Reading and Writing: An Interdisciplinary Journal **16:** 573–609, 2003.

Language specific speech perception and the onset of reading

DENIS BURNHAM
MARCS Auditory Laboratories, University of Western Sydney, Sydney, Australia

Abstract. In two studies the relationship between the onset of reading and language specific speech perception, the degree to which native speech perception is superior to non-native speech perception, was investigated. In Experiment 1 with children of 4, 6, and 8 years, language specific speech perception occurred maximally at 6 years and was positively related to reading ability for age and language comprehension level. In Experiment 2, with an expanded range of ages and various stimulus and task changes, the relationship between reading and language specific speech perception still held, and maximal language specific speech perception occurred around the onset of reading instruction for three different sets of speech contrasts, but not for a control set of non-speech contrasts. The results show that language specific speech perception is a linguistic rather than an acoustic phenomenon. Results are discussed in terms of early speech perception abilities, experience with oral communication, cognitive ability, reading ability, alphabetic versus logographic languages, phonics versus whole word reading instruction, and the effect of age *versus* instruction.

Key words: Perceptual re-organisation, Phoneme awareness, Phonemic perception, Reading, Reading instruction, Speech perception

Introduction

Considerable research has been conducted on the relationship between reading ability and phonological processing abilities in children. Wagner and Torgesen (Wagner & Torgesen, 1987; Wagner, Torgesen, Laughton, Simmons & Rashotte, 1993; Wagner, Torgesen & Rashotte, 1994) identified three aspects of phonological processing. These are (a) phonological awareness, consisting of phonological analysis (phoneme segmentation and categorisation tasks) and phonological synthesis (phoneme blending); (b) retrieval of phonological codes from long-term store, either in isolated or serial speeded naming tasks; and (c) phonological recoding to maintain information in working memory (using tasks such as digit span and sentence repetition recall). All three are relatively stable over development (Wagner et al., 1993, 1994; Hansen & Bowey, 1994), and related to reading ability. There is evidence for (a) a causal relationship between phonological awareness and reading ability (Calfee, Lindamood & Lindamood, 1973; Bradley & Bryant, 1985; Goswami & Bryant, 1990; Wagner et al., 1994), and for a reciprocal causal relationship between phonological awareness and reading

(Morais, Cary, Alegria & Bertelson, 1979; Read, Zhang, Nie & Ding, 1986; Bradley & Bryant, 1978; Wagner & Torgesen, 1987, Wagner et al., 1994); (b) a positive relationship between children's phonological recoding in lexical access and their reading ability (Wolf, 1984; Wagner & Torgesen, 1987); and (c) a positive relationship between children's phonological recoding to maintain information in working memory and their reading ability (e.g., Mann & Liberman, 1984; Hansen & Bowey, 1994).

Are these three aspects of phonological processing different abilities or imperfect measures of the same underlying ability? Wagner and Torgeson (1987; Wagner et al., 1993) support a model in which one underlying ability accounts for phonological analysis and working memory performance, and another for phonological synthesis and performance on retrieval of phonological code tasks. The first is the most important source of individual differences in kindergarten children's phonological processing and is argued to have a phonological basis, whereby the quality of children's phonological representations affects their memory span performance (Wagner et al., 1993). This is thus a perceptual primacy hypothesis: if phonological representations are misaligned or degraded (due to inaccurate or imperfect speech perception mechanisms), then phonological awareness and memory span performance will be impaired (Wagner et al., 1993) as will, in turn, reading ability. To investigate this hypothesis it is important that the relationship between speech perception, phonological processing, and reading be investigated, and this will be done here.

Research on phonological processing and reading has mostly been conducted independently of speech perception research. In an exception to this McBride-Chang (1996) tested whether speech perception affects reading ability directly, or indirectly via its effect on phonological processing. Structural modeling with eight- to ten-year-old children revealed the most parsimonious model was one in which speech perception affects phonological awareness, which in turn affects word reading. Similar results have been found with reading disabled adults (Watson & Miller, 1993). Gibbs (1996) found similar but delayed effects at the onset of reading, with speech perception ability at five and six years predicting phonological awareness at six and seven years, respectively. Together these results support the contention that basic speech perception abilities are indeed implicated in reading difficulties via their effect on phonological analysis skills (Wagner et al., 1994).

A basic speech perception skill is categorical speech perception in which stimuli differing along physical continua such as formant transitions or voice onset time are perceived to belong to distinct categories (Eimas, Siqueland, Jusczyk & Vigorito, 1971; Burnham, Earnshaw & Clark, 1991; Aslin, Pisoni, Hennessy & Perey, 1981; Burnham, 1986). A number of studies have shown that children with reading difficulties produce less sharp category boundaries than their normal reading counterparts (Brandt & Rosen, 1980, but see Werker & Tees, 1987; Godfrey, Syrdal-Lasky, Millay & Knox, 1981; Hurford & Sanders, 1990; Manis, McBride-Chang, Seidenberg, Keating, Doi, Munson & Petersen, 1997; Joanisse, Manis, Keating & Seidenberg, 2000; Werker & Tees, 1987). Werker and Tees (1987) take such results to show that phonological categories are less robust in reading disabled than normal reading children, and that group differences are functionally significant and especially evident under stressful conditions (Werker & Tees, 1987; Brady, Shankweiler & Mann, 1983). However, as to the basis of this difference between disabled and normal reading, the underlying perceptual cause is as yet uncertain. Again we can see that the relationship between speech perception and reading begs further enquiry (Werker & Tees, 1987).

A possible perceptual factor underlying reading and reading disability is the ability to process rapid changes in acoustic stimuli. Children with specific language impairment have difficulty discriminating speech sound pairs distinguished by rapid frequency changes in formant transitions (consonants), but not pairs of steady-state (vowel) sounds (Tallal & Piercy, 1974; Reed, 1989), or in processing rapid *visual* stimuli (Reed, 1989). Tallal (1980; Tallal & Piercy, 1973a, b, 1974) suggests that the difficulty lies in discriminating the temporal sequence of rapidly-presented acoustic (not necessarily linguistic) stimuli, i.e., that there is a primary perceptual deficit. There are training studies to support this notion (Tallal, Miller, Bedi, Byma, Wang, Nagarajan, Schreiner, Jenkins & Merzenich, 1996; Merzenich, Jenkins, Johnston, Schreiner, Miller & Tallal, 1996), however, the specific benefit of training on non-speech as opposed to speech stimuli on reading is yet to be demonstrated, and there is opposing evidence that the difficulty is specifically with phonological contrasts rather than more general acoustic features (Mody, Studdert-Kennedy & Brady, 1997). So, it is by no means clear that the perceptual factor of which Tallal speaks is a primary acoustic factor as claimed, rather than a specifically linguistic factor.

Pointers to a perceptual basis for reading skills may be gleaned from considering developmental processes in speech perception. Young infants' astounding abilities to discriminate almost any phonologically-relevant or irrelevant phonetic contrast are well-documented (e.g., Eimas et al., 1971; Streeter, 1976; Best, McRoberts & Sithole, 1988, Burnham et al., 1991; see Burnham, 1986, and Goodman & Nusbaum, 1994 for reviews). This initial speech perception is then modified through processes of attunement (Aslin, Pisoni, Hennessy & Perey, 1981), prototype abstraction for native language phones (Kuhl, 1991; Kuhl, Williams, Lacerda, Stevens & Lindblom, 1992), and especially attenuation for non-native phones (Burnham et al., 1991;

Burnham, 1986; Tees & Werker, 1984; Werker, Gilbert, Humphrey & Tees, 1981; Werker & Tees, 1983, 1984a; Werker & Lalonde, 1988; Werker, 1994). These processes result in the perceptual reorganisation of initial phonetic classes into language specific phonological categories around 6 to 9 months (Werker, 1994).

Based on such evidence Werker and Tees (1987) suggest that infants' initial phonetic sensitivities are the building blocks of phonological categories, and that these categories are then sharpened by experience with oral communication. In line with this, an alternative model to Tallal's (1980) primary perceptual deficit is that there is a secondary perceptual deficit based on subtle differences in the efficiency and accuracy with which children reorganise their speech perception in accord with the specific phonology of their language environment.

Additionally we have obtained evidence that phonological categories may also be sharpened by experience with *written* language. Burnham (1986; Burnham, et al., 1991) investigated English language environment listeners' ability to perceive two synthetic stop contrasts; one phonologically *relevant*, and another phonologically *irrelevant*. Using the Infant Speech Identification method (Burnham, Earnshaw & Quinn, 1987) perceptual ability for each of the two contrasts was tested at 10 months, $2^1/_2$ years, $6^1/_2$ years, and in adults. Analysis of category boundary sharpness showed perception of the phonologically-relevant [p]–[p^h] contrast, perceived as /b/ and /p/ by English speakers, became more categorical with increasing age, especially between two and six years. However, for the phonologically-irrelevant contrast, [b]–[p], both perceived as /b/ by English speakers, boundary sharpness improved between infancy and two years, reverted to chance level at six years, and then improved slightly by adulthood.

Burnham (1986; Burnham et al., 1991) speculated that the intensification of language specific speech perception between two and six years is related to the onset of reading instruction. Encouragement to process language in terms of separate phonemes, and instruction in phoneme-to-grapheme mapping would, it was reasoned, promote speech perception bias *towards* phoneme classes of the ambient language and *away from* phonologically irrelevant phonetic differences (e.g., [b] *versus* [p] for English language environment listeners). This language specific speech perception, as it may be called, would facilitate the phoneme-to-grapheme mapping process, one of the most difficult tasks in reading acquisition. Based on this reasoning, we may hypothesise that around school age, the degree of language specific speech perception will be positively related to children's reading ability. The Burnham et al. (1991) results also show significant improvement for the non-native contrast by adulthood; presumably as a result of reading becoming more automatic, fine-grained phonetic discriminations again are possible.

The two experiments reported here were conducted to investigate this proposed relationship between language specific speech perception and reading. Both experiments concern young children's perception of phonologically-relevant and phonologically-irrelevant speech contrasts, and their performance on reading-related tests. Four questions are investigated.

1. What is the developmental locus of the intensification of language specific speech perception observed by Burnham et al. (1991)? Does it occur specifically at six years and does it abate at some later stage?
2. Is language specific speech perception related to reading acquisition?
3. Is the language specific speech perception phenomenon robust – is improvement for the native [p]–[p^h] and attenuation for the non-native [b]–[p] contrast also evident for other native/non-native speech contrasts, and in other test situations?
4. Is language specific speech perception a specifically linguistic phenomenon or can it be explained by more general auditory processes?

EXPERIMENT 1

The two aims of Experiment 1 are to define the developmental locus of the intensification of language specific speech perception, and to investigate whether it is related to reading ability. Burnham et al. (1991) found a trough in the perception of non-native speech sounds at six years compared with two years and adulthood. As there were wide differences between ages, the locus could not be specified. Here four-, six-, and eight-year-old English language environment children were tested on each of the two contrasts used in our earlier study, the native [p]–[p^h] and the non-native [b]–[p] (Burnham et al., 1991).

