Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Brain and Language xxx (2005) xxx–xxx
Brain and Language
www.elsevier.com/locate/b&l

Discrimination training of phonemic contrasts enhances phonological processing in mainstream school children[☆]

David R. Moore[a,b,c,*], Joy F. Rosenberg[a], John S. Coleman[d]

[a] MindWeavers Ltd, Oxford Centre for Innovation, Mill Street, Oxford OX2 0JX, UK
[b] MRC Institute of Hearing Research, University Park, Nottingham NG7 2RD, UK
[c] University Laboratory of Physiology, Parks Road, Oxford OX1 3PT, UK
[d] Phonetics Laboratory, University of Oxford, 41 Wellington Square, Oxford OX1 2JF, UK

Accepted 30 November 2004

Abstract

Auditory perceptual learning has been proposed as effective for remediating impaired language and for enhancing normal language development. We examined the effect of phonemic contrast discrimination training on the discrimination of whole words and on phonological awareness in 8- to 10-year-old mainstream school children. Eleven phonemic contrast continua were synthesised using linear interpolation coding from real speaker endpoints. Thirty children were pre-tested on the Word Discrimination Test (WDT) and the Phonological Assessment Battery (PhAB). Eighteen then trained for 12 × 30 min sessions over 4 weeks using an adaptive three interval two alternative phonemic matching task. The remaining children participated in regular classroom activities. In Post-testing, trained children significantly increased their age-equivalent scores on both the WDT and PhAB by about 2 years. For the PhAB, no improvement was found in the controls. Enhanced performance in the trained children was maintained in a delayed test 5–6 weeks following training. Enhancements on the trained discriminations were weak and variable. The results indicate a dramatic improvement in phonological awareness following phonemic discrimination training without matching perceptual learning.

Keywords: Perceptual learning; Language impairment; Language development; Dyslexia; Phoneme; Phonological awareness; Adaptive learning; Computer training; Word discrimination; Auditory

---

1. Introduction

Much recent debate on receptive language impairments in people with audiometrically sensitive hearing has focused on whether the impairment is due primarily to sensory (auditory perception) or linguistic (e.g., semantics, pragmatics) problems. On the sensory side, Tallal and Piercy (1973), Tallal (1980), Tallal, Stark, and Mellits (1985) have argued not only that childhood language impairments, including specific language (SLI) and reading (SRI) impairment, are due to perceptual problems, but that they are due to a specific problem usually referred to as 'auditory temporal processing.' Other advocates of a perceptual involvement in language impairments have found a wider variety of auditory processing problems in many children and adults with SLI (McArthur & Hogben, 2001; Wright et al., 1997), SRI (e.g., Amitay, Ahissar, & Nelken, 2002; Ramus et al., 2003), attention deficit disorder (Kraus et al., 1996), autistic spectrum disorders (Siegal & Blades, 2003), and behaviour problems (Hill, 2000; King

---

[☆] David Moore is a founder, shareholder, and non-executive director of MindWeavers Ltd. Joy Rosenberg is a consultant to the company without other financial interest. John Coleman has done consultancy work for MindWeavers on behalf of Oxford University.

[*] Corresponding author. Fax: +44 115 951 8503.
E-mail address: davem@ihr.mrc.ac.uk (D.R. Moore).

0093-934X/$ - see front matter © 2004 Published by Elsevier Inc.
doi:10.1016/j.bandl.2004.11.009

& Stephens, 1992). On the linguistic side, Mody, Studdert-Kennedy, and Brady (1997) have presented empirical data and theoretical arguments in favour of a speech-specific failure in phonological representation in language impaired individuals. Other researchers (Bishop, Carlyon, Deeks, & Bishop, 1999; Rosen, 2003; Rosen & Manganari, 2001) have been unable to find auditory processing difficulties consistently in those individuals, leading to suggestions that sensory problems may be "neither necessary nor sufficient for causing language impairment in children" (Bishop et al., 1999). A conciliatory view is that while auditory processing difficulties contribute to many cases of language impairment, there are other cases where no evidence for auditory processing difficulties has been found (cf. Ramus, 2003).

A further and somewhat independent issue is whether training on auditory processing tasks can ameliorate those impairments. There is a substantial literature on the use of phonological training to treat language impairments (e.g., Wise, Ring, & Olson, 1999, 2000). This training typically uses a variety of approaches, including teacher and computer delivered instruction of syllable manipulation and grapheme–phoneme matching. In contrast to methods used to measure auditory processing, phonological training does not adapt trial by trial to the trainees performance. In well-controlled studies, intensive phonological training has been found to improve phonological awareness, reading, and other language skills. The mechanisms of that improvement, however, remain obscure. In this study we took a more focussed perspective, asking whether adaptive training directed specifically at improving auditory processing would also improve receptive language skills.

Practice effects and other non-stationary influences (e.g., attention) on hearing have been recognised for as long as hearing has been measured. However, for most psychoacoustic studies those influences have been regarded as negative and something that had to be controlled (Moore, 2001; Zwsilocki, Maire, Feldman, & Rubin, 1958). Data on what became known as perceptual learning have been available for over 100 years, with reasonably distinct epochs of peak activity in the 1950s and 1960s (Gibson, 1967) and over the last 10 years (Ahissar, 2001; Fahle & Poggio, 2002; Moore, Amitay, & Hawkey, 2003). Perceptual learning has recently been defined broadly as "relatively long-lasting changes to an organism's perceptual system that improve its ability to respond to its environment and are caused by this environment" (Goldstone, 1998, p.586). While perceptual learning occurs to some extent through passive exposure to stimuli, the most efficient way to promote the learning is by active training (Gibson, 1967). In short, perceptual learning is most efficient when the trainee is alert, well motivated, and working hard. We use the term perceptual learning here to include the use of phoneme and syllable speech stimuli, consistent with Goldstone's definition.

Although the benefit of training for applied sensory tasks (e.g., wine tasting; Goldstone, 1998) has been long recognised, it was not until quite recently that the therapeutic opportunity of perceptual learning began to be appreciated (Hurford & Sanders, 1990; Polat, Ma-Naim, Belkin, & Sagi, 2004). Therapeutic auditory training appears to have been mostly clinic based, scientifically underpinned by known deficits in auditory processing abilities (Chermak & Musiek, 1997), and known practice/training effects in psychoacoustics (as reviewed above). However, in 1996, Merzenich, Tallal, and their colleagues showed that intensive, adaptive training on a variety of auditory tasks could dramatically improve the ability of children with 'language-based learning impairments' (LLIs) to perform standardised tests of auditory processing (Merzenich et al., 1996) and language (Tallal et al., 1996). Based on findings that children with LLI have difficulty identifying or sequencing two brief sounds (Tallal & Piercy, 1973; Tallal et al., 1985), particularly when presented rapidly, the research showed that digitally processing speech and non-speech sounds to extend them in time and to amplify rapid transitions improved the ability of those children to distinguish the sounds. In the study of language learning (Tallal et al., 1996), the children were given, in a 'Pre-training' phase, several standardised tests of language and the 'Tallal Repetition Test' (TRT; Tallal, 1980), a measure of sound sequencing for two brief tones presented at variable interstimulus intervals. As would be expected, the LLI children did poorly on these tests in terms of age-equivalent scores, relative to typically developing children. Next, in a 'training' phase, the LLI children conducted a cycle of 10 different listening exercises that included adaptive training, computer-based games (Merzenich et al., 1996), and exposure to acoustically modified speech of varying complexity, some of which was also presented as interactive computer games. A total of 88–116 h of training was given over 4 weeks. Some children received equivalent exposure to the games and the language exercises, but without adaptive training and with natural (i.e., not temporally modified) speech. 'Post-training' testing showed that all children improved in the standardised language tests, but that those who underwent adaptive training with modified speech improved more. Age-equivalent language scores of adaptively trained children improved by approximately 2 years.

The research of Tallal, Merzenich, and their colleagues was ground breaking in several respects. Based on theory, it identified a possible training solution with novel use of both computer games and psychophysical procedures. It achieved highly statistically significant results while, unusually in the fields of educational and remedial software, maintaining high standards of scientific rigor. Because of the potential importance of the results, however, many issues raised by these studies require further investigation. In the study reported here we focussed on four issues. First, whether training with just one type of sound stimulus and one procedure could produce learning. Second, whether less extensive training could produce learning. Third, whether training of typically developing children could improve their language skills. Fourth, whether the learning would persist after the training.

