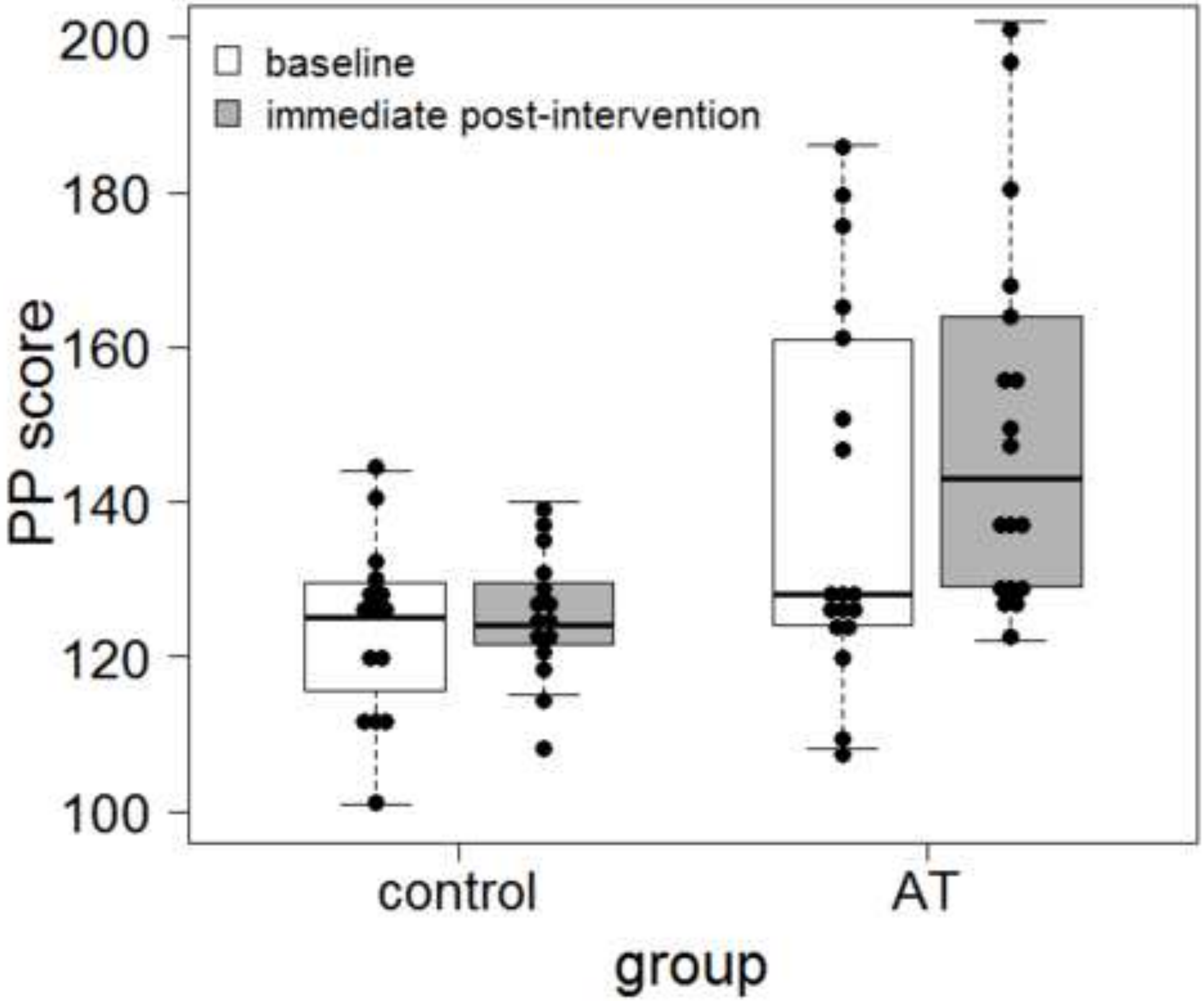
☐ baseline
▨ immediate post-intervention
PP score
200
180
160
140
120
100
control
AT
group