To investigate the relationship with reading ability, children were also given an articulation, a language comprehension, and for the six- and eight-year-olds, a reading test. It was expected that language specific speech perception (given by the perceptual ability for native (N) minus the ability for non-native (NN) speech contrasts) should occur maximally at six years of age, and that children's degree of language specific speech perception should be positively related to their reading ability.

Method

Subjects

Thirty-six monolingual native English-speaking predominantly Anglo background children were tested, 12 four-year-olds (6 boys, 6 girls; $\overline{X}$ = 4 years 8.01 months, 4–8.01 years hereafter; range = 4–5 to 4–10.9 years), 12 six-year-olds (6 boys, 6 girls; $\overline{X}$ = 6–7.98 years; range = 6–5.8 to 6–10.4 years), and 12 eight-year-olds (5 boys, 7 girls; $\overline{X}$ = 8–8.27 years; range = 8–5.7 to 8–11 years). (An additional 2 four-year-olds were tested but failed to reach criterion on the non-native speech discrimination test. All six- and eight-year-olds completed all tests.) The four-year-old group attended pre-schools and the six- and eight-year-olds attended schools near the university. The schools and preschool in the area surrounding the university were in a predominantly middle class area. The school reading curriculum used programs that were oriented towards a whole word method but with a degree of phonics instruction incorporated into the program.

Apparatus and materials

Speech perception test
Apparatus for the speech perception test is described in detail in Burnham et al. (1991), and more briefly here. Participants sat facing a large screen painted blue to the left and orange to the right of their midline. The screen housed central attention-getting lights, a central audio speaker through which sounds were presented, and two lateral reinforcement windows each of which could be illuminated to reveal animated toy rabbits. In front of the child were two paddles side-by-side, color-coded to match the two sides of the screen. When one of the two training sounds was presented centrally, pressing the paddle on the side pre-assigned for that sound resulted in reinforcement delivery from the window on that side, and similarly for the other sound and side. Incorrect paddle presses were not reinforced.

Stimulus sounds were synthesised bilabial stop consonants plus the vowel [a:]. For the native contrast condition the training sounds were synthesised bilabial stop consonants of 0 msec voice onset time (VOT), [pa:], and +70 msec VOT, [p^ha:] (/ba/ and /pa/ in English), and the generalisation test sounds were these endpoints plus the six intervening sounds at 10 msec VOT intervals, i.e., bilabial stops with VOTs of 0, 10, 20, 30, 40, 50, 60, and 70 msec. For the non-native contrast condition the training sounds were synthesised bilabial stops of 0 msec VOT, [pa:], and –70 msec VOT, [ba:] (both /ba/ in English), and the generalisation test sounds were these plus the six intervening sounds at 10 msec VOT intervals, i.e., bilabial stops with VOTs of 0, –10, –20, –30, –40, –50, –60, and –70 msec. Sounds were played on an Otari MX5050 reel-to-reel tape recorder through an Acoustics Research BXi speaker. Each stimulus syllable was 500 msec in duration, and in each trial 10 repetitions of each syllable were presented with 250 msecs of silence between repetitions. Thus each trial was of 7.5 secs duration.

Language tests

The Queensland Articulation Test (Kilminister & Laird, 1978), the Bankson Language Screening Test (Bankson, 1977), and the Daniels and Diack Reading Test Number 1 (Daniels & Diack, 1974) were used. The Queensland Articulation Test (QAT hereafter) is a 64-item Australian language test similar to the Edinburgh Articulation Test, in which subjects are asked to name pictured objects. The items are designed such that all English language consonants are tested in initial, medial and final positions with appropriate phonotactic constraints ([h] initial only, [ŋ] medial and final only, [w], [ɹ], and [j] initial and medial only, and [ʒ] medial only). Productions are scored as correct or incorrect with a possible total of 64. Children were tested on a subset of the 39 most discriminating items of the Bankson Language Screening Test (Bankson hereafter), including items to test syntactic, semantic, and pragmatic understanding (Bankson, 1977). Each item was scored as correct or incorrect with a possible total of 39. Six- and eight-year-olds, but not four-year-olds, were also given the Daniels and Diack Reading Test Number 1 (Reading hereafter), consisting of 36 items which become progressively more difficult. Children are required to read the items aloud and continue until they make errors on three consecutive items. Items 1 to 26 require a yes/no answer, the answer being recorded simply to maintain the child's interest, and each item is awarded 3 points if no reading errors occur. The remaining items are questions with four multiple-choice answers, with one point given if the question is read correctly and one for each of the 4 alternatives read correctly. The maximum possible score is 128.

Procedure

All subjects were given the native contrast perception test, the non-native contrast perception test, the QAT, and the Bankson. The six- and eight-year-olds were also given the Reading test. Subjects were usually tested first on one of the perception tests, then the language tests, and then the other perception test. This order was sometimes modified if the child found the perception test initially daunting. Despite such deviations order of presentation of the native and non-native perception tests was counterbalanced between subjects, as was order within the language tests. The procedure for the QAT, the Bankson, and Reading followed the prescribed published procedures (Kilminister & Laird, 1978; Bankson, 1977; Daniels & Diack, 1974).

The procedure for the speech perception tests was identical for the native and non-native contrasts, except that the stimuli, and the random presentation orders of test stimuli differed. Children sat in front of the screen, were told they would hear one of two sounds, and instructed to press one paddle for one sound and the other for the other sound. For example, the child was told

"When you hear the 'ba' sound press this blue paddle here (*on the left*) and when you hear the 'pa' sound press this orange paddle here (*on the right*)". Similar instructions were given for the non-native contrast with the prevoicing in the –70 msec stimulus being pronounced "mba". The consequences of pressing the correct paddle (reinforcement by animated bunny) were demonstrated. Children were told to use only one hand to press the paddles, and instructed to keep the agreed-upon hand on the table on an appropriately-sized red handprint until the stimulus sound began, with the other hand kept under the table. One experimenter stayed next to the child for the initial part of training. For the six- and eight-year-olds the experimenter then retired but for the four-year-olds the experimenter remained to monitor performance and maintain the child's interest.

In the initial period, reinforced training with randomly-presented end point sounds continued until at least three unprompted correct paddle presses to each side occurred. Then any three consecutive correct responses to the randomly-presented endpoints was followed by an unreinforced generalisation test trial in which one of the eight possible test sounds was presented. After each test trial children returned to reinforced training until a further three consecutive correct responses occurred. Reinforcement was given for correct training trials but no reinforcement was given in test trials. Testing continued until all eight test trials had been presented once.[1]

The time taken to complete the tests varied with age. The four-year-olds required two testing sessions totalling around 1.5 hours. This included short breaks if required by the child. The six- and eight-year-olds completed the tests in a single session of between 50 and 75 minutes.

Results

Speech perception tests

Each of the eight native contrast test sound responses were scored as appropriate for the 0 msec VOT or the +70 msec VOT endpoint; and the 8 non-native contrast test sound responses as appropriate for the 0 msec or the –70 msec VOT endpoint. Group identification curves are shown in Figure 1. The native contrast identification functions appear to become progressively steeper over age, whereas those for the non-native contrast appear to slope in the appropriate direction for the four-year-olds, give way to a flatter function at 6 years, followed by a more appropriately sloped curve at 8 years.

Individuals' patterns of responses on the native and on the non-native contrast were converted to categorical perception scores via a method, which is affective with small numbers of test responses, unlike methods such as d' or

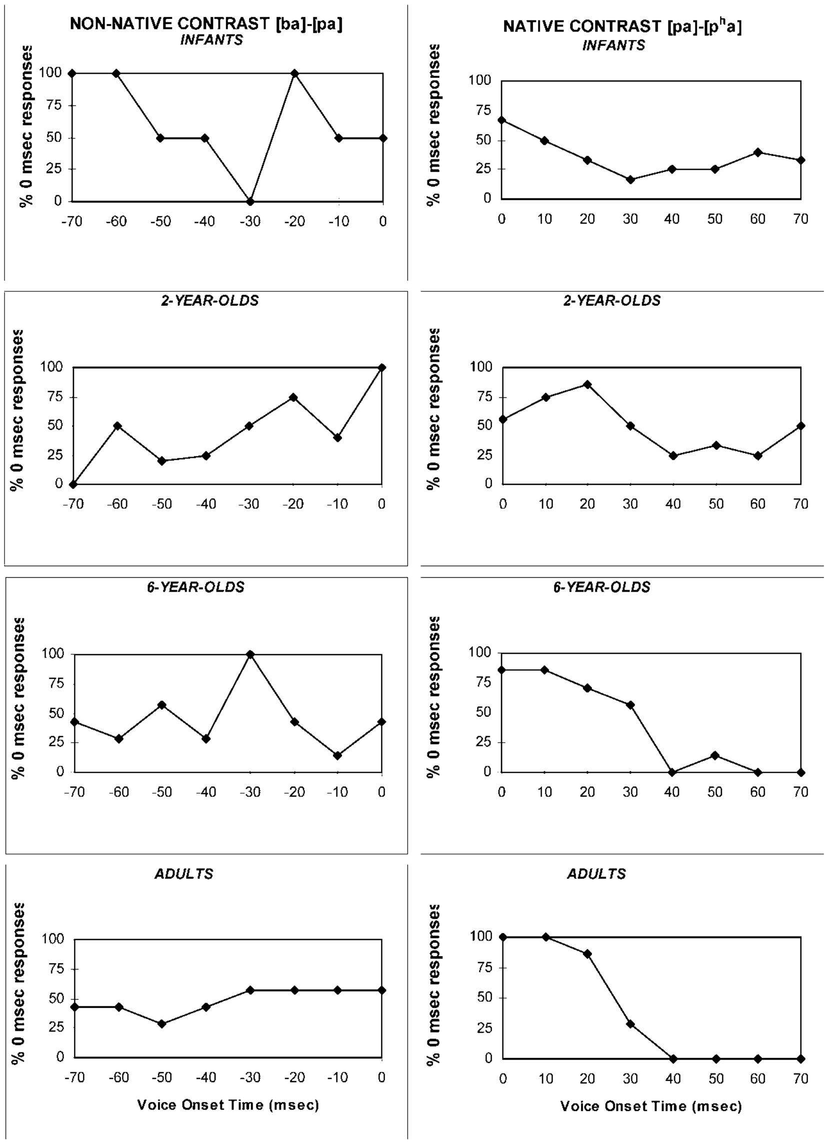

Figure 1. Group identification functions for English language environment four-year-olds, six-year-olds, and eight-year-olds for native, [pa]–[pʰa], and non-native, [ba]–[pa], speech contrasts in Experiment 1.