To address these issues, we designed a new type of adaptive language training computer game that was based on discrimination of phonemes representative of the major phonological categories of (British) English. The choice of this type of training material was based on the poor stimulus generalisation often reported in studies of perceptual learning (Ahissar & Hochstein, 1997; Irvine, Martin, Klimkeit, & Smith, 2000), coupled with our desire to train real language skills. The training was active and adaptive—we knew that learning occurred most efficiently for alert, well-motivated listeners who were performing at near threshold levels; at their 'edge of competence.' We chose a training task that did not require categorical judgements ('naming'), since we wished to control for variations in response criterion and because some listeners may have a specific problem with naming, presumably due to the high cognitive demand of such a task. To promote motivation, training was provided in relatively short chunks (30 min), and the task was embedded in a computer game with graphics designed by a commercial game developer. Auditory training was interleaved with play on an arcade-style game whose purpose was purely fun. Together, these design features were used to ask whether appropriately presented phoneme discrimination training can be used to improve receptive language skills in 8- to 10-year-old typically developing children.

2. Materials and methods

2.1. Participants

Thirty 8- to 10-year old children enrolled in year 4 of a mainstream primary school in Oxford, UK. The children were assigned to one of two groups by the school head teacher on a whole-class basis and without detailed knowledge of what each group would do. Eighteen children (the 'Trained' group; ages 8:07–9:11; 11F, 7M) completed 4 weeks of training, and both Pre- and Post-training assessments. Sixteen of these children completed a further, Delayed assessment 5–6 weeks after completion of training. Two additional children withdrew from the study prior to completion of training and two children were unavailable for the Delayed assessment. Twelve children (the 'Control' group; ages 8:07–9:06; 6F, 6M) received two assessments, separated by 4 weeks and identical to those of the Trained group. For comparison with the Trained group, these will also be referred to as Pre- and Post-tests. One child in the Trained group and 3 children in the Control group had first languages other than English. Letters of invitation to participate were sent to candidate schools, and then to parents of individual children. The letters explained the nature of the research and the general procedures to be used. Positive responses were obtained from the parents of all participants. Other criteria for participation were 'normal' hearing (by parental and teacher report) and ability to use a computer mouse, space bar, and arrow keys.

2.2. Training

2.2.1. Training game

The game ('Phonomena') had two main parts, a training section (the 'Sound Game') and a reward, arcade-style section (called '3's Company'). The Sound Game (Fig. 1A) was presented as a learning exercise in which, on each trial, a tutor (a dinosaur character, 'Rex- T') first 'mimed' a syllable, drawn from a library of sound sets. Next, two furry cavemen characters ('Mic' and 'Mac'), each mimed a syllable, one of which was identical to that produced by Rex-T. The players' task was to choose who of Mic or Mac produced the sound that matched the sound made by Rex-T. The 'miming' involved a simple flip from mouth closed to mouth open. There were no other facial or mouth movements. In the following trials, an adaptive staircase procedure was followed to vary the difficulty level of the matching task. A total of 60 trials was presented in the Sound Game, with the same sound set used throughout. Correct or incorrect responses on each trial were indicated by a bell or a hooter sound, respectively, immediately following the response. The cumulative correct trial score was indicated numerically in the top right corner of the screen. The difficulty level of the game (i.e., adaptive level) was indicated by an 'elevator gauge' in the top left corner of the screen and by text ('good,' 'great,' 'excellent') that appeared next to the gauge when performance reached criterion levels. A brief animation (~10 s) followed each completed sound game.

The 3's Company game (Fig. 1B) consisted, briefly, of 3 faces that were rolled sequentially and in random order along a conveyor belt to a catapult. The user aimed the catapult at other faces aligned under a plunger and attempted to make groups of 3 faces by aiming the catapult. When complete, each group of faces earned points and dropped out of the game. The object was to complete as many groups in the time available, initially 1 min. After each complete Sound Game, the number of faces to be grouped and the time for completion of 3's Company increased slightly to motivate players further.

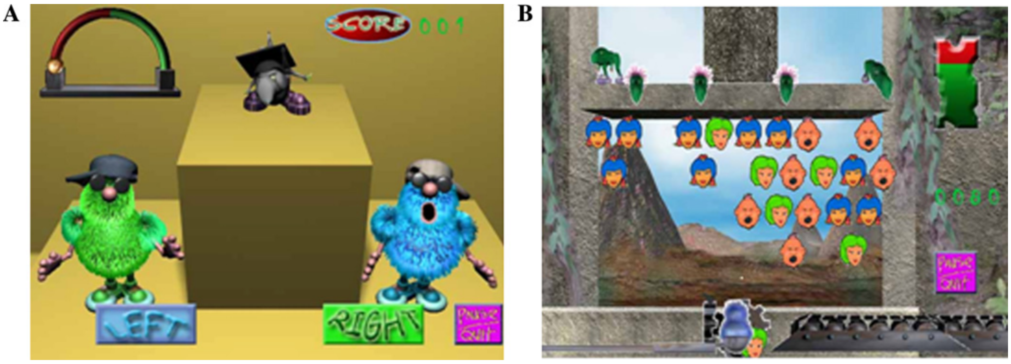

Fig. 1. Computer screenshots of (A) the phoneme discrimination (Sound) game and (B) the reward (3s-company) game. In the Sound game, the Rex-T character uttered a sound (Table 1), combined with a visible 'mouth opening.' This was followed by the other characters (Mic and Mac) attempting to copy the sound. The user's task was to choose who of Mic or Mac correctly copied the sound. In 3s-company, the user's task was to form matching sets of Baby, Blue hair, Green hair, from a supply provided by a conveyor belt at lower right.

Phonomena was designed to run from a CD on Windows-based computers containing at least Pentium II processors running at >500 MHz.

2.2.1.1. Sound sets. Sound sets ($n = 11$; Table 1) exemplifying almost the full range of phonological contrasts in British English were constructed as continua of 96 sound files compressed into a single data file. The sound files were simple syllables consisting of either a single vowel or a consonant–vowel combination. Most of the syllables were meaningless; some of them (*lee*, *err*, *or*/*awe*, *bee*, *dee*, *mar*, *shah*) were real words, mostly of low usage frequency. The endpoints of each continuum (e.g., *bee* and *dee)* were derived by analysis and re-synthesis of naturally spoken recordings of those syllables. The 94 files between the endpoints were obtained by interpolating between the acoustic parameters of the endpoint files.

The sound set to be used for a Sound Game was selected in a password accessible 'Administrator login' section of Phonomena. Sound set selection was controlled by one of three adult coaches who supervised the children as they trained during the first 2 weeks. Only 2 coaches were needed during the second 2 weeks when routines were well established. Sound sets were changed each time a student completed both parts of Phonomena. Training began with the default (b–d) and progressed with other sound sets in the order shown in Table 1, starting back at b–d when all the sets were completed. All children completed each sound set at least once, many completed or came close to completing a second round, and some sets were used further; up to 7 times in one case (see Section 2.2.1.1).