Figure 2

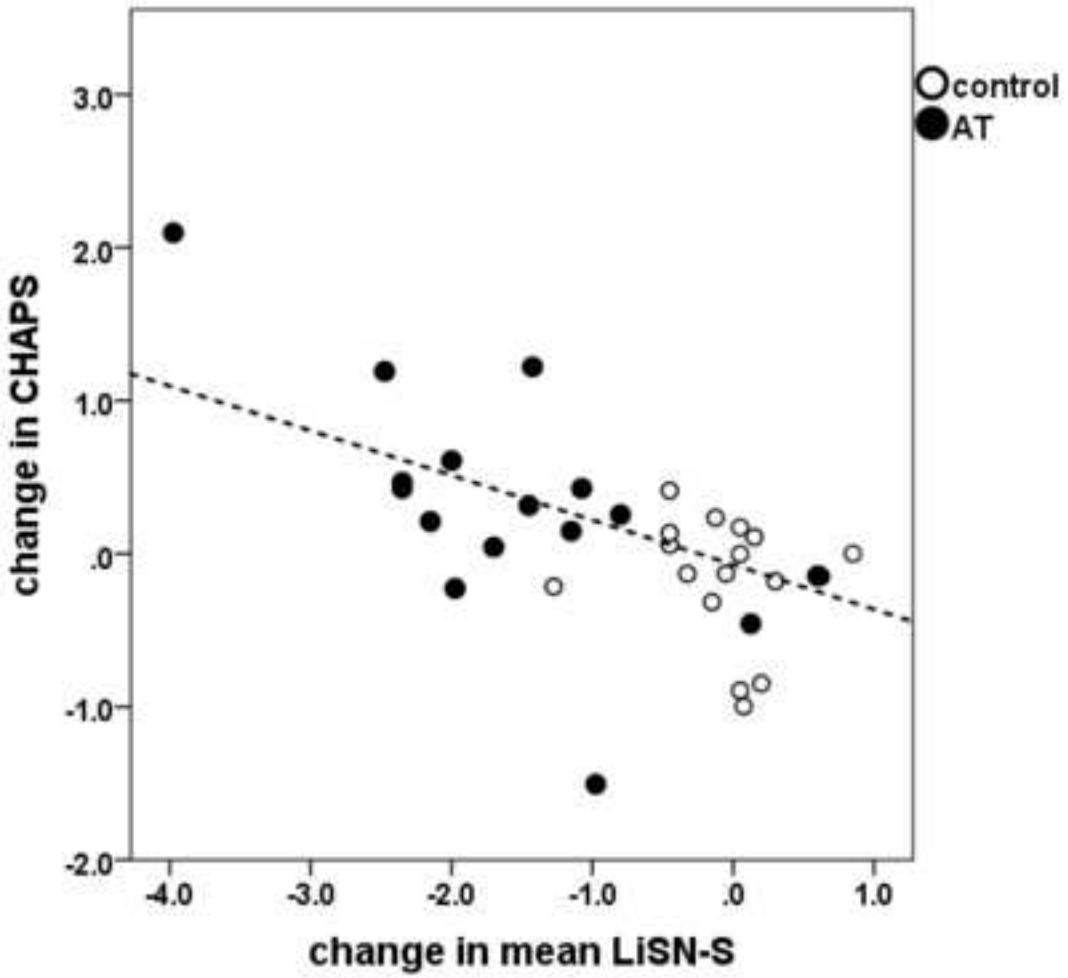
control
AT
change in CHAPS
3.0
2.0
1.0
.0
-1.0
-2.0
-4.0
-3.0
-2.0
-1.0
.0
1.0
change in mean LiSN-S

Figure 3

Table 1: The Auditory Processing (AP) Test Battery

AP tests & Technical Information	Presentation level & number of stimuli	Task	Scoring
Frequency Pattern Test (FPT) Auditec – Child version Low: 880 Hz; High: 1430 Hz; Tone duration: 500 msec; Inter-tones interval: 300 msec; Inter-pattern interval:10 sec	50 dB HL monaurally, 30 stimuli per ear	Label the tone pattern verbally as high or low in a sequence of 3 tones (e.g. high-low-low)	% correct per ear
Duration Pattern Test (DPT) Auditec Tone: 1000 Hz; Tone durations: 250 msec (short) or 500 msec (long); Inter-tones interval: 300 msec; Inter-pattern interval: 10 sec	50 dB HL monaurally, 30 stimuli per ear	Label the tone pattern verbally as long or short in a sequence of 3 tones (e.g. long-short-short)	% correct per ear
Random Gap Detection Test (RGDT) Auditec Stimuli: 0.5, 1, 2, & 4 kHz; Inter-stimuli intervals: 0, 2, 5, 10, 15, 20, 25, 30, and 40 msec. in random order.	50 dB HL binaurally, 4 sets of stimuli at different frequencies	Respond verbally to indicate whether 1 or 2 sounds were heard	Average of gap detection thresholds for 4 stimuli (ms)
Masking Level Differences (500Hz) – MLD Auditec 5 tone bursts (500 Hz; 300 msec) in 3sec bursts of narrow band noise 10 SoNo conditions (1- to -17dB S/N); 12 SπNo conditions (-7 to -29 dB S/N), and 11 no tone conditions.	50 dB HL, binaurally 33 presentation	Respond verbally whether tone pulses were heard or not within the buzzing noise.	SπNo threshold minus SoNo threshold
Dichotic Digits Test (DDT) Auditec Male voice; 25 pairs of double digits (1 to 9 except 7)	50 dB HL, binaurally	Repeat verbally all the 4 numbers	% correct per ear

Table 2

Table 2: A Summary of the Baseline Data (AP, Language, Phonological Skills, Memory and NVIQ) for the AT and Control Groups