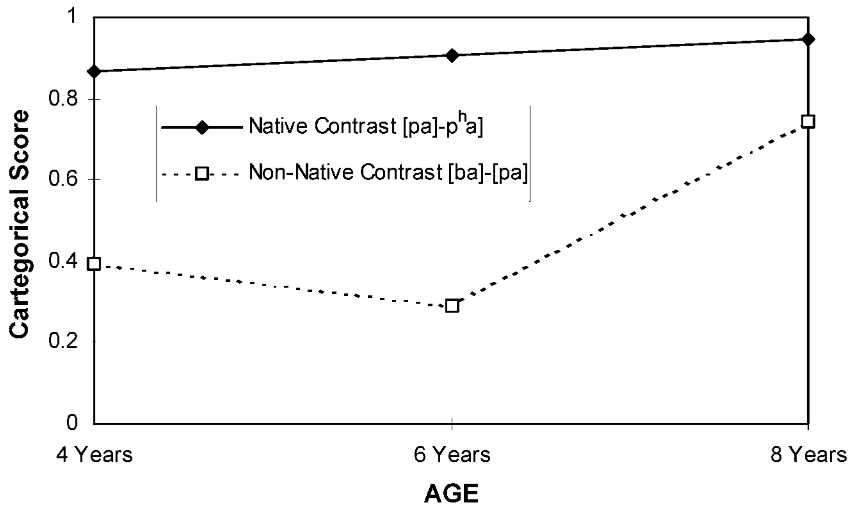

Figure 2. Mean categorical perception scores for English language environment four-year-olds, six-year-olds, and eight-year-olds for native, [pa]–[p^ha], and non-native, [ba]–[pa], speech contrasts. Chance level = 0; perfect categoricality = 1.

logistic functions (Burnham et al., 1991). To derive these categorical scores, each individual's pattern of responses for the 8 test trials are compared with 7 ideal categorical curves constructed with sharp crossovers between the two response types ("ba" *versus* "pa" for the native, and "ba" *versus* "mba" for the non-native speech perception tests). These 7 curves use the narrowest possible VOT width here (10 msec), with category crossover centered on 5, 15, 25, 35, 45, 55, or 65 msec VOT (positive for the native, and negative for the non-native contrast). Once the best fit is found, then the closeness of fit is determined by a method similar to testing planned contrasts in an analysis of variance. The resultant degree of fit, between –1 and +1, measures the sharpness of the individual's pattern of responses, and provides individual categorical scores for each participant to be used in group analyses.

Mean categorical scores are plotted in Figure 2. An age × speech contrast, 3 × (2) analysis of variance with repeated measures on speech contrast revealed significantly higher categorical perception scores on the native ($\overline{X}$ = 0.91) than the non-native ($\overline{X}$ = 0.48) contrasts, $F(1,33) = 24.00$, Mean Square Error (MSE) = 0.15. The overall speech contrast × age effect was not significant, nor was speech contrast × linear age trend, $F(1,33) = 1.46$, $P > 0.05$, MSE = 0.15, or the speech contrast × quadratic age trend, $F(1,33) = 3.06$, $P < 0.10$, MSE = 0.15, despite the suggestion in Figure 2 of a quadratic U-shaped component centred on 6 years for the non-native native contrast.

Table 1. Mean percentage correct on Queensland Articulation Test (QAT), Bankson Language Comprehension Test (Bankson), and Daniels and Diack Reading Test (Reading) in Experiment 1.

Test	Age		
	4 Years	6 Years	8 Years
QAT	89.97	97.53	98.57
Bankson	56.84	92.11	97.15
Reading	—	57.16	96.10

Language tests

Percent correct scores for QAT, Bankson, and Reading are shown in Table 1. Separate analyses of variance revealed there was a significant linear improvement over age for the QAT, $F(1,33) = 18.44$, MSE = 24.03, and a significant linear and quadratic improvement over age for the Bankson, $F(1,33) = 56.77$ & 10.64, MSE = 171.69. For Reading, eight-year-olds were significantly better than six-year-olds, $F(1,22) = 24.91$, MSE = 365.18.

Relationship between speech perception and language tests

Burnham et al. (1991) found that at 6 *versus* 2 years of age there was a significant reduction of scores for a non-native speech contrast, despite continued improvement for a similar native contrast, and that this difference decreased by adulthood. In that experiment native and non-native contrasts was a between subject factor, and here it was deliberately made within subjects, to allow correlations and regression analyses. Each subject's non-native contrast categorical score was subtracted from their native categorical score, (N-NN), as a measure of language specific speech perception. Initial exploratory correlations of the native, non-native, and (N-NN) scores with age and QAT, Bankson, and Reading scores revealed an interesting complex of results. Age was positively correlated with both native and non-native scores ($r = 0.20$, 0.21, respectively) but *negatively* correlated with (N-NN) scores ($r = -0.15$). On the other hand all the language tests were positively correlated with native, negatively correlated with non-native, and positively correlated with (N-NN) scores ($r = 0.08$, –0.01, and 0.03 respectively for QAT; 0.11, –0.17, and 0.18 for Bankson; and 0.29, –0.10, and 0.16 for Reading). While the magnitude of some of these is negligible, the *pattern* of results suggests that N-NN scores may be tapping something different than either native or non-native scores alone. This is not inconsistent with our notion that N-NN should measure language specific speech perception, independently of native and non-native scores. Accordingly, the scores for the three language tests and subjects' exact ages were regressed onto the N-NN language specific speech perception scores. Two separate analyses were conducted, one with all subjects in which the reading scores were omitted, and one for just the six- and eight-year-olds in which reading scores were included.

The regression with all subjects was significant $R^2 = 0.29$, $F(3,32) = 4.42$, $P < 0.01$, MSE = 0.27, and the regression equation was *(N-NN) = 0.48 QAT + 0.38 Bankson – 0.79 Age*. F to enter or remove was significant for the QAT, $F(1,32) = 7.17$, $P = 0.01$ and Age, $F(1,32) = 10.82$, $P = 0.002$, but not the Bankson, $F(1,32) = 2.91$, $P = 0.01$. Dividing each coefficient by the highest coefficient and rounding the resultant coefficient to +1, 0 or –1 is a way to simplify the regression equation to facilitate meaningful verbal descriptions (DeSarbo, Hausman, Shen & Thompson, 1982; Harris, 1985). Using this method the resultant equation is (N-NN) = QAT – Age. This, along with the significance tests shows that language specific speech perception is highest in those participants with relatively good articulation ability for their age.

The regression which included Reading scores conducted with just the six- and eight-year-olds was significant, $R^2 = 0.53$, $F(4,19) = 5.34$, $P < 0.005$, MSE = 0.216, and the regression equation was *(N-NN) = 0.88 Reading + 0.23 QAT – 0.90 Age – 0.49 Bankson*. *F* to enter or remove were significant for Reading, $F(1,19) = 9.02$, $P = 0.01$, Age $F(1,19) = 13.34$, $P = 0.002$, and Bankson $F(1,19) = 4.36$, $P = 0.05$, but not for QAT, $F(1,19) = 1.95$, $P = 0.18$. The simplified regression equation, (N-NN) = Reading – Age – Bankson, and the significance tests can be interpreted to show that language specific speech perception is highest in those participants with relatively good reading ability for their age and level of language comprehension.

Discussion

Burnham et al. (1991) showed that there is a period of increased language specific speech perception some time between 2 years and adulthood. The speech perception results here suggest that this occurs specifically around 6 years. In addition, the results are consistent with the notion that language specific speech perception subsides around 8 years; while categorical perception scores for the native contrast continued to increase from 4 to 6 to 8 years, there was a U-shaped trend for the non-native contrast suggesting that categorical speech perception of the non-native contrast was worse at 6 than 4 or 8 years. However, this should be interpreted cautiously, as while a similar quadratic trend component was significant in Burnham et al. (1991), it was not here. This effect here may have been ameliorated for two possible reasons. Firstly, the categorical scores here were generally high (see Figure 2) relative to those in Burnham et al. (1991; see Figure 2 there), possibly because in the earlier study children were simply told they would hear two sounds and left to deduce the sound/button-press contingencies themselves, whereas here more specific phonologically-relevant labels for the speech sounds were given to reduce test duration, and allow the other language tests to be completed. These generally higher scores here may have reduced any potential quadratic trend. A second possibility is that the ages 4, 6, and 8 years represent the trough of a broader dip in non-native speech perception performance, and that if younger and older ages are included this trend may become more evident. To investigate this, finer gradations and a wider range of ages are used in Experiment 2, and power is increased by more than doubling subject numbers in each age group. So the first aim of this study has been only partly realised – it appears that the locus of the intensification of post-infancy language specific speech perception occurs around 6 years, i.e., around the onset of reading, but further evidence is required.

The regression analysis with the four-, six-, and eight-year-olds showed that language specific speech perception is positively related to children's level of articulation ability for age. This makes good intuitive sense: if particular children are articulating well for their age then it should follow that they have greater perceptual bias towards the speech sounds present in their own language. The direction of causality between speech perception and speech production, or whether causality is uni- or bi-directional (Werker, 1993) cannot be determined here, but it is notable that very few instances of a relationship between children's speech perception and speech production skills have been reported previously (Zlatin & Koenigsknecht, 1976; Strange & Broen, 1981). The reason that we found a perception-production relationship here may be that we used language specific perceptual bias (N-NN) scores based on two different speech contrasts, rather than perceptual ability on a single task. Further study of the relationship between language specificity in speech perception and speech production is warranted.

The regression analysis with just the six- and eight-year-old children, and with reading added, showed that children who are good readers for their age and level of language comprehension are those who manifest greater language specific speech perception. (Articulation ability was no longer a significant predictor, suggesting that after the onset of reading instruction, reading is a more powerful predictor of language specific speech perception than is articulation.) Again the regression solution makes sense intuitively. Children who have high language specific speech perception should have a good grasp of the phoneme-to-grapheme conversion rules involved in reading, and should be able to apply these rules to unfamiliar words. In addition, such children should have a good grasp of which phonetic variations signal phonological differences in their language and which are simply allophones of a native language phoneme class. Again the direction of causality cannot be determined here, though either causal direction would appear intuitively feasible.

EXPERIMENT 2

Experiment 2 has four aims, all related to the language specific speech perception found in Experiment 1. To investigate the *developmental locus* of the intensification of language specific speech perception in a more fine-grained manner, a wider range of ages was included. Children at the beginning of their first year of school (called ‘Kindergarten’ year in Sydney schools), children at the end of their Kindergarten year, and children in subsequent years, Grades 1, 2, 3, and 4, as well as a control group of adults were tested. Power is also greatly improved with 28 subjects per group here compared with 12 per group in Experiment 1.

To investigate the *relationship between language specific speech perception and reading*, measures of phonological segmentation and reading ability were included.

Regarding the *nature of language specific speech perception* it is unclear from the results of Experiment 1 whether language specific speech perception involves a bias towards native as opposed to non-native speech sounds, or a bias towards *familiar* (native) sounds and away from all *unfamiliar* sounds, be they non-native speech or even non-speech. To evaluate this, non-speech analogues of voicing contrasts, tone onset time (TOT) stimuli, were included. If performance on non-native contrasts were to be attenuated for six-year-olds *below* the level of these non-speech TOT sounds then the bias could be said to be specifically phonological.

To investigate the *generalisability* of the results of Experiment 1, three changes were made. Firstly, naturally-produced as opposed to synthetic syllables were used. Secondly, perception of two additional speech contrasts was tested. And thirdly, a different method, discrimination rather than identification was used. If such changes do not affect the basic relationship between speech perception and reading ability found in Experiment 1, then we can conclude that this relationship is robust.