2.2.1.2. Contrast recording and synthesis. Four tokens of each of the syllables listed in Table 1 were produced by a young adult male speaker of a south-east British English dialect. To elicit the tokens, the informal syllable spellings (Table 1) were presented in random order on a PC monitor at intervals of 4 s. Mono sound recordings were obtained in a sound-insulated booth at a sampling rate of 44.1 kHz. These recordings were low-pass filtered and down-sampled to 11.025 kHz. The leading and trailing silences of each file were manually edited to time-

Table 1
The 11 phonemic contrasts used in training

Sound set name	Phonetic representation	Phonemic contrast	Informal description of syllables	Training order
b–d	/biː/–/diː/	[±labial]	"bee" vs. "dee"	1
d–g	/dɑː/–/gɑː/	[±coronal]	"dar" vs. "gar"[a]	2
e–a	/ɛ/–/a/	[±low]	"eh" vs. short "a"	3
er–or	/əː/–/ɔː/	[±round]	"err" vs. "or" (or "awe")	4
i–e	/ɪ/–/ɛ/	[±high]	"ih" vs. "eh," as in "bit" vs. "bet"	5
l–r	/liː/–/ɹiː/	[±lateral]	"lee" vs. "ree"	6
m–n	/mɑː/–/nɑː/	[±labial]	"mar" vs. "nar"	7
s–sh	/sɑː/–/ʃɑː/	[±anterior]	"sar" vs. "shah"	8
s–th	/sɑː/–/θɑː/	[±distributed]	"sar" vs. "thar"	9
v–w	/vɑː/–/wɑː/	[±sonorant]	"var" vs. "wah"	10
a–uh	/a/–/ʌ/	[±back]	short "a" vs. "uh," as in "bat" vs. "but"	11

Consonant phonemes were incorporated into CV utterances, the vowel component of which was acoustically identical for all 96 tokens. Training with each set progressed in the order shown (see Section 2).

[a] The final "r" is not pronounced in the standard, southern variety of British English (RP).

align the onset of the vowels for all tokens of the endpoints of each continuum.

Sound files in each continua were generated using linear prediction analysis and re-synthesis. Only spectral parameters (reflection coefficients) were manipulated: the natural durations of the original stimuli were not adjusted, and a single set of source parameters (e.g., voicing and f_0) from one or another of the endpoints were used for all of the tokens in a continuum. To determine which particular tokens of the endpoint files to use for a given continuum, all of the tokens were first analysed into linear prediction parameters. The spectral parameters were 15 reflection coefficients over a 6.7 ms window and the four voice source parameters included f_0 and voicing. Source and spectral parameters were thus obtained at the rate of 150 vectors/s. With these settings it was possible to obtain quite high-quality synthetic reproductions of natural speech, though the naturalness was somewhat degraded by the subsequent processing steps.

The endpoint recordings were re-synthesised using the source parameters of the *other* endpoint of the continuum in order to assess which endpoint's source parameters to use for all of the files in the continuum. For example, synthetic versions of *dar* were generated that combined the spectral parameters of *dar* with the source parameters of *gar*, and synthetic *gar* tokens were generated using the source parameters of *dar*. These "cross-synthesised" stimuli were impressionistically assessed for naturalness by members of the research team. Stimuli which contained obvious synthesis errors (e.g., clicks, pops, and buzzes) were removed, and the most natural-sounding of the remainder were retained for use in generating the complete continua.

Though the source parameters were held constant for each continuum, files of intermediate *spectral* parameter values were generated by linear interpolation between the endpoint values, using in-house software. In this way, the generation of sound sets was automated, once the initial impressionistic choice of endpoint tokens had been decided. After generation, the amplitude of each sound file was scaled so that all of the files in a sound set had the same overall RMS power.

2.2.2. Training procedure

The school's computer suite was booked for an hour thrice weekly (Mondays, Wednesdays, and Thursdays) and the Training group was divided in half to allow for 30 min sessions. The Monday and Thursday sessions were conducted at the start of the school day and the Wednesday sessions were conducted after lunch. Predictably, Wednesdays' after-lunch sessions provided the most challenge in helping students to stay on task. Based on the experience of pilot studies, an extrinsic reward scheme was put in place. The coaches (see Section 2.2.1.1) provided encouragement verbally and redirected attention when needed. They awarded stickers for every 5 min of play, approximately the time needed to complete one game. Students collected stickers on a chart and could redeem them for small prizes daily (e.g., pens and rubbers), medium prizes weekly (e.g., balls and books), or a large prize at the end of training (e.g., craft kits). Five students opted for the longer term delayed gratification of large prizes.

2.2.2.1. Adaptive procedure. The Sound Game, as described above, was a 3 interval, 2 alternative, 'XAB' task that was presented, by default and throughout training, as a '3 down, 1 up' adaptive staircase. This is an efficient adaptive method (Leek, 2001) that tracks the 79% correct response threshold. On the first 3 trials of each Sound Game, the two endpoint utterances (i.e., stimuli #1 and #96) in the selected sound set (Table 1), the digitised versions of the speaker's actual utterances, were presented. If these were correctly discriminated, a staircase step size of 5 was imposed and the next pair of stimuli (#6 and #91) were presented. This cycle was repeated until a single error occurred, at which point the stimuli presented by the program separated by a further 5 at each end. The step size was kept at 5 until 4 'reversals' (changes in direction from incrementing to decrementing stimulus separation, or vice versa) had occurred. The step size was then reduced to 3, then 2, and, finally, to 1 on each successive 4 reversals. The final sensitivity ('score') was taken as the mean number of the lower valued stimulus over the last 20 trials. Thus, the more closely the stimuli approximated each other during the game, the higher the value of the final discrimination sensitivity score.

2.3. Assessments

Two forms of assessment ('outcome measures') additional to the final discrimination sensitivity score described above were used. These were conducted by one of the authors (J.F.R.) and an Oxford University graduate student, working under supervision. The testers were not blind to the children's intervention status.

2.3.1. Phonological Assessment Battery

The Phonological Assessment Battery (PhAB) (Frederickson, Frith, & Reason, 1997) was chosen as an independently derived and validated, broadly based assessment of receptive phonological skills. The PhAB has four subtests that measure receptive abilities, specifically perception and manipulation of sounds in words and the ability to decode non-words. The subtests used—and the abilities they were designed to assess—were:

- Alliteration—isolate initial sounds in single syllable words.

- Rhyme—identify the rhyme in single syllable words.
- Spoonerisms—segment single syllable words and then synthesise the segments to provide new words or combinations.
- Non-word reading—decode letter strings, tapping only the phonological processing involved in reading non-words.

The PhAB has been normalised by subtest using a large British cohort, and age-equivalent scores are available for the range 6–14 years.

2.3.2. Word Discrimination Test

To assess hearing and listening skills that rely minimally on cognitive ability, we developed (Rosenberg & Moore, 2003) and used the MindWeavers (www.mindweavers.com) Word Discrimination Test (WDT). The WDT consists of 40 test pairs of words. Seven pairs are the same, and 33 differ only by one phoneme. Words in the WDT are short, high (word corpus) frequency, and encompass the range of phonological categories of English. The words are spoken by an adult female (south-east English dialect), digitised, embedded in broadband (pink) noise (the level of which was calibrated in pilot studies to optimise the difficulty of the task for 7- to 8-year-old children), and presented through headphones via a laptop PC. The noise bursts last for 2 s and are separated by 2.5 s of silence. Words within a pair are separated by 1 s and temporally centred within the noise bursts. For each pair of words, the listener tells the tester whether the words are the same or different. Three additional, practice pairs of words (male speaker) are presented before the test.

Prior to the present study, we administered the WDT to 185 children, aged 5:6–9:6 years, in a different Oxford mainstream school. The object of this exercise was to standardise the WDT. The function relating the mean number of 'different' pairs correctly identified (out of 33) in each 6 month age group was well fit by a linear regression ($r = .91$). We were therefore able to convert WDT scores to 'age-equivalent' scores.

2.4. Control procedure

Children in the Control group participated in normal school classroom activities while their classmates trained.

3. Results

The most notable finding (Fig. 2) was that this form of discrete phoneme training led to a dramatic and lasting improvement in a broad outcome measure of phonological awareness (the PhAB).