Measures		**AT, n=20**			**Control, n = 19**			**p**
Demographic	Age [mean (SD)]	9.1 (1.33)			9.0 (1.32)			0.735
	Sex (Male)	16			16			
Non-normally distributed		**Median**	**Range**	**N (failed test)**	**Median**	**Range**	**N (failed test)**	**p**[*]
Behavioural AP	**DDT_R**[†]	88	66-98	4	90	24-98	3	.91
	DDT_L[†]	84	68-96	5	85	42-93	4	1.00
	FPT_R[†]	80	7-100	6	76	25-100	6	.91
	FPT_L[†]	80	33-100	7	76.5	25-100	7	.71
	DPT_R[†]	50	10-100	8	40	10-100	10	.59
	DPT_L[†]	60	0-100	8	51.50	0-90	10	.52
	RGDT[‡]	8.75	3-25	8	6.75	3-25	5	.30
	MLD[§]	12	4-14	6	12	4-18	5	.84
Phonological awareness (PhAB)	**Alliteration**	100	77-101	1	96	76-101	3	.08
	Rhyming	93	69-113	4	92	69-113	8	.99
	Spoonerism	103	71-119	2	106	69-128	3	.72
	Nonword reading	109	93-131	0	115	84-131	2	.79
Auditory memory (TAPS-R)	**ANFM**	97	79-127	4	92	72-133	3	.87
	ANBM	100	81-130	2	98	76-118	4	.55
	AWM	90	70-100	6	85	72-116	8	.97
	ASM	91	70-110	5	87	72-110	6	.79
Normally distributed		**Mean**	**SD**		**Mean**	**SD**		**p**[¶]

Language (CELF-4)	**Core Language**	85.6	13.3	11	79.5	15.6	11	.20
Nonverbal IQ (TONI)	**NVIQ score**	108.0	13.4	0	109.7	13.7	0	.69

ANBM = auditory number backward memory; ANFM = auditory number forward memory; ASM = auditory sentence memory; AWM = auditory word memory; DDT = dichotic digits test; DPT = duration pattern test; FPT = frequency pattern test; MLD = masking level differences; RGDT = random gap detection test; R = right ear; L = left ear, [*] Mann-Whitney test; ¶ t-test. **Note:** Unless stated otherwise, value is standard score. [†] score in %; [‡] score in ms; [§] score in dB;

Table 3: Descriptions of the Auditory Training Tasks Used.

Games	Type of training	Target Speech	Types of masker	Response mode	Algorithm
CCRM Children's Coordinate Response Measure: an expanded version of the WiNiCS task (Messaoud-Galusi, Hazan, & Rosen, 2011) based on Bolia, Nelson, Ericson, & Simpson (2000)	Speech-in-noise for words in sentences spoken by an adult British female talker (target speech and masker manipulated with respect to relative location: 0° azimuth; ± 90° azimuth).	Sentences of the form "Show the dog where the [*colour*] [*number*] is" spoken by a female adult; *number* could be 1 to 9 (excluding the bisyllabic 7); *colour* could be black, red, white, blue, green or pink.	Theatre noise; multitalker babble; competing speech by a male talker with identical sentence structure but different animal, colour and number; steady-state speech-shaped noise.	Click on the target number and colour. Corrective feedback given.	Adaptive procedure tracking 79% correct, stopping after 6 reversals or 30 trials.
Who-Is-Right?	Speech-in-noise for isolated CVC monosyllabic words spoken by an adult British female talker.	A target word and two other non-word foils differing by one feature in the initial consonant (e.g. boat, woat, poat) are presented in a random order.	Continuous steady-state speech-shaped noise.	Click on one of 3 cartoon figures to indicate the position of the target word specified previously by a picture and spoken by a male talker. Corrective feedback given.	Adaptive procedure tracking 79% correct, stopping after 42 trials.
Story-in-noise Faulkner, Rosen, & Green (2012)	Speech-in-noise for words in phrases spoken by an adult British female talker.	Phrases from a connected narrative taken from books aimed at foreign learners of English.	Continuous steady-state speech-shaped noise.	Click on 1-3 keyword(s) present in the target phrase from a set of 2-6 options (each foil being phonetically similar to the target). Corrective feedback given. The phrase is replayed every time a wrong choice is made.	Fixed at +10 dB SNR, stopping after 15 minutes.
TATP Temasek Auditory	Dichotic listening (9 different games	Digits, mono- and bi-syllabic words;	Competing speech stimuli that are similar to the target	Indicate on a computer GUI the items	Adaptive procedure

Training Programme	varying in terms of target speech stimuli and response mode). Spoken by an adult Singaporean male talker.	sentences not longer than 8 words.	speech presented simultaneously to the contralateral ear, at various SNRs for sounds across the two ears.	presented to one ear whilst ignoring the other. Attended ear varied over training. No corrective feedback given.	tracking 50% correct, stopping after 16 reversals.

Table 4: An overview of a week 1 training programme for children in the AT group

Day	**Training 1 (15 min)**	**Training 2 (15 min)**
Monday	TATP_1	Keywords extraction_1
Tuesday	DOGGY_1	Keywords extraction_1
Wednesday	TATP_2	Keywords extraction_1
Thursday	WHO-IS-RIGHT	Keywords extraction_1
Friday	TATP_3	Keywords extraction_1

Note. The above training schedule was repeated for 12 weeks with different tasks being pre-programmed in each listening game session.