Method

Subjects

A total of 196 monolingual English speakers from the same pool as in Experiment 1 were tested, six groups of 28 children and one group of 28 adults, with 14 males and 14 females in each group. Children were tested in schools close to the university. The youngest group of children (Young Kindergarten; $\overline{X}$ age = 5–4.46 years, range = 4–2 to 5–10 years) were in their first year of schooling and tested in April, close to the start of the school year. The Australian school year begins in early February, so these children had approximately 2 months of schooling at the time of testing. The second (Kindergarten) group of children were also in the preparatory kindergarten class but were tested late the previous school year (December). Allowing for school holidays, they had had approximately 8 months of schooling, and their $\overline{X}$ age was 5–9.75 years (range = 5–0 to 6–8 years). The remaining four child groups were taken from Grades 1, 2, 3 and 4 and were tested in the first half of the school year, giving them approximately 1–3, 2–3, 3–3, and 4–3 years of schooling respectively. Their mean ages were respectively 6–6.54 years (range = 5–8 to 7–4 years); 7–6.43 years (range = 6–10 to 8–5 years), 8–5.96 years (range = 7–8 to 9–3 years) and 9–5.93 years (range = 8–7 to 10–4 years). The adults were university students (with at least 13 years of schooling), and their mean age was 23–7.44 years (range = 17 to 41 years; sd = 6.11 months).

Apparatus and materials

Children were given speech perception, phoneme segmentation, and reading tests.

Speech perception tests
An AX discrimination paradigm, in which pairs of sounds were presented for same/different judgements, was used. To ensure that children understood the meanings of "same" and "different" visual training materials, a card with two identical side-by-side red squares ("same") and a card with a red square and a blue square side-by-side ("different"), and cards of nonsense animals which differed in the number of legs, type of neck and patterning on the body were used. Some animal cards showed two identical animals ("same"), others showed very different pairs ("different"), and yet others showed slightly different pairs ("just a little bit different but still different").

There were five blocks of trials, and in all five the interstimulus interval (ISI) between the offset of sound 1 and the onset of sound 2 was 500-msec.

Participants were required to report whether the two sounds were the "same" or "different". Stimulus sounds were recorded from a master tape on an Otari MX5050 reel-to-reel tape recorder with Ampex 631 tape onto TDK SA (Type II) cassette tapes. A Marantz portable cassette recorder was used to present the sounds.

There were three blocks of speech sound stimuli, one block of non-speech stimuli, and an auditory post-test block, always presented last. The three speech sound blocks consisted of natural speech tokens of consonants paired with the [a:] vowel. Each of these three speech sound blocks contained two speech contrasts: one which was phonologically-relevant for English language environment listeners (produced by a male native English speaker) and one which was phonologically irrelevant for English language listeners (produced by a male native Hindi speaker). The productions of the English and Hindi speaker were selected and edited to ensure that intonation contour, pitch, and duration of the tokens were perceptually equivalent via initial screening by a panel of experienced judges, and further acoustic screening with a speech editor. Two different tokens of each phone were used to ensure acoustic variability, while maintaining phonetic identity. Sounds used in the five blocks are described below and acoustic and phonetic details of native and non-native sounds are given in Appendix 1A, 1B, and 1C for blocks A, B, and C, respectively.

Block A – bilabial voicing contrasts. Native – voiceless unaspirated [pa:] vs voiceless aspirated [p^{h}a:] bilabial stops: This contrast is phonologically-relevant in both Hindi and English. The two sounds are categorised by English speakers as the phonemes /b/ and /p/. *Non-native – voiced [ba:] vs voiceless [pa:] unaspirated bilabial stops:* This contrast is relevant in Hindi but English listeners categorise both as /b/.

Block B – alveolar/dental voicing contrasts. Native – voiceless unaspirated [ta:] vs voiceless aspirated [t^{h}a:] alveolar/dental stops: This contrast is phonologically relevant in both Hindi and English. The sounds are categorised as /d/ and /t/ in English. *Non-native – breathy prevoiced [d^{h}a:] vs aspirated voiceless [t^{h}a:] alveolar/dental stops:* This contrast is phonologically relevant in Hindi but in English both sounds tend to be categorised as /t/.

Block C – place contrasts. Native – bilabial [p^{h}a:] vs alveolar/dental [t^{h}a:] voiceless stops: This contrast is phonologically-relevant in both Hindi and English, being represented in the latter by the phonemes /p/ and /t/. *Non-native – unvoiced unaspirated dental [t̪a:] vs unvoiced unaspirated retroflex* *[t̪̊a:] stops:* This Hindi contrast is not phonologically relevant in English, with both tending to be categorised by English language listeners as /d/.

Block D – non-speech (TOT) analogues of voicing contrasts. This block consisted of tone onset time (TOT) stimuli, considered to be non-speech analogues of VOT contrasts because they are composed of a high frequency (1500 Hz) component modelled on the high frequency broad-band burst in stop consonants and a low frequency (500 Hz) component modelled on the low frequency voice bar in stop consonants (Pisoni, 1977); and because they tend to be perceived categorically by adults (Pisoni, 1977) and infants (Jusczyk, Pisoni, Walley & Murray, 1980).[2] *Positive TOT Contrast, 0 msec vs +70 msec TOT:* These sounds are used as analogues of voiceless unaspirated 0 msec and voiceless aspirated +70 msec VOT stop consonant contrasts (see native voicing contrasts, Block A and B above). For the 0 msec stimulus, onset of the two sinusoidal tones, one of 500 Hz, and another 1500 Hz, was simultaneous and each lasted for 245 msec. For the +70 msec stimulus, the duration of the 500 Hz tone was 245 msec, and its onset preceded the 175 msec 1500 Hz tone by 70 msec. In both stimulus sounds the offset of the two tones was simultaneous, with the total duration of each stimulus being 245 msec. *Negative TOT Contrast, 0 msec vs –70 msec TOT:* These sounds are used as analogues of 0 msec voiceless unaspirated and –70 msec voiced unaspirated VOT stops (see non-native voicing contrasts in Blocks A, B above). The 0 msec stimulus was described above. For the –70 msec stimulus, the duration of the low frequency 500 Hz tone was 315 msec and its onset preceded that of the 245 msec 1500 Hz tone by 70 msec, so the total duration of this sound was 315 msec.

Block structure. In blocks A, B, C and D, 16 AX pairs were presented. For each pair the ISI was 500 msec, employed to ensure that subjects processed sounds at a phonetic level (Werker & Logan, 1985; Werker & Tees, 1984b). Of the 16 pairs in each block, 8 were of the appropriate native contrast and 8 were of the appropriate non-native contrast (positive and negative TOT contrasts respectively for Block D). Within contrasts half (4) were pairs of the same phone (or TOTs) and half (4) were pairs of different phones (or TOTs). For the same pairs half (2) were the same phones (or TOTs) for one member of the contrast (AA pairs) and half (2) were the same phones (or TOTs) for the other member of the contrast (BB pairs). For "same" pairs in Blocks A, B, and C *different* tokens of the same phone were used to ensure acoustic but not phonetic variability. For the different pairs half (2) were presented in each order (AB or BA) and within each order the tokens of each phone type varied across trials.

Auditory post-test. To ensure that children were applying the same/different labels correctly and that their hearing was adequate, a series of eight AX pairs of harmonic tones was presented at the end of the four test blocks. There was a high harmonic tone (based on 1500 Hz), and a low harmonic tone (based on 500 Hz). Four of the AX trials were same stimuli (two low-low and two high-high pairs) and 4 were different (high-low pairs with counterbalanced order).

Phoneme segmentation test

The phoneme segmentation test consisted of 24 items in which the targeted phoneme occurred either initially or finally, and as a singleton or in a cluster, such that there were 6 items in each initial/final × singleton/cluster subgroup. Two versions of the test were given: a receptive deletion task and a productive identification task. In the former, children were told for example, "Say 'ice' without the /s/ sound". In the latter they were asked, "What sound is in 'ice' but not in 'eye'?". Six practice items were given.

Reading test

The Daniels and Diack Reading Test 1 was administered to all the children (see 'Apparatus and materials' in Experiment 1).

Procedure

Each child was tested individually in a quiet room at their school. Children sat on a small chair at a table facing the experimenter. Each child was the given speech perception, phoneme segmentation, and reading tests with order of presentation counterbalanced between subjects. Children mostly completed all tests in one sitting ranging between 40 and 50 minutes. Breaks were given when the child required them and one experimenter gave the speech tests and another the reading and phoneme segmentation tests to provide some variety. The adults were only given the speech perception tests.

Speech perception test

Training with the visual same/different training materials always occurred first. Children were shown one card with two red squares and one with a blue and red square and asked to identify which was "same" and "different". These were then put face up on the table, one on each side of the child. Children were then asked to place pairs of nonsense animals in the "same" or the "different" pile. When the child categorized the animals correctly they then proceeded to the four auditory blocks with order of block presentation counterbalanced between subjects. At the end of each block two more nonsense animal pairs (one same, one different) were presented to ensure continued understanding. Following the fourth block, the auditory post-test was given. Subjects were excluded from later analysis if they did not score at least 6 correct on the 8 items in the auditory post-test. Adults were tested in a similar fashion except that the same/different training materials were not used.

Phoneme segmentation test

Children were first presented with the six practice items, and did not proceed to the test items until they had responded correctly to at least 5 of these, and/or until the experimenter considered that the child understood the task. Half of the children in each age group were given the receptive deletion form and half were given the productive identification form of the task.

Reading test

Children were asked to read the items aloud and testing continued until errors occurred on three consecutive items (see 'Apparatus and materials', Experiment 1).

Intelligence

For the purpose of accounting for variability in regression analyses, each child was given an estimate of general intelligence by their teacher. This was a simple coarse-grained measure of "Above Average", "Average" or "Below Average", and these were converted to scores of 3, 2 and 1 respectively. More detailed intelligence tests could not be given due to the policy of the schools that participated in the study.

Scoring

For the speech perception test discriminative index (DI) scores were calculated, given by DI = (Correct 'Different' responses on Different Trials "Hits" – Incorrect 'Different' responses on Same Trials "False Positives")/(Number of Different (AB + BA) trials), with a denominator of 4 for all contrasts. Eight such DI scores were calculated for each subject: a native and the non-native score for Blocks A, B, and C, and an index each for the positive TOT contrast and the negative TOT contrast in Block D. The maximum DI score is 1 and chance level is 0.[3]

Items in the phoneme segmentation test were scored as correct or incorrect with a possible maximum of 24. For the reading test the scoring protocol given by Daniels and Diack (1974), with a maximum possible score of 128 was used. Both phoneme segmentation and reading scores were converted to percentages.

Results

Speech perception tests

Preliminary analysis revealed that there were no significant differences in DI scores as a function of order of block presentation, $F(3,168) = 0.52$, MSE = 1.133. Mean DIs for each of the four sets of two contrasts are plotted against age in Figures 3a, b, c, and d, and were analysed by planned contrasts in an age × (speech contrasts × native/non-native), 7 × (4 × 2) analysis of variance with repeated measures on the speech contrast and native/non-native factors.