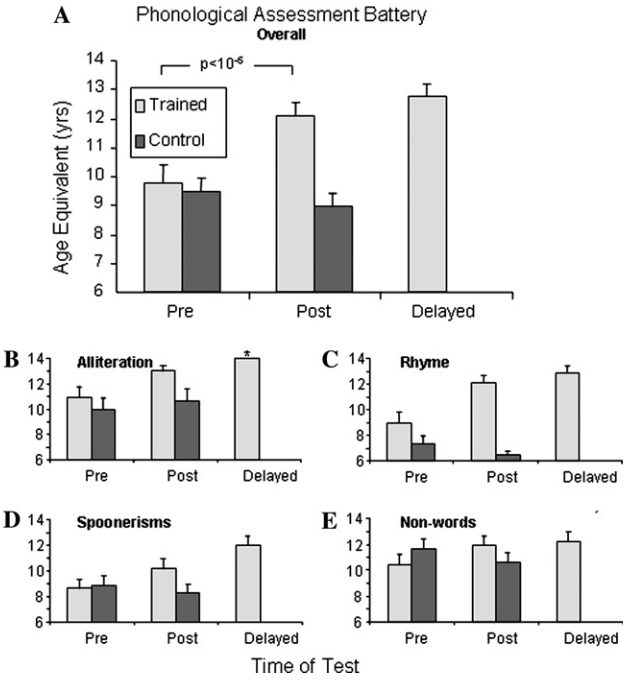

Fig. 2. Phonological Assessment Battery (PhAB) group results. Performance of the trained and control children before (Pre), immediately after (Post), and 5–6 weeks after (Delayed) the 4 week training period. All scores are referenced to the (age-equivalent) normalised British values (Frederickson et al., 1997). Overall performance (A) was the mean of performance on each of the four receptive sub-tests (B–E). Each data bar shows the mean (+*SEM*) level of all the children in that group. The asterisk (*) in (B) indicates a saturation ceiling effect where *all* the children in the trained group obtained the maximum alliteration score.

3.1. Training enhanced phonological awareness

Overall performance of both the Trained and Control groups on the PhAB (Fig. 2A) was high, with both groups scoring just above the UK normalised values for their mean chronological ages (9:2 and 9:1 years, respectively) in the Pre-tests. Individual age-equivalent scores were quite variable (range 6:0–14:0; Fig. 3), as was mean performance on the sub-tests of the PhAB (Figs. 2B–E). However, training resulted in a clear, substantial, and highly statistically significant improvement in performance of the trained group, both on the overall PhAB ($t_{17} = -6.50$, $p < .0001$) and on each of the subtests (Table 2). In contrast, the Control group did not significantly change their performance between the Pre- and Post-tests.

The correlation between Pre- and Post-overall test scores (Fig. 3A) was significant in both the Trained and Control groups, adding confidence to the validity of the PhAB in these samples. The individual performances of the Trained children all improved following training, whereas those of the Control group actually showed more decreases than increases on the Post-test relative to the Pre-test, confirming that the apparent

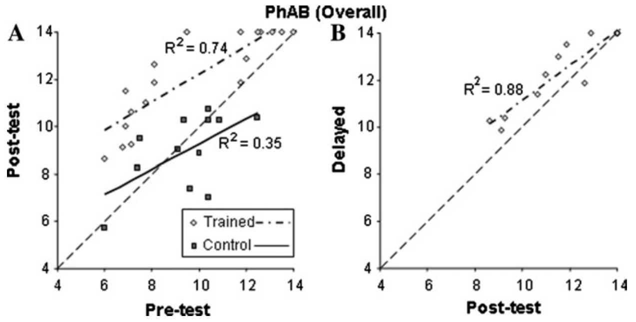

Fig. 3. Phonological Assessment Battery individual results. Comparison of age-equivalent scores (A) from Pre-test to Post-test, and (B) from Post-test to delayed. The diagonal, hatched line denotes equal performance, so points above the line show individual improvement from the earlier to the later test.

improvement in the Trained group was not simply due to test repetition. The slope of the regression line fitted to the Trained group data was less than unity, suggesting that children in this group who performed more poorly prior to training improved their performance more than those who initially performed better. However, a ceiling effect is apparent in the Post-test data, possibly contributing to the limited improvement shown by the more able children. Curiously, the slope of the regression for the Control children was also less than unity, despite the lack of either a training or a ceiling effect in this group.

Trained children continued to improve their performance over the weeks between the Post-test and the Delayed test, significantly in the case of the alliteration sub-test (Fig. 2B, Table 2), with a very high correlation between overall performance on these two tests (Fig. 3B). The slope of the regression line was, again, less than unity, despite the reduced number of children performing at ceiling level in this slightly smaller sample.

3.2. Training and repeat testing enhanced word discrimination

Both Trained and Control children improved their performance on the WDT between the Pre- and Post-tests (Fig. 4, Table 3). However, the performance gap that existed between the groups in the Pre-test (0.3 years), possibly because of the slightly lower age of the Control children, widened in the Post-test (to 0.6 years)

Table 2
PhAB comparisons—mean age-equivalent differences between groups and *t* tests

	Mean difference	*t*	*df*	*p*	α = 0.05	α(Bon) = 0.05[a]
PhAB (all)						
Trained vs. Control (Pre)	0.29	0.32	28	.75276		
Trained vs. Control (Post)	3.11	4.62	28	.00008	x	x
Pre vs. Post (Trained)	−2.33	−6.50	17	.00001	x	x
Pre vs. Post (Control)	0.49	1.11	11	.28920		
Post vs. Delayed	−0.70	−1.16	32	.25354		
PhAB (alliteration)						
Trained vs. Control (Pre)	1.01	0.81	28	.42266		
Trained vs. Control (Post)	2.38	2.53	28	.01729	x	
Pre vs. Post (Trained)	−2.03	−2.70	17	.01509	x	
Pre vs. Post (Control)	−0.67	−0.65	11	.53021		
Post vs. Delayed	−1.00	−2.07	32	.04618	x	
PhAB (rhyme)						
Trained vs. Control (Pre)	1.71	1.51	28	.14313		
Trained vs. Control (Post)	5.67	7.30	28	6.0E−08	x	x
Pre vs. Post (Trained)	−3.08	−4.53	17	.00029	x	x
Pre vs. Post (Control)	0.88	1.28	11	.22821		
Post vs. Delayed	−0.79	−0.95	32	.34814		
PhAB (Spoonerism)						
Trained vs. Control (Pre)	−0.31	−0.29	28	.77291		
Trained vs. Control (Post)	2.93	2.91	28	.00700	x	x
Pre vs. Post (Trained)	−2.61	−4.10	17	.00075	x	x
Pre vs. Post (Control)	0.63	1.60	11	.13720		
Post vs. Delayed	−0.78	−0.80	32	.42773		
PhAB (Non-word reading)						
Trained vs. Control (Pre)	−1.25	−1.07	28	.29561		
Trained vs. Control (Post)	1.46	1.42	28	.16578		
Pre vs. Post (Trained)	−1.58	−2.77	17	.01297	x	
Pre vs. Post (Control)	1.13	1.76	11	.10652		
Post vs. Delayed	−0.25	−0.26	32	.79708		

[a] Bonferroni correction for multiple comparisons.

so that, following correction for multiple comparisons, only the improvement in the trained group was significant. Nevertheless, the results suggest that, in contrast to the PhAB, repeat testing also improved performance in the WDT. The Trained group continued to improve in the third, Delayed test, but this improvement was not significant.

Comparison between individual performances on the WDT and the PhAB showed interesting differences (Fig. 5). For the WDT (Fig. 5A), the Trained group showed a modest, but significant correlation between Pre- and Post-training scores, as found in the PhAB. But the Control group failed to show any correlation, suggesting that their improved performance as a group on re-testing may have been due to factors other than test item familiarity. Pre-test performance on the PhAB and WDT is compared directly in Fig. 5B. No correlation was found, suggesting that these two tests evaluated different aspects of performance. As in other analyses, age-equivalent performance was generally higher on the PhAB, presumably due to differences in normalisation procedures between the tests.

Exclusion from analysis of the children for whom English was not the first language did not alter the statistical significance or conclusions reached from either outcome measure. However, it is possible that the larger number of these children in the Control group contributed to the slightly poorer performance of this group.

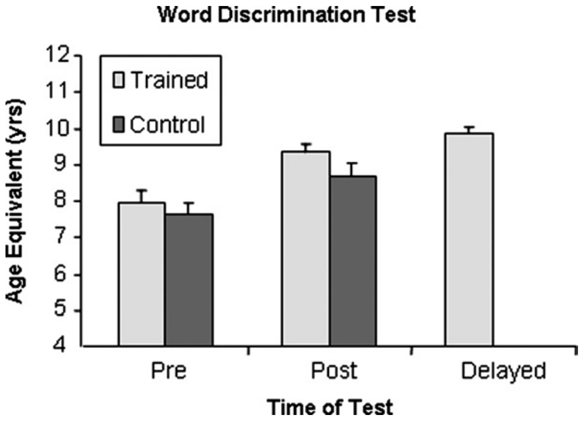

Fig. 4. Word Discrimination Test (WDT) group results. Scores are referenced to previously normalised values from another Oxford mainstream school (see Section 2). Data bars show the mean (+*SEM*) age-equivalent score for each group at the various tests (see Fig. 2).