Bilabial stop voicing contrasts. For the bilabial stop voicing contrasts the native contrast was more easily discriminated than the non-native contrast, $F(1,189) = 750.18$, MSE = 0.06 ($\overline{X}_{\text{native}} = 0.91$, $\overline{X}_{\text{non-native}} = 0.24$). There was a significant interaction of native versus non-native contrasts with the linear age trend, $F(1,189) = 24.73$, and of native versus non-native contrasts with the quadratic age trend, $F(1,189) = 22.99$, MSE = 0.06. These significant interactions in conjunction with inspection of Figure 3a indicate that the age trend for the non-native contrast was more linear and quadratic than for the native contrast, with a trough centred on the Grade 1 children. These results are similar to the age trends with the synthetic versions of these contrasts in Experiment 1 here, and in Burnham et al. (1991).

Alveolar/dental stop voicing contrasts. For the dental voicing contrasts similar results were found. The native contrast was discriminated more easily than the non-native contrast, $F(1,189) = 764.55$ ($\overline{X}_{\text{native}} = 0.93$, $\overline{X}_{\text{non-native}} = 0.25$); and there were interactions of speech contrast with linear $F(1,189) = 45.29$, and quadratic age trends $F(1,189) = 4.20$, MSE = 0.06. Consideration of Figure 3b indicates that, as for the bilabial voicing contrasts, the age trend for the non-native contrast was more linear and more quadratic than for the native contrast. The quadratic trough here seems more focussed than for the bilabials and is centred on the Kindergarten group.

Place contrasts. Similar results were also found for the place contrasts. The native contrast was discriminated more easily than the non-native contrast, $F(1,189) = 326.36$, MSE = 0.07 ($\overline{X}_{\text{native}} = 0.89$, $\overline{X}_{\text{non-native}} = 0.42$), though the native / non-native difference was not as great here as for the bilabial and alveolar stop consonants, $F(1,189) = 44.45$, MSE = 0.06, especially for the adults, $F(1,189) = 4.15$. Figure 3c, in conjunction with the significant contrast × linear age interaction, $F(1,189) = 41.63$, MSE = 0.07, and the significant contrast × quadratic age interaction $F(1,189) = 5.38$, MSE = 0.07, shows that the trend for the non-native contrasts is both more linear and more quadratic with a trough centred on the Kindergarten group.

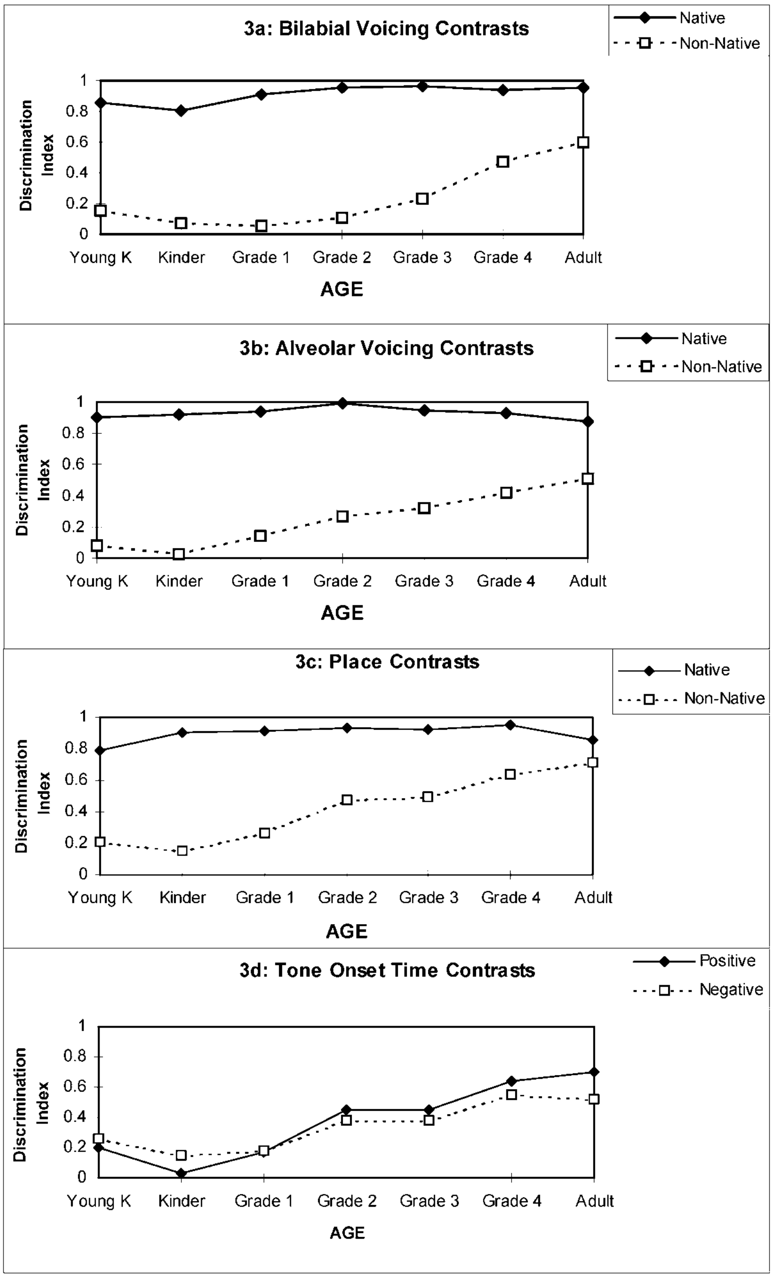

Figure 3. Discrimination of Native and Non-Native Speech Contrasts, Positive and Negative Tone Onset Time (TOT) Contrasts: (a) bilabials, (b) alveolars, (c) place, (d) TOT.

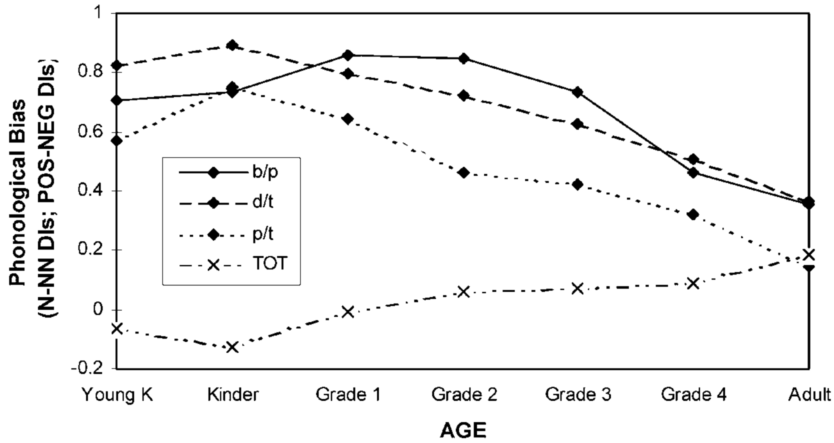

Figure 4. Language specific perceptual bias (N-NN) scores for Speech – bilabial stop, alveolar stop, and place contrasts – and Positive-Negative Scores for TOT sounds.

Non-speech contrasts. In contrast to the results for the speech contrasts, there was no significant difference in discrimination ability for the positive and negative TOT contrast, $F(1{,}189) = 1.34$, MSE = 0.07 ($\overline{X}_{\text{positive}} = .38$, $\overline{X}_{\text{negative}} = .35$), and no significant difference between the quadratic age trends for the two contrasts, $F(1{,}189) = 0.09$, MSE = 0.07, both showing a trough centred on Kindergarten (see Figure 3d). However, there was a difference in the linear component of the age trends, with the age trend for the negative TOT contrast being flatter than that for the positive TOT contrast, $F(1{,}189) = 11.53$, MSE = 0.069 (see Figure 3d).[4]

Speech vs non-speech differences. The results for the three sets of speech contrasts were strikingly similar, and quite different to those for the non-speech contrasts. Language specific speech perception measured by (N-NN) scores generally peaked around the end of Kindergarten year / start of Grade 1 though there was a significant bilabial/alveolar × native/non-native × quadratic age trend, $F(1{,}189) = 4.64$, MSE = 0.05, reflecting slightly different peaks for the bilabial and alveolar contrast sets (see Figure 4). This relatively common peak in (N-NN) scores presumably reflects children's focus on native language sounds as a function of age or schooling, although it is possible that this peak may be a product of some acoustic quirk of the particular native and non-native sounds we used. If this were indeed the case, then speech analogue sounds with similar acoustic characteristics should show a similar "language specific" peak. To investigate this, (N-NN) language specific speech perception scores were calculated for each subject for the three speech contrasts, along with positive minus negative

Table 2. Mean percentage correct in phoneme segmentation and reading tests in Experiment 2.

Test	Age/grade					
	Young K	Kinder	Grade 1	Grade 2	Grade 3	Grade 4
Phoneme segmentation	24	53	75	78	81	88
Reading	0	12	59	73	95	98

(P-N) TOT native bias scores.[5] The resultant mean native bias scores for speech contrasts were significantly greater than the positive TOT (P-N) bias scores, $F(1,189) = 339.79$, MSE = 0.07 (see Figure 4), and while the speech (N-NN) showed a peak in the early years of schooling, the TOT (P-N) scores showed no such peak, $F_{\text{speech/non-speech}\times\text{linear age}}(1,189) = 59.81$, $F_{\text{speech/non-speech}\times\text{quadratic age}}(1,189) = 7.19$, MSE = 0.07. These results strongly support the notion that there is an accentuation of language specific speech perception around late kindergarten to Grade 1, and that this is *specifically due to the phonological relevance of the speech contrasts*, rather than any particular acoustic characteristics.

Phoneme segmentation and reading tests

Percent correct phoneme segmentation and reading scores are shown in Table 2. For the phoneme segmentation scores, there was a general increase over age, $F_{\text{linear}}(1,144) = 179.23$, a general levelling out in Years 1, 2, 3, and 4, $F_{\text{quadratic}}(1,144) = 30.89$, and sharper rises at earlier (Young Kindergarten to Year 1) and later (Year 3 to Year 4) than in the middle age ranges, $F_{\text{cubic}}(1,44) = 4.74$. Within the phoneme segmentation task, productive identification ($\overline{X} = 46.17\%$) was more difficult than the receptive deletion version of the task ($\overline{X} = 55.88\%$) $F(1,144) = 6.41$, MSE = 2.9, but this did not interact with age. Further details of the phoneme awareness test and detailed results can be found in Burnham and Tam (in preparation).

For reading there was a significant increase over age $F_{\text{linear}}(1,156) = 621.61$, a significant levelling out at the older ages $F_{\text{quadratic}}(1,156) = 19.38$, and a less rapid improvement between Year 1 and 3 than between late Kindergarten and Year 1, $F_{\text{cubic trend}}(1,156) = 9.6$. Generally, girls had higher mean reading scores than boys (77.36% vs 66.41%), $F(1,156) = 8.51$, but this difference was ameliorated by Grades 3 and 4, $F(1,156) = 7.17$, MSE = 568.75. Thus for both reading and phoneme segmentation there was, as expected, a general improvement over age, with performance asymptoting at the older ages.

Table 3. Results of the regression of intelligence and sex on the first step, and age, phoneme segmentation, and reading on the second, against phonological bias scores for Experiment 2.

Contrast	Step 1					Step 2					
	Coefficients		R^2	F	P	Coefficients			ΔR^2	ΔF	P
	Intell	Sex				Age	Phon	Read			
Speech – bilabials	0.028	0.136	0.019	1.60	0.21	–0.460	0.137	0.236	0.062	3.67	0.01
Speech – alveolars	0.039	–0.019	0.002	0.158	0.85	–0.253	0.146	–0.213	0.128	7.94	0.0001
Speech – place	–0.038	0.030	0.002	0.203	0.81	–0.332	0.095	–0.048	0.099	5.98	0.001
Non-speech – TOT	0.164	–0.034	0.029	2.43	0.09	0.195	–0.167	0.088	0.039	2.25	0.08

Relationships between speech perception, reading, and phoneme segmentation

Four regression analyses were conducted in which the outcome measures were language specific speech perception scores (N-NN) for each of the three sets of speech contrasts, and positive TOT bias (P-N) scores for the non-speech contrasts. The predictor variables in each were rated intelligence, sex, age, phoneme segmentation, and reading. As significant relationships were found between (N-NN) and age and language variables in Experiment 1, here intelligence and sex were entered first to partial them out, and then the contribution of age, phoneme segmentation, and reading was ascertained. The results of these analyses, R^2, F-values and coefficients are shown in Table 3 for the three speech contrasts and the non-speech contrast, and the results are summarised below.

Speech contrasts. For all three speech contrasts, sex and intelligence did not significantly predict the (N-NN) dependent variable, whereas the subsequent addition of age and the language variables did. The coefficients are presented in Table 3. Applying the same interpretive device as in Experiment 1 (DeSarbo et al., 1982; Harris, 1985), the results for each contrast can be described as follows. For the bilabials, (N-NN) = Reading – Age, which is the same solution as was found for synthetic versions of these same contrasts in the identification task in Experiment 1, except that language comprehension (not measured here) appeared in the equation there. A similar verbal description is applicable, children who have a high degree of language specific speech perception tend to be good readers for their age. For the alveolars, (N-NN) = Phoneme Segmentation – Reading – Age, so for these contrasts children with a high degree of language specific speech perception are those who are relatively good segmenters for their age and level of reading ability. This is a similar but not identical result as that for bilabials and suggests that under some conditions a relationship between language specific

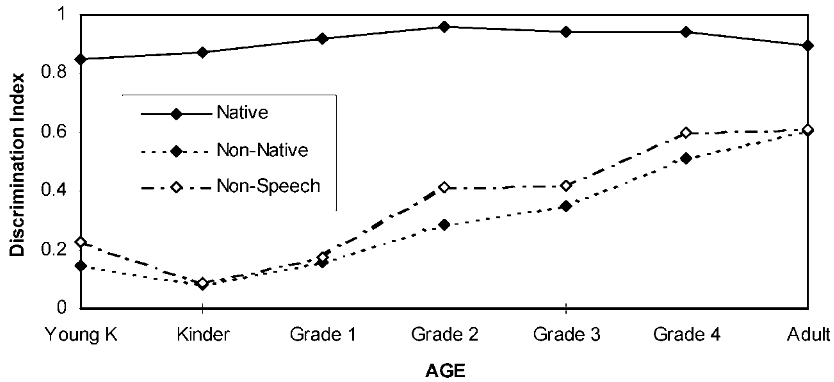

Figure 5. Discrimination of Native Speech, Non-Native Speech, and Non-Speech.

speech perception and phoneme segmentation may be observed. For the place contrasts, (N-NN) = – Age. This significant relationship does not include either of the linguistic factors, reading or phoneme segmentation.[6] However, it is in the same spirit as the other solutions, for there is a negative relationship with age, i.e., older children have greater language specific speech perception scores.

Non-speech contrasts. For the non-speech contrasts the regression for sex and intelligence was not significant, and additionally age and the language variables did *not* add significantly to the prediction. Thus applying the same analysis rationale as for the speech contrasts led to no significant relationship between any of the predictor variables and the dependent variable here, TOT (P-N). The only individual variable which even approached significance was intelligence ($P = 0.053$), and the positive beta weight (see Table 3) shows that (P-N) scores were positively related to intelligence (whereas for all the speech contrasts beta weights for intelligence were negligible). Thus it can be concluded (P-N) scores for the TOT contrasts are unaffected by linguistic variables (phonological segmentation and reading). This is in direct contrast to the speech perception results, which show a relationship between (N-NN) and linguistic variables, and/or a negative relationship with age.

Non-native speech vs non-speech. Given the better discrimination of native over non-native contrasts, it is of interest to determine the relationship between perception of non-native speech and non-speech. Mean discrimination indices for the three native contrasts, the three non-native contrasts, and the two non-speech contrasts are plotted against age in Figure 5. Overall discrimination of native contrasts was better than for non-native and non-speech contrasts combined, $F(1,189) = 1548.51$, MSE = 0.08, and there was a significant interaction of native versus non-native and non-speech combined with the linear $F(1,189) = 104.34$, quadratic $F(1,189) = 22.18$, and cubic $F(1,189) = 4.53$, MSE = 0.08, age trends. Discrimination of the non-native contrasts was better than for the non-speech contrasts, $F(1,189) = 7.21$, MSE = 0.11, but age trends for the two did not differ. To investigate the non-native speech versus non-speech effect further post-hoc tests at each age revealed that discrimination of non-native speech and non-speech was statistically equivalent in Kindergarten and Grade 1 children, and also for adults. However, discrimination of non-speech was significantly greater than non-native speech in Young Kindergarten, $F(1,189) = 4.16$, $P < 0.05$, Grade 2, $F(1,189) = 13.25$, $P < 0.01$, and Grade 4, $F(1,189) = 5.28$, $P < 0.01$, but despite a similar trend, failed to reach significance in Grade 3, $F(1,189) = 3.17$. This pattern of results could be described as a general superiority of non-speech over non-native speech, which is ameliorated around Kindergarten and Grade 1.

Discussion

The results are discussed with respect to the four aims of this experiment.

Ontogenetic locus of language specific speech perception

In Australian schools instruction in sound-symbol correspondences, phoneme segmentation, recognition of simple high-frequency words and so on, begins in Kindergarten year, while reading instruction proper begins in Grade 1. Here we found a sharp increase in reading and phoneme segmentation scores from the beginning to the end of Kindergarten year (see Table 2). For each of the three sets of speech contrasts language specific speech perception intensifies at the end of the Kindergarten year, subsequently to be attenuated at the end of the first (Kindergarten) or second (Grade 1) year of school. Peak levels of language specific speech perception thus occur just when reading and phoneme segmentation are rapidly improving, *after* the start of schooling.

Relationship between language specific speech perception and reading

While there was a slightly different regression solution for the place versus the two voicing contrasts,[6] it can generally be concluded that the results support the notion that there is a significant relationship between language specific speech perception and skills associated with the onset of reading instruction.

Nature of language specific speech perception

Peaks in language specific speech perception were found for the speech contrasts after the onset of schooling, but for the non-speech TOT contrasts there is a distinctively different developmental trend. This, plus the fact that the (N-NN) language specific speech perception is positively related to reading ability whereas positive (P-N) TOT bias is not, strongly implies that language specific speech perception is a specifically linguistic phenomenon.

Generalisability of language specific speech perception

Similar results to those in Experiment 1 were found here, despite a number of procedural differences: natural rather than synthetic speech sounds, a discrimination rather than an identification task, and two further sets of speech contrasts – alveolar stops, and a bilabial versus alveolar/dental place contrast. For all three speech contrasts, similar results regarding the locus of language specific speech perception were obtained, in all three a negative relationship with age was found, and in two of the three a relationship between reading and language specific speech perception was obtained. The similar results in Experiments 1 and 2 here and in Burnham et al. (1991) point to the pervasiveness of heightened language specific speech perception following the onset of reading.

General discussion

Together the results of the two experiments show that language specific speech perception is a specifically linguistic phenomenon, and that heightened language specific speech perception is related to the onset of reading.

The nature of language specific speech perception

Building upon previous research, the results of Experiment 1 and especially Experiment 2 (see Figure 5) suggest that language specific speech perception is comprised of three components, as follows.

There is *facilitation* of the perception of native speech sounds due to linguistic experience. There is *attenuation* of the perception of all other sounds, both non-native speech and non-speech. There is *additional attenuation* of the perception of non-native speech sounds *below the normal level of perceptual acuity*. These first two processes are evident throughout development, at least after the onset of phonemic speech perception in infancy (Kuhl, 1991; Kuhl et al., 1992; Tees & Werker, 1984; Werker et al., 1981; Werker &

Tees, 1983, 1984a; Werker & Lalonde, 1988; Werker, 1994), and are certainly evident over the ages studied here. The third process is generally evident over the ages studied here; though around the onset of reading perception of both non-native speech sounds and non-speech sounds are equally and maximally suppressed. This may be described as a floor effect, but it may be more heuristic to characterise this as an efficient strategy whereby incoming sounds are best assigned to orthographically defined categories if perceptual attention to *all* sounds which are not native speech is equally attenuated. (Discrimination of non-native speech and non-speech is also equal for adults, and this could be due to adults being better able to adopt strategies to perform a particular laboratory task.) Clearly these explanations are ad hoc, but they highlight the need for further investigation of the dynamic interaction of speech perception processes such as facilitation and attenuation right across development, in infants, children, and adults (Aslin, 1981; Gottlieb, 1976).

Language specific speech perception and the onset of reading

What happens to language specific speech perception with the onset of reading? Heightened language specific speech perception in the early years of schooling has now been found in three separate experiments (Burnham et al., 1991; Experiments 1 and 2 here), and to be significantly related to reading acquisition in Experiments 1 and 2 here. This is consistent with the notion that reading instruction in an alphabetic language directs perceptual attention to matching graphemes to phonemes, which in turn requires efficient phoneme segmentation, and efficient assignment of allophonic variants to strictly defined phoneme categories. The continued ability to perceive the phonetic characteristics which distinguish native speech phonemes and selectively suppress just those phonetic characteristics which distinguish non-native speech phones from each other (Werker, 1994; Burnham, 1986), and from similar native phones (Best, 1994), that is, language specific speech perception, would be very useful at this time. What we have found here is that there is heightened language specific speech perception after children begin school, and this is related to the onset of reading. Results here suggest that around reading onset children may adopt a strategy in which perception of *all* contrasts that are not phonologically relevant is suppressed, while both before and after this time perceptual performance for non-native speech is additionally suppressed.

To the extent that the evidence presented here suggests that the additional suppression for non-native speech contrasts is based on their phonetic characteristics rather than any acoustic characteristics they share with non-speech sounds, there is a prima facie argument against the involvement of a low-level *auditory* deficit as a precursor to reading difficulties (Tallal, 1980).

Moreover, it is clear that language specific speech perception is in place in early infancy, and thus is a likely candidate to play a role in the acquisition of later linguistic abilities (Kuhl, 1991; Kuhl et al., 1992; Tees & Werker, 1984; Werker et al., 1981; Werker & Tees, 1983, 1984a; Werker, 1994). Indeed, the results of Experiment 1 here suggest that around 4 years language specific speech perception is related to articulation ability. Nevertheless, there is not yet sufficient evidence to make definitive judgements on this issue; it remains possible that underlying individual differences in speech production, speech perception, general cognitive ability (Lalonde & Werker, 1995), and even auditory perception (Tallal, 1980) could contribute to more sharply defined phonological categories in some pre-school children than in others. Longitudinal studies from preschool through to post-reading age are clearly required.