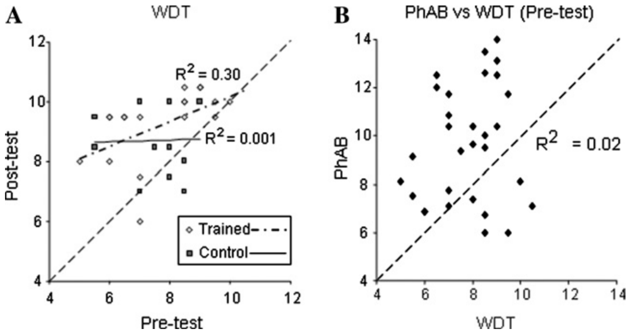

Fig. 5. (A) Word Discrimination Test individual results. Comparison of age-equivalent scores from Pre-test to Post-test (see Fig. 3 for further details). (B) Comparison between PhAB and WDT age-equivalent scores at the Pre-test.

3.3. Training produced small and variable improvements in phoneme discrimination

Children in the Trained group completed between 22 and 43 Sound Games (mean = 30), with 1–7 repetitions of each sound set, during the 12 training sessions. Performance varied widely, both between sound sets and between individuals (Fig. 6). Some sound set pairs (Fig. 6A; i–e, e–a, er–or, and l–r) were easily discriminated by almost all the children from the first game in which they were heard. For these sets, performance generally remained high throughout subsequent repetitions. Performance on about half the sets (s–sh, v–w, m–n, b–d, a–uh, and s–th) started and remained modest. Finally, sound set d–g proved extremely difficult for the children to discriminate. Little clear evidence of training is seen in these data, although some improvement is apparent in those sets for which more than three games were played. However, this may be attributable to the relatively small number of high performing children who played more than three games.

Individual children were equally variable. Some children performed consistently well and completed several repetitions of each sound set. Others tried hard, completing a large number of games, but with generally poor or inconsistent results. Still others produced monotonic learning curves. Fig. 6B shows an assortment of performances and depicts the 5 individuals who repeated a given sound set at least 5 times. This criterion was chosen so that a comparison of performance and learning could be made without being biased by the number of games

Table 3
WDT comparisons—mean age-equivalent differences between groups and *t* tests

WDT	Mean difference	*t*	*df*	*p*	$\alpha = 0.05$	α(Bon) = 0.05[a]
Trained vs. Control (Pre)	0.29	0.55	28	.58990		
Trained vs. Control (Post)	0.60	1.35	28	.18698		
Pre vs. Post (Trained)	−1.39	−4.34	17	.00044	x	x
Pre vs. Post (Control)	−1.08	−2.28	11	.04388	x	
Post vs. Delayed	−.57	−1.77	15	.09785		

[a] Bonferroni correction for multiple comparisons.

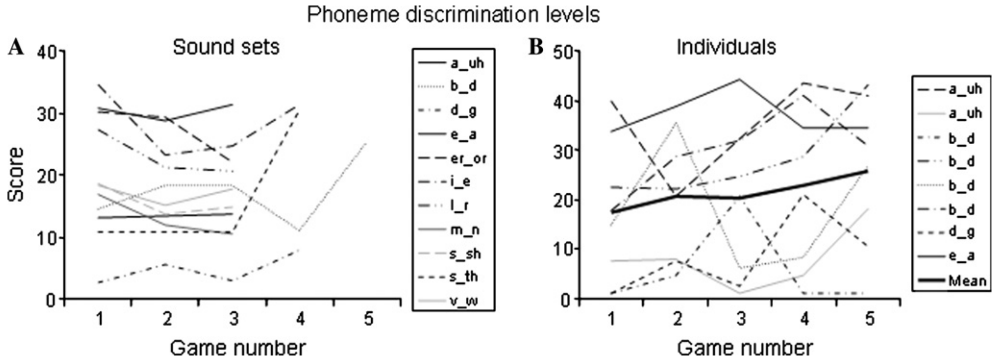

Fig. 6. Phoneme discrimination scores during training. The data points are an index (see Section 2) of the mean performance over the last 20 trials in each game. (A) Mean performance of all (trained) children on each of the 11 sound sets and for each (sequential) game completed by at least 4 children. (B) Scores of individual children on sequential games. Data show all instances for which at least 5 games were played. Note that for the b_d and a_uh sound sets, more than one child completed at least 5 games.

played (cf. Fig. 6A). Note that one child repeated two sound sets at least 5 times and one child repeated 3 sound sets at least 5 times. The data in Fig. 6B are not representative, since children who completed a larger number of games tended to be more able. Nevertheless, both the variability and the learning are apparent in these data. The mean of the 8 curves is in bold font and shows that, overall, a modest improvement in performance occurred.

Three children performed at or near chance level on the majority of sound sets (data not shown). Even these children, however, obtained relatively high scores (>25) on at least one (and usually more) sound set, showing that they understood the rules of the task. Typically, these children performed well in some of the early sessions of training, but their performance became erratic later and, as mentioned above, they failed to complete as many games as the more able children.

To focus on discrimination performance that was clearly not based on chance alone, we examined the data of individuals who scored at least 5 on all games except the first one of each sound set. This selection produced between 2 and 14 (mean = 9) individual data sets for each sound set. Representative examples are shown in Fig. 7. For moderate (Fig. 7A; a–uh) or low (Fig. 7B; d–g) levels of initial performance, improvements were usually seen in later games. However, in the extreme case of the sound set i–e (Fig. 7C), for which *all* children performed well in the first game, performance mostly became poorer in later games. These trends were seen generally in the other sound sets. Thus, children who initially performed poorly, but above chance, tended to improve with increasing practice. Those who initially performed well often got poorer. And those who performed at chance more than once tended not to improve.

3.4. Trained phoneme discrimination ability predicted outcomes

The relation between phoneme discrimination and the two outcome measures in individual children of the Trained group is shown in Fig. 8. Two indices of phoneme discrimination were used. The ‘Games’ measure was the total number of games played by each child, excluding those for which the score was less than 5 on any except the first game on each sound set. The Games measure did not correlate significantly with either Pre-test performance or ‘improvement’ (Post minus Pre age equivalent) in either the PhAB or the

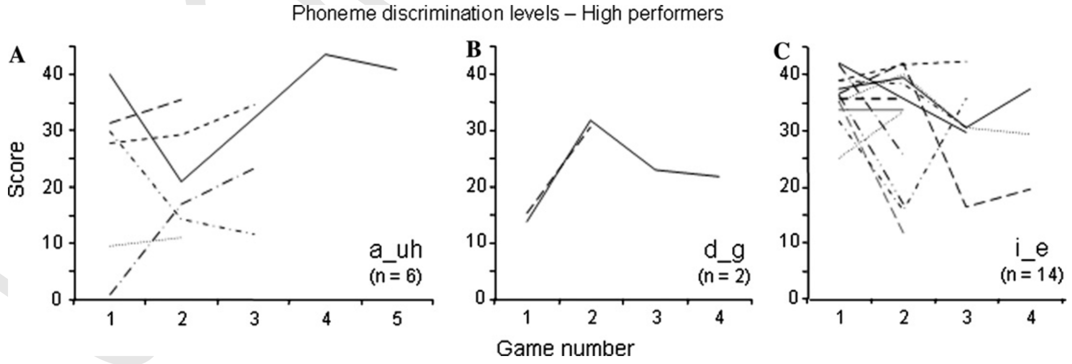

Fig. 7. Phoneme discrimination scores of high performers. Data show all individual instances for which at least two games were played on the specified sound set ((A) a_uh; (B) d_g; and (C) i_e) and the score was at least 5 on all games except the first.