Language specific speech perception in infancy and childhood

The heightened language specific speech perception uncovered here around reading onset is additional to, and in no way diminishes the importance of the well-documented perceptual reorganisation between 6 and 12 months (Polka & Werker, 1994; Kuhl et al., 1992; Werker & Lalonde, 1988; Werker & Tees, 1992; Werker, 1994; Lalonde & Werker, 1995). One main difference between earlier "phonemic" and the later "orthographic" periods (Burnham, Tyler & Horlyck, 2002) is that while in infancy perceptual reorganisation occurs due to the natural ontogenetic processes of perceiving and producing speech, in the later period, heightened language specific speech perception results from the imposition of the "unnatural" requirement to read an alphabetic language.

It is the similarities between these two periods that are of most interest. In both periods language specific speech perception does not involve any loss of speech perception ability. Werker and colleagues showed that if the ISI in an AX discrimination task is shortened to allow phonetic or acoustic processing (Werker & Tees, 1984b; Werker & Logan, 1985), or if extensive exposure is given (Werker & Logan, 1985), adults *can* discriminate non-native speech contrasts which under normal circumstances they appear unable to discriminate. Thus, there is no sensori-neural loss of speech perception; exposure to a particular language results in perceptual selectivity rather than perceptual loss. In a similar vein, the orthographic language specific speech perception uncovered here is transitory, occurring specifically around the onset of formal language instruction, and relinquishing hold around Grade 2 to Grade 3, possibly due to reading becoming more proficient.

Another important similarity between these two periods relates to the work of Lalonde and Werker (1995), who suggest that there is an initial perceptual analysis underlying all speech perception, which is untouched by specific linguistic experience or general cognitive ability. The output of such perceptual analysis then is redescribed at increasingly complex and abstract levels throughout development (Lalonde & Werker, 1995; Karmiloff-Smith, 1991). So, according to Lalonde and Werker, the same phonetic detail is available to infants before and after the six- to nine-month reorganisation period, but during this period there is a redescription of the output from phonetic analysers in order to reflect native speech categories. "The redescription results in a second representation superimposed on the first, leaving the original detail of language-general discrimination intact . . ." (Lalonde & Werker, 1995: 472). In this way Lalonde and Werker's model allows the infant's ability for redescription to be related to general cognitive abilities, even if underlying phonetic or acoustic abilities are more or less equivalent. This model may be expanded if more were known about how the processes of speech perception interact with other linguistic processes occurring at the time of redescription. For example, does perceptual assimilation (Best, 1994) operate equivalently at all periods of development, and do particular phonetic characteristics of speech sounds (Burnham, 1986; Stevens & Keyser, 1989) differentially affect perceptual performance in different developmental contexts?

The orthographic period uncovered here could be seen to be a further redescription, which is superimposed on the earlier redescription (see Burnham et al., 2002, for elaboration). And like the earlier redescription, there appear to be later vestiges of this. For instance, in auditory phoneme deletion and addition tasks, Tyler and Burnham (2000) found adults had longer reaction times for orthographic interference items, such as WORTH-EARTH – "add /w/ to /ɜθ/" or "delete /w/ from /wɜθ/", than for non-interfering items, such as FACT-ACT – "add /f/ to /ækt/" or "delete /f/ from /fækt/" (see also Taft & Hambly, 1985).

Implications

There are important implications of these results for reading instruction. Three of these are set out below.

Is reading instruction in an alphabetic language a necessary condition for the intensification of language specific speech perception? In order to test this, six-year-old children who were not being taught to read (either due to a later onset or absence of schooling; Morais et al., 1979), and six-year-olds being taught to read a non-alphabetic script (Read et al., 1986) should be compared on native and non-native speech perception. Additionally, given the supposed role of phoneme-to-grapheme correspondences in language specific speech perception, it would be of interest to investigate the degree of language specific speech perception when children learn a language with a transparent orthography such as Spanish or Turkish, in which phoneme-to-grapheme conversion rules are both feedforward and feedback consistent, versus an opaque language such as English or Irish, with irregular grapheme-to-phoneme mappings.

Does phonics reading instruction result in greater language specific speech perception than whole word reading instruction? Our arguments above support the use of phonic rather than whole word methods of early reading instruction, because the phonics method should facilitate acquisition of the arbitrary language-specific phoneme-to-grapheme conversion rule conventions. If, as we would predict, children taught with phonics-based methods (see Liberman & Liberman, 1992; Byrne & Fielding-Barnsley, 1993; Schwartz, 1988) show greater language specific speech perception independent of reading ability (Andrews, 1992a, b) than children taught with a whole word method (Goodman & Goodman, 1979; Goodman, 1986), then this would suggest that heightened language specific speech perception is not only a *consequence* of learning to read, but also an *index* of reading instruction efficacy.

Is heightened language specific speech perception the result of reading instruction or maturation? The heightened language specific speech perception found here in Experiments 1 and 2 appears to be related to both age and reading ability. It is not yet clear whether 'reading ability' relates to the amount of reading instruction, the type of reading instruction, the child's ability to benefit from reading instruction, pre-reading skills that children bring to school, each child's ability for graphemic redescription of the speech input, or children's general cognitive ability (Lalonde & Werker, 1995). These need to be investigated in a design in which reading instruction and age are independently manipulated. Such a study is currently being conducted in our laboratory.

Acknowledgements

The assistance of Lynda Earnshaw, Genevieve O'Connor, Brian O'Grady, Monica Bray, Lynnette Mason, Agatha Pupaher, Jane Lawrence, and Esme Nasser in stimulus construction, testing of children, and collation and analysis of data is greatly appreciated. We thank Professor David Pisoni of the Psychology Department at Indiana University for supplying the TOT stimuli, and Mr Salahuddin Bihar for producing the Hindi speech sounds. This research was supported by Australian Research Council grants to the author (A78616331, AC8932421, AC0231201, A79601993, and A00001283).

Notes

1. In speech perception studies with older children and adults, multiple trials are often given for each stimulus. This was not possible here due to the attention span of the youngest children, the four-year-olds. In the current study each child was required to complete two speech perception tests, i.e., the native and non-native tests, and a battery of language tests, so presentation of more than one exemplar for each test stimulus was not possible. Despite this limitation, the method for converting responses to categorical scores (see Burnham et al., 1991), adequately takes into account single presentations of test stimuli.
2. These TOT sounds are the same as those used in Pisoni's laboratory (Pisoni, 1977; Jusczyk et al., 1980), and we thank Professor Pisoni for supplying them.
3. d' or a non-parametric equivalent (e.g., A') was not appropriate here as the total number of AB and BA trials was small (4). In such cases d' or A' is often 0.
4. This difference in the linear components is underscored by a shift over age in relative performance on the two contrasts: at the younger ages there is superior performance on the negative contrast (Young Kindergarten & Kindergarten combined: $\overline{X}_{\text{positive}}$ = 0.11, $\overline{X}_{\text{negative}}$ = 0.21), in Grade 1 performance is equivalent (Grade 1: $\overline{X}_{\text{positive}}$ = 0.17, $\overline{X}_{\text{negative}}$ = 0.18), while at the older ages performance is better on the positive contrast (Grades 2, 3, 4, and adults: $\overline{X}_{\text{positive}}$ = 0.56, $\overline{X}_{\text{negative}}$ = 0.46). This latter result confirms adults' superiority for positive region contrasts found by Pisoni (1977), and extends it to children as young as about 7 years. The opposite effect for the four- and five-year-old Kindergarten children, i.e., superior performance on the negative region contrast and the crossover in Grade 1, is intriguing and deserves further study.
5. Note that the P-N TOT DIs are analogues of the N-NN DIs for the speech contrasts, because the positive region TOT contrast is an analogue of voicing contrasts which are phonologically relevant in English, while the negative region TOT contrast is an analogue of voicing contrasts which are phonologically irrelevant in English.
6. A possible reason for this may be that the non-native place contrast was the easiest of the non-native contrasts to discriminate (see Figure 3c). Of the three sets of speech contrasts the place contrasts had the lowest mean N-NN scores for the combined child groups ($\overline{X}_{\text{bilabials}}$ = 0.72, $\overline{X}_{\text{alveolars}}$ = 0.77, $\overline{X}_{\text{place}}$ = 0.53) and this was attributable to the fact that while the native contrasts were equally difficult in each of the three pairs ($\overline{X}_{\text{bilabials}}$ = 0.90, $\overline{X}_{\text{alveolars}}$ = 0.94, $\overline{X}_{\text{place}}$ = 0.90), the non-native contrast was considerably easier for the place contrasts ($\overline{X}_{\text{bilabials}}$ = 0.18, $\overline{X}_{\text{alveolars}}$ = 0.17, $\overline{X}_{\text{place}}$ = 0.37). It is thus quite likely that the relationship between language specific speech perception and related reading skills is only evident when difficult to perceive non-native speech contrasts are used.

Appendix 1: Characteristics of speech sounds in Experiment 2

1A and *1B*: Duration (msecs) and *VOT (msecs)* of stimuli in Block A: bilabial voiced stops; and Block B: alveolar voiced stops (The subscripts 1 and 2 refer to the different tokens used for each phoneme type).

	Native				Non-native			
A: Bilabial	[pa]$_1$	[pa]$_2$	[p^{h}a]$_1$	[p^{h}a]$_2$	[ba]$_1$	[ba]$_2$	[pa]$_1$	[pa]$_2$
	334 – *16*	321 – *79*	312 *87*	323 *91*	325 – *88*	320 – *99*	258 – *6*	246 – *3*
B: Alveolar	[ta]$_1$	[ta]$_2$	[t^{h}a]$_1$	[t^{h}a]$_2$	[d^{h}a]$_1$	[d^{h}a]$_2$	[t^{h}a]$_1$	[t^{h}a]$_2$
	329 – *12*	326 – *33*	324 *88*	340 *109*	431 – *121*	435 – *125*	312 *64*	300 *47*

1C: Duration, VOT, and formant details for Block C: place contrasts (the subscripts 1 and 2 refer to the different tokens used for each phoneme type).

Measure	Native				Non-native			
	[p^{h}a]$_1$	[p^{h}a]$_2$	[t^{h}a]$_1$	[t^{h}a]$_2$	[t̪a]$_1$	[t̪a]$_2$	[t̥a]$_1$	[t̥a]$_2$
Duration (msec)	319	319	327	328	265	295	266	287
VOT (msec)	94	91	85	89	–9	–19	–9	–13
F_1 start transition (Hz)	1300	1189	1327	1382	331	387	221	248
F_1 end transition (Hz)	1161	1134	1161	1189	470	414	525	525
F_2 start transition (Hz)	2129	2378	2351	2351	2019	1825	2102	1991
F_2 end transition (Hz)	2295	2268	2295	2268	1770	1438	1825	1714
$\overline{X}$ F_3 transition (Hz)	3039	3198	3062	3032	2877	2925	2881	2858

References

Andrews, S. (1992a). A skills approach: Optimising initial reading instruction. In A. Watson and A. Badenhop (Eds.), *The prevention of reading failure* (pp. 84–97). Sydney: Ashton Scholastic.