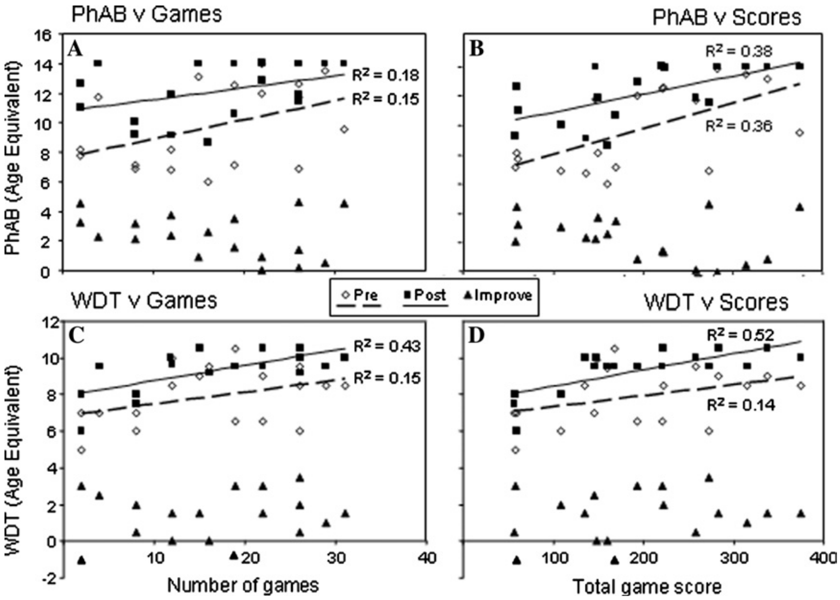

Fig. 8. Relation between phoneme discrimination and outcome measures. The indices of phoneme discrimination (see Section 2.2.1.1) were 'Games,' the extent of game play by high performers (as defined in Fig. 7), and 'Scores,' the total score of all the games played by each child, as defined by the Games measure.

WDT (Figs. 8A and C). However, for the Post-test performance, the correlation was near significance for the PhAB ($r = .42$, $df = 16$) and significant ($r = .66$, $p < .01$) for the WDT. The second index of phoneme discrimination, the 'Scores,' was the total score of all the games played by each child, as defined by the Games measure. It correlated significantly ($p < .01$) with both Pre- ($r = .60$) and Post- ($r = .60$) performance on the PhAB, and with Post ($r = .72$, $p < .001$) on the WDT (Figs. 8B and D). None of the improvement measures correlated with either discrimination measure. Thus, training generally produced a closer alignment between discrimination and outcome, but this was not related to individual improvement.

4. Discussion

This research has shown that training on a computer game incorporating an adaptive phoneme discrimination task improved phonological awareness and word listening skills. This is the first time, to our knowledge, that adaptive training using only phonemes has been found to influence performance on a broadly based language outcome measure. The performance enhancements were dramatic, lasting, and obtained in typically developing children.

4.1. Duration and generalisation of learning

The finding of such wide-ranging learning following relatively short periods of training using a limited stimulus set is both important and surprising. The possibility that automated learning as described here could be used either as part of or in addition to the school curriculum has implications for educational resourcing, teaching methods, and parental and community attitudes towards computer games, to name but a few. In the UK, the government's 'National Literacy Strategy' specifies that "Pupils should be taught to discriminate between the separate sounds in words . . . read words by sounding out and blending their separate parts . . . [and] write words by combining the spelling patterns of their sounds" (http://www.standards.dfes.gov.uk/literacy). The training offered by the process described here could potentially have a major impact on the implementation of this strategy. Our previous work (Moore and Rosenberg, 2003) has suggested that the same training method is also effective for developing word listening skills in children attending speech and language therapy.

Previous research on the application of auditory training to the development of language skills includes studies of word learning, both in native (Schwab, Nusbaum, & Pisoni, 1985) and additional (Morosan & Jamieson, 1989; Lively, Pisoni, Yamada, Tohkura, & Yamada, 1994) languages. This research has shown that, following extensive training, it is possible to improve the discrimination, identification, and production of non-native words and speech sounds (e.g., /l/–/r/ for Japanese listeners) and that this learning persists for at least several months after training. An important component of the training emphasised in all these studies was the 'high variability' of the trained stimuli.

In contrast to this hypothesis, a relatively short (140 min) period of training on a task involving matching of visual and auditory (tone) sequential patterns was found to improve reading and to enhance a tone-evoked brain potential (the mismatch negativity, MMN) in dyslexic children (Kujala et al., 2001). Like the study reported here, Kujala and colleagues found that training using simple stimuli was able to enhance performance on a broad and indirect outcome measure. Unlike our study, several distinct training games were used, the stimuli were non-speech sounds, the training was not adaptive, and the participants were reading impaired. The studies of Merzenich et al. (1996) and Tallal et al. (1996), discussed above, are the most influential in terms of application ('Fast ForWord'; see www.scilearn.com). They also used long and wide-ranging training and language impaired participants. A comparison of those studies with ours and that of Kujala et al. (2001) suggests that long and elaborate training may not be necessary to effect improvements in broad-based measures of language. Moreover, our results suggest that relatively long-lasting improvements are possible with the more restricted training we used. Another recent study (Hawkey, Amitay, & Moore, 2004) has shown that the most dramatic improvements in auditory perceptual learning, at least for the discrimination of tones, occur very rapidly indeed—within the first couple of hundred trials.

Several recent studies have questioned the efficacy of Fast ForWord (FFW) training for improving outcomes of language impaired children. A study of reading skills in children with parent-reported reading difficulties (Hook, Macaruso, & Jones, 2001) compared FFW with untrained controls and found that while, in the short term, FFW improved performance on a composite speaking and syntax index, it did had no long-lasting effect on this or other indices when the children were tested lor 2 years following training. In a very recent study (Agnew, Dorn, & Eden, 2004), it was found that FFW training enhanced auditory (but not visual) duration judgements without improving performance on tests of phonological awareness and non-word reading among a small heterogeneous group of children receiving training at a private clinic. In a third study (Cohen et al., in press), a randomised control trial, language skills among severely language-impaired children receiving concurrent intensive specialist therapy were not further enhanced by the training. Arguments could be made in connection with these studies that the samples tested were too small, too heterogeneous or too closeted. It does seem, however, that further positive results will be required before FFW achieves wider acceptance. In a recent editorial (Nature Neuroscience, 2004), it was argued that the onus is on future studies to be more methodologically rigorous and extensive in testing what seems to be an interesting and important phenomenon, the analogy being made with the huge trials required to certify a new drug. Our present view is that there will be some programmes that will work with some children, and that a broad diversity of studies each focussing on a particular population and specific aspects of the training protocol will best move the field forward.

4.2. Auditory and cognitive learning

Rapid, broad-based and long-lasting learning following training with simple speech sounds, the discrimination of which did not consistently improve, raises several questions about the nature of the learning. In particular, it might be asked whether the learning really was auditory perceptual learning or one or more other training-induced phenomena such as improved attention or memory. From an applied perspective, this may not matter, but consideration of the theoretical basis of the results may help improve future training designs, stimulate other applications, and bear on the neural mechanisms of the learning. Possible mechanisms are considered below. Here, we address the nature of the learning.

Performance on most or, possibly, all auditory tasks improves with training; there are countless examples of improved discrimination of both simple and complex sounds (see Gibson, 1967). Auditory learning of single phonemes along a contrast continuum has previously been demonstrated in adults by Kraus et al. (1995) and Tremblay et al. (1997, 1998) and shown to generalise from one continuum to another (Tremblay et al., 1997). In several unpublished preliminary studies and student projects, we have found that both adults and 8- to 12-year-old children learn to discriminate along some of the phoneme contrasts used in this study. Large numbers of typically developing children have been trained using the 'i–e' sound set and smaller numbers have been trained, either additionally or separately, with the 'b–d' and '1–r' sound sets. Briefly, the results of those studies showed consistent and monotonic improvements in discrimination over at least the first 3–5 games (each of 60 or 80 trials) when either 1 or 3 sound sets were used. The learning curves resembled those obtained routinely in perceptual learning studies using more orthodox methods (e.g., Amitay, Hawkey, & Moore, in press). In general, when small numbers of sound sets were used to train the discrimination of phoneme contrasts, classic auditory perceptual learning was observed.

There are several possible reasons why clear perceptual learning was not readily apparent in this study. The most obvious is that, with a large number of sound sets, changed after each game, the children did not experience consistent and continuous auditory stimulation. Counter points are that, in one of the studies referred to above (Jamison, 2002), involving game-wise rotation of 3 sound sets, more consistent training was observed. Also, if any generalisation of learning occurred between the sound sets, we would still expect clear training. Another possible reason for the apparent lack of training is that, as a consequence of the large number of sound sets, the training on each set by most children was relatively brief. Some data from children who completed large amounts of training were available and, on average, these did show a small, but clear training effect. However, no correlation was found between the extent of training and improvement on the main outcome measures. Thus, while some evidence for perceptual learning was found, it was inconsistent and did not predict receptive language enhancements.

To establish the nature of those enhancements it is necessary to consider other possible effects of the training. It is well recognised that young children perform more poorly than adults on psychoacoustic tasks and that their poor performance is due to both 'sensory' and 'non-sensory' components (Nozza, 1995; Werner & Bargones, 1991; Wightman, Allen, Dolan, Kistler, & Jamieson, 1989). The nature of the non-sensory component remains obscure, but is usually couched in terms of cognitive immaturity (or 'internal noise'; Nozza, 1995) and has been inferred from modelling. Studies involving both children (Deary, 1994; Hartley, Wright, Hogan, & Moore, 2000) and adults (Ahissar et al., 2000) have provided converging evidence by showing correlations between performance on cognitive and psychoacoustic tasks. A leading candidate for this non-sensory contribution is poor attention. Some recent work has attempted to address this link by studying auditory processing in children clinically diagnosed with attention deficit/hyperactivity disorder (ADHD). Unsurprisingly, many of these children performed poorly on tests of auditory frequency discrimination. But when these same children received stimulant medication ('Ritalin') to treat ADHD, their frequency discrimination improved to normal for age (Sutcliffe, 2003). It is therefore possible that the training effects seen here and in other studies are due to a combination of perceptual learning and improved attention. The key issue then becomes the definition of perception, learning, and attention. Perhaps perceptual learning is a type of improved attention, or that the constructs cannot be experimentally separated. Other types of cognitive explanation (e.g., memory enhancement) would seem equally difficult to disprove. It does seem clear, however, that an explanation based purely in terms of classic, 'sensory' perceptual learning is likely to be incomplete, at best.

4.3. Mechanisms of learning

Perceptual learning and allied phenomena have been variously explained in terms of changes in neural 'activation' in the cerebral cortex (Tremblay, Kraus, & McGee, 1998; Temple et al., 2003), shifts in the topography of sensory neuron sensitivity (Gilbert, 1998; Recanzone et al., 1992), and synaptic enhancement and gene expression (Kandel, 2001). Because studies of sensory neuron remapping inspired much of the current interest in perceptual learning, it is common in recent reports to read of explanations based in these terms (Irvine et al., 2000; Thai-Van, Micheyl, Moore, & Collet, 2003). However, neuroimaging studies offer the possibility of examining learning-induced changes occurring across the whole brain and, potentially, of dissociating the contributions of sensory, linguistic, and attention processes to the learning. Kujala et al. (2001) showed enhancement of the auditory MMN following successful audiovisual training. Since the MMN is thought to derive from 'pre-attentive' processing, the suggestion was that the training was influencing low level ("bottom up") pathways, in line with a perceptual account of the learning. The contribution of distinct cortical regions to this training was not, however, examined. Recently, Temple et al. (2003), using fMRI, showed a more complex pattern of enhanced activation across multiple cortical regions, bilaterally, following use of the Fast ForWord training package. In addition to showing changes in brain activity in areas normally associated with phonological processing (the left temporo-parietal and frontal cortex), the study found increased activity in the anterior cingulate gyrus, a region associated with attention, and the left hippocampal gyrus, associated with memory. Although the training package used in that study (Temple et al., 2003) was designed to enhance attention and memory as well as perceptual processing, it is also possible that, as for the phonological enhancement reported here, the outcome of the training was purely based on improved attention and memory rather than on a specific enhancement of perceptual processing. In this context, it is interesting that a recent attempt (Brown, Irvine, & Park, 2004) to show sensory neuron remapping in the primary auditory cortex during an auditory frequency discrimination task failed, despite the occurrence of clear perceptual learning. It therefore appears that neither changes in primary cortical topography nor substantial sensory learning (present study) are likely to be essential elements in the improvement of language skills following auditory discrimination training.

5. Uncited references

Bradlow, Akahane-Yamada, Pisoni, and Tohkura (1999) and Karni and Sagi (1991).

Acknowledgements

Thanks to the staff, students, and parents of St. Andrews First School, Headington, Oxford, for their enthusiastic participation in this research. Thanks also to Daniel Sokol for his help with the preparation of the stimuli, to Prof. Dorothy Bishop and Dr. Sygal Amitay for helpful comments on the paper, and to Bruce Robinson, CEO, MindWeavers Ltd., for his support of and contributions to this research.

References

Agnew, J. A., Dorn, C., & Eden, G. F. (2004). Effect of intensive training on auditory processing and reading skills. *Brain and Language, 88*, 21–25.

Ahissar, M. (2001). Perceptual training: A tool for both modifying the brain and exploring it. *Proceedings of the National Academy of Sciences of the United States of America, 98*, 11842–11843.

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

Amitay, S., Ahissar, M., & Nelken, I. (2002). Auditory processing deficits in reading disabled adults. *Journal of the Association for Research in Otolaryngology, 3*, 302–320.

Amitay, S., Hawkey, D. J. C., & Moore, D. R. (2004). Auditory frequency discrimination learning is affected by stimulus variability. *Perception and Psychophysics*, in press.

Bishop, D. V., Carlyon, R. P., Deeks, J. M., & Bishop, S. J. (1999). Auditory temporal processing impairment: neither necessary nor sufficient for causing language impairment in children. *Journal of Speech, Language and Hearing Research, 42*, 1295–1310.

Bradlow, A. R., Akahane-Yamada, R., Pisoni, D. B., & Tohkura, Y. (1999). Training Japanese listeners to identify English /r/ and /l/: Long-term retention of learning in perception and production. *Perception and Psychophysics, 61*, 977–985.

Brown, M., Irvine, D. R., & Park, V. N. (2004). Perceptual learning on an auditory frequency discrimination task by cats: Association with changes in primary auditory cortex. *Cerebral Cortex*, epub ahead of print.

Chermak, G. D., & Musiek, F. E. (1997). *Central auditory processing disorders*. San Diego: Singular.

Cohen, W., Hodson, A., O'Hare, A., Boyle, J., Durrani, T., McCartney, et al. (2004). Effects of computer based intervention using acoustically modified speech (Fast ForWord-Language) in receptive language impairment: Outcomes from a randomised control trial. *Journal of Speech, Language and Hearing Research*, in press.

Deary, I. J. (1994). Intelligence and auditory-discrimination—separating processing speed and fidelity of stimulus representation. *Intelligence, 18*, 189–213.

Fahle, M., & Poggio, T. (Eds.). (2002). *Perceptual learning*. Cambridge, MA, USA: MIT Press.

Frederickson, N., Frith, U., & Reason, R. (1997). Phonological assessment battery: Standardised edition. Available from: www.nfer-nelson.co.uk.

Gibson, E. J. (1967). *Principles of perceptual learning and development*. New York, USA: Appleton-Century-Crofts.

Gilbert, C. D. (1998). Adult cortical dynamics. *Physiological Review, 78*, 467–485.

Goldstone, R. L. (1998). Perceptual learning. *Annual Review of Psychology, 49*, 585–612.

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*, 1402–1415.

Hawkey, D. J. C., Amitay, S., & Moore, D. R. (2004). Early and rapid perceptual learning. *Nature Neuroscience, 7*, 1055–1056.

Hill, P. R. (2000). Auditory temporal processing in emotionally and behaviourally disturbed adolescents. Unpublished MSc thesis. University of Oxford.

Hook, P. E., Macaruso, P., & Jones, S. (2001). Efficacy of Fast ForWord training on facilitating acquisition of reading skills by children with reading difficulties—a longitudinal study. *Annals of Dyslexia, 51*, 75–96.

Hurford, D. P., & Sanders, R. E. (1990). Assessment and remediation of a phonemic discrimination deficit in reading disabled 2nd and 4th graders. *Journal of Experimental Child Psychology, 50*, 396–415.

Irvine, D. R., Martin, R. L., Klimkeit, E., & Smith, R. (2000). Specificity of perceptual learning in a frequency discrimination task. *Journal of Acoustical Society of America, 108*, 2964–2968.

Jamison, H. (2002). The effect of auditory training on the receptive language abilities of children. Unpublished MSc thesis. University of Oxford.

Kandel, E. R. (2001). The molecular biology of memory storage: A dialogue between genes and synapses. *Science, 294*, 1030–1038.

Karni, A., & Sagi, D. (1991). Where practice makes perfect in texture discrimination: Evidence for primary visual cortex plasticity. *Proceedings of the National Academy of Science of the United States of America, 88*, 4966–4970.

King, K., & Stephens, D. (1992). Auditory and psychological factors in 'auditory disability with normal hearing'. *Scandinavian Audiology, 21*, 109–114.

Kraus, N., McGee, T., Carrell, T. D., King, C., Tremblay, K., & Nicol, T. (1995). Central auditory system plasticity with speech discrimination training. *Journal of Cognitive Neuroscience, 7*, 25–32.

Kraus, N., McGee, T. J., Carrell, T. D., Zecker, S. G., Nicol, T. G., & Koch, D. B. (1996). Auditory neurophysiologic responses and discrimination deficits in children with learning problems. *Science, 273*, 971–973.

Kujala, T., Karma, K., Ceponiene, R., Belitz, S., Turkkila, P., Tervaniemi, M., et al. (2001). Plastic neural changes and reading improvement caused by audiovisual training in reading-impaired children. *Proceedings of the National Academy of Science of the United States of America, 98*, 10509–10514.

Leek, M. R. (2001). Adaptive procedures in psychophysical research. *Perception and Psychophysics, 63*, 1279–1292.

Lively, S. E., Pisoni, D. B., Yamada, R. A., Tohkura, Y., & Yamada, T. (1994). Training Japanese listeners to identify English /r/ and /l/. III. Long-term retention of new phonetic categories. *Journal of the Acoustical Society of America, 96*, 2076–2087.

McArthur, G. M., & Hogben, J. H. (2001). Auditory backward recognition masking in children with a specific language impairment and children with a specific reading disability. *Journal of Acoustical Society of America, 109*, 1092–1100.

Merzenich, M. M., Jenkins, W. M., Johnston, P., Schreiner, C., Miller, S. L., & Tallal, P. (1996). Temporal processing deficits of language-learning impaired children ameliorated by training. *Science, 271*, 77–81.

Mody, M., Studdert-Kennedy, M., & Brady, S. (1997). Speech perception deficits in poor readers: Auditory processing or phonological coding. *Journal of Experimental Child Psychology, 64*, 199–231.

Moore, D. R. (2001). Sensorimotor training and special education—can practice make perfect. *British Journal of Special Education, 28*, 138–141.

Moore, D. R., Amitay, S., & Hawkey, D. J. (2003). Auditory perceptual learning. *Learning and Memory, 10*, 83–85.

Morosan, D. E., & Jamieson, D. G. (1989). Evaluation of a technique for training new speech contrasts: Generalization across voices, but not word-position or task. *Journal of Speech, Language and Hearing Research, 32*, 501–511.

Nature Neuroscience (2004). Better reading through brain research. *Nature Neuroscience, 7*, 1.

Nozza, R. J. (1995). Estimating the contribution of non-sensory factors to infant–adult differences in behavioral thresholds. *Hearing Research, 91*, 72–78.

Polat, U., Ma-Naim, T., Belkin, M., & Sagi, D. (2004). Improving vision in adult amblyopia by perceptual learning. *Proceedings of the National Academy of Science of the United States of America, 101*, 6692–6697.

Ramus, F. (2003). Developmental dyslexia: Specific phonological deficit or general sensorimotor dysfunction?. *Current Opinion in Neurobiology, 13*, 212–218.

Ramus, F., Rosen, S., Dakin, S. C., Day, B. L., Castellote, J. M., White, S., et al. (2003). Theories of developmental dyslexia: Insights from a multiple case study of dyslexic adults. *Brain, 126*, 841–865.

Rosen, S. (2003). Auditory processing in dyslexia and specific language impairment: Is there a deficit? What is its nature? Does it explain anything?. *Journal of Phonetics, 31*, 509–527.

Rosen, S., & Manganari, E. (2001). Is there a relationship between speech and nonspeech auditory processing in children with dyslexia?. *Journal of Speech, Language and Hearing Research, 44*, 720–736.

Rosenberg, J. F., & Moore, D. R. (2003). Winning game (Auditory training in speech and language therapy: A field trial). *Bulletin of the Royal College Speech Language Therapy*, June Edition, 5–6.

Schwab, E. C., Nusbaum, H. C., & Pisoni, D. B. (1985). Some effects of training on the perception of synthetic speech. *Human Factors, 27*, 395–408.

Siegal, M., & Blades, M. (2003). Language and auditory processing in autism. *Trends in Cognitive Sciences, 7*, 378–380.

Sutcliffe, P. (2003) Auditory processing performance in young children: Attention is needed. Unpublished D.Phil., thesis, University of Oxford.

Tallal, P. (1980). Auditory temporal perception, phonics, and reading disabilities in children. *Brain and Language, 9*, 182–198.

Tallal, P., Miller, S. L., Bedi, G., Byma, G., Wang, X., Nagarajan, S. S., et al. (1996). Language comprehension in language-learning impaired children improved with acoustically modified speech. *Science, 271*, 81–84.

Tallal, P., & Piercy, M. (1973). Defects of non-verbal auditory perception in children with developmental aphasia. *Nature, 241*, 468–469.

Tallal, P., Stark, R. E., & Mellits, E. D. (1985). Identification of language-impaired children on the basis of rapid perception and production skills. *Brain and Language, 25*, 314–322.

Temple, E., Deutsch, G. K., Poldrack, R. A., Miller, S. L., Tallal, P., Merzenich, M. M., et al. (2003). Neural deficits in children with dyslexia ameliorated by behavioral remediation: Evidence from functional MRI. *Proceedings of the National Academy of Science of the United States of America, 100*, 2860–2865.

Thai-Van, H., Micheyl, C., Moore, B. C., & Collet, L. (2003). Enhanced frequency discrimination near the hearing loss cut-off: A consequence of central auditory plasticity induced by cochlear damage?. *Brain, 126*, 2235–2245.

Tremblay, K., Kraus, N., Carrell, T. D., & McGee, T. (1997). Central auditory system plasticity: Generalization to novel stimuli following listening training. *Journal of Acoustical Society of America, 102*, 3762–3773.

Tremblay, K., Kraus, N., & McGee, T. (1998). The time course of auditory percpetual learning: neurophysiological changes during speech–sound training. *NeuroReport, 9*, 3557–3560.

Werner, L. A., & Bargones, J. Y. (1991). Sources of auditory masking in infants: Distraction effects. *Perception and Psychophysics, 50*, 405–412.

Wightman, F., Allen, P., Dolan, T., Kistler, D., & Jamieson, D. (1989). Temporal resolution in children. *Child Development, 60*, 611–624.

Wise, B. W., Ring, J., & Olson, R. K. (1999). Training phonological awareness with and without explicit attention to articulation. *Journal of Experimental Child Psychology, 72*, 271–304.

Wise, B. W., Ring, J., & Olson, R. K. (2000). Individual differences in gains from computer-assisted remedial reading. *Journal of Experimental Child Psychology, 77*, 197–235.

Wright, B. A., Lombardino, L. J., King, W. M., Puranik, C. S., Leonard, C. M., & Merzenich, M. M. (1997). Deficits in auditory temporal and spectral resolution in language-impaired children. *Nature, 387*, 176–178.

Zwsilocki, J. J., Maire, F., Feldman, A. S., & Rubin, H. (1958). On the effect of practice and motivation on the threshold of audibility. *Journal of Acoustical Society of America, 30*, 254–262.