Andrews, S. (1992b). Ideology and methods of reading instruction: Why is phonics an F-word? In *English literacy round table papers.* Sydney: Board of Studies, New South Wales.

Aslin, R.N. (1981). Experiential influences and sensitive periods in perceptual development: A unified model. In R.N. Aslin, J.R. Alberts & M.R. Petersen (Eds.), *Development of perception: Psychobiological perspectives. Volume 2: The visual system* (pp. 45–93). New York: Academic Press.

Aslin, R.N., Pisoni, D.B., Hennessy, B.L. & Perey, A.J. (1981). Discrimination of voice onset time by human infants: New findings and implications for the effects of early experience. *Child Development, 52*, 1135–1145.

Bankson, N.W. (1977). *Bankson language screening test*. Baltimore: University Park Press.

Best, C.T. (1994). The emergence of native-language phonological influences in infants: A perceptual assimilation model. In J.C. Goodman & H.C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 167–224). Cambridge, Massachusetts: MIT Press.

Best, C.T., McRoberts, G.W. and Sithole, N.M. (1988). Examination of perceptual reorganization for nonnative speech contrasts: Zulu click discrimination by English speaking adults and infants. *Journal of Experimental Psychology: Human Perception and Performance, 14*, 345–360.

Bradley, L. & Bryant, P. (1978). Difficulties in auditory organization as a possible cause of reading backwardness. *Nature, 271*, 746–747.

Bradley, L. & Bryant, P. (1985). *Rhyme and reason in reading and spelling*. Ann Arbor: University of Michigan Press.

Brady, S., Shankweiler, D. & Mann, V. (1983). Speech perception and memory coding in relation to reading ability. *Journal of Experimental Child Psychology, 35*, 345–367.

Brandt, J. & Rosen, J.J. (1980). Auditory phonemic perception in dyslexia: Categorical identification and discrimination in stop consonants. *Brain and Language, 9*, 324–337.

Burnham, D.K. (1986). Developmental loss of speech perception: Exposure to and experience with a first language. *Applied Psycholinguistics, 7*, 207–240.

Burnham, D.K. & Tam, H. (in preparation) An Australian phoneme deletion and addition test.

Burnham, D.K., Earnshaw, L.J. & Quinn, M.C. (1987). The development of categorical identification of speech. In B.E. McKenzie & R.H. Day (Eds.), *Perception in infancy: Problems and issues* (pp. 237–275). Hillsdale, New Jersey: Lawrence Erlbaum.

Burnham, D.K., Earnshaw, L.J. & Clark, J.E. (1991). Development of categorical identification of native and non-native bilabial stops: infants, children and adults. *Journal of Child Language, 18*, 231–260.

Burnham, D.K., Tyler, M.D. & Horlyck, S. (2002). Periods of speech perception development and their vestiges in adulthood. In P. Burmeister, T. Piske & A. Rohde (Eds.), *An integrated view of language development: Papers in honor of Henning Wode* (pp. 281–300). Trier, Germany: Wissenschaftlicher Verlag Trier.

Byrne, B. & Fielding-Barnsley, R. (1993). Evaluation of a program to teach phonemic awareness to young children: A one-year follow-up. *Journal of Educational Psychology, 85*, 104–111.

Calfee, R.L., Lindamood, P. & Lindamood, C. (1973). Acoustic-phonetic skills and reading in kindergarten through twelfth grade. *Journal of Educational Psychology, 64*, 293–298.

Daniels, J.C. & Diack, H. (1974). *The standard reading tests*. London: Shatto & Windus.

De Sarbo, W.S., Hausman, R.E., Shen, L. & Thompson, W. (1982). Constrained canonical correlation. *Psychometrika, 47*, 489–516.

Eimas, P.D., Siqueland, E.R., Jusczyk, P. & Vigorito, J. (1971). Speech perception in infants. *Science, 171*, 303–306.

Gibbs, S. (1996). Categorical speech perception and phonological awareness in the early stages of learning to read. *Language and Communication, 16*, 37–60.

Godfrey, J.J., Syrdal-Lasky, A.K., Millay, K.K. & Knox, C.M. (1981). Performance of dyslexic children on speech perception tests. *Journal of Experimental Child Psychology, 32*, 401–424.

Goodman, J.C. & Nusbaum, H.C. (eds.) (1994). *The development of speech perception: The transition from speech sounds to spoken words*. Cambridge, Massachusetts: MIT Press.

Goodman, K. (1986). *What's whole in whole language: A parent-teacher guide*. Portsmouth, NH: Heinemann.

Goodman, K. & Goodman, Y. (1979). Learning to read is natural. In L. Resnick & P. Weaver (Eds.), *Theory and practice in early reading* (Vol. 1, pp. 137–154). Hillsdale, New Jersey: Erlbaum.

Goswami, U. & Bryant, P. (1990). *Phonological skills and learning to read*. Hove, UK: Erlbaum.

Gottlieb, G. (1976). The roles of experience in the development of behavior and the nervous system. In G. Gottlieb (Ed.), *Neural and behavioural specificity* (pp. 25–53). New York: Academic Press.

Hansen, J. & Bowey, J.A. (1994). Phonological analysis skills, verbal working memory, and reading ability in second-grade children. *Child Development, 65*, 938–950.

Harris, R.J. (1985). *A primer of multivariate statistics* (2nd ed.). New York: Academic Press.

Hurford, D.P. & Sanders, R.E. (1990). Assessment and remediation of a phonemic discrimination deficit in reading disabled second and fourth graders. *Journal of Experimental Child Psychology, 50*, 369–415.

Joanisse, M.F Manis, F.R., Keating, P. & Seidenberg, M.S. (2000). Language deficits in dyslexic children: Speech perception, phonology, and morphology. *Journal of Experimental Child Psychology, 77*, 30–60.

Jusczyk, P.W., Pisoni, D.B., Walley, A. & Murray, J. (1980). Discrimination of relative onset time of two-component tones by infants. *Journal of the Acoustical Society of America, 67*, 262–270.

Karmiloff-Smith, A. (1991). Beyond modularity: Innate constraints and developmental change. In S. Carey & R. Gelman (Eds.), *The epigenesis of mind: Essays on biology and cognition* (pp. 171–197). Hillsdale, New Jersey: Erlbaum.

Kilminister, M.G.E. & Laird, E.M. (1978). Articulation development in children aged three to nine years. *Australian Journal of Communication Disorders, 6*, 23–30.

Kuhl, P.K. (1991). Human adults and human infants show a "perceptual magnet effect" for prototypes of speech categories, monkeys do not. *Perception and Psychophysics, 50*, 93–107.

Kuhl, P.K., Williams, K.A., Lacerda, F., Stevens, K.N. & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science, 255*, 606–608.

Lalonde, C.E. & Werker, J.F. (1995). Cognitive influences on cross-language speech perception in infancy. *Infant Behavior and Development, 18*, 459–475.

Liberman, I.Y. & Liberman, A.M. (1992). Whole language versus code emphasis: underlying assumptions and their implications for reading instruction. In P.B. Gough, L.S. Ehri & R. Treiman (Eds.), *Reading acquisition* (pp. 343–366). Hillsdale, New Jersey: Erlbaum.

Manis, F.R., McBride-Chang, C., Seidenberg, M.S., Keating, P., Doi, L.M., Munson, B. & Petersen, A. (1997). Are speech perception deficits associated with developmental dyslexia? *Journal of Experimental Psychology, 66*, 211–235.

Mann, V.A. & Liberman, I.Y. (1984). Phonological awareness and verbal short-term memory. *Journal of Learning Disabilities, 17*, 592–599.

McBride-Chang, C. (1996). Models of speech perception and phonological processing in reading. *Child Development, 67*, 1836–1856.

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*, 77–80.

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.

Morais, J., Cary, L., Alegria, J. & Bertelson, P. (1979). Does awareness of speech as a sequence of phones arise spontaneously. *Cognition, 7*, 323–331.

Pisoni, D.B. (1977). Identification and discrimination of the relative onset time of two-component tones: implication for voicing perception in stops. *Journal of the Acoustical Society of America, 61*, 1352–1361.

Polka, L. & Werker, J.F. (1994). Developmental changes in perception of nonnative vowel contrasts. *Journal of Experimental Psychology: Human Perception and Performance, 19*, 421–435.

Read, C., Zhang, Y., Nie, H. & Ding, B. (1986). The ability to manipulate speech sounds depends on knowing alphabetic writing. *Cognition, 24*, 31–44.

Reed, M.A. (1989). Speech perception and the discrimination of brief auditory cues in reading disabled children. *Journal of Experimental Child Psychology, 48*, 270–292.

Schwartz, S. (1988). A comparison of componential and traditional approaches to training reading skills. *Applied Cognitive Psychology, 2*, 189–100.

Stevens, K.N., & Keyser, S.J. (1989). Primary features and their enhancement in consonants. *Language, 65*, 81–106.

Strange, W. & Broen, P.A. (1981). The relationship between perception and production of /w/, /r/, and /l/ by three-year-old children. *Journal of Experimental Child Psychology, 31*, 81–102.

Streeter, L.A. (1976). Language perception of 2 month-old infants shows effects of both innate mechanisms and experience. *Nature, 259*, 39–41.

Taft, M. & Hambly, G. (1985). The influence of orthography on phonological representations in the lexicon. *Journal of Memory and Language, 24*, 320–335.

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

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

Tallal, P. & Piercy, M. (1973b). Developmental aphasia: Impaired rate of non-verbal processing as a function of sensory modality. *Neuropsychologia, 11*, 389–398.

Tallal, P. & Piercy, M. (1974). Developmental aphasia: Rate of auditory processing and selective impairment of consonant processing. *Neuropsychologia, 12*, 83–93.

Tallal, P., Miller, S.L., Bedi, G., Byma, G., Wang, X., Nagarajan, S.N., Schreiner, C., Jenkins, W.M. & Merzenich, M.M. (1996). Language comprehension in language-learning impaired children with acoustically modified speech. *Science, 271*, 81–83.

Tees, R.C. & Werker, J.F. (1984). Perceptual flexibility: Maintenance or recovery of the ability to discriminate non-native speech sounds. *Canadian Journal of Psychology, 38*, 579–590.

Tyler, M.D. & Burnham, D.K. (2000). Orthographic influences on initial phoneme addition and deletion tasks: The effect of lexical status. In B. Yuan, T. Huang & X. Tang (Eds.), *Proceedings of the sixth international conference on spoken language processing* (Vol. 3, pp. 494–497). Beijing, China: China Military Friendship Publish.

Wagner, R.K. & Torgesen, J.K. (1987). The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin, 101*, 192–212.

Wagner, R.K., Torgeson, J.K., Laughton, P., Simmons, K. & Rashotte, C.A. (1993). Development of young readers' phonological processing abilities. *Journal of Educational Psychology, 85*, 83–103.

Wagner, R.K., Torgeson, J.K. & Rashotte, C.A. (1994). Development of reading-related phonological processing abilities: New evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology, 30*, 73–87.

Watson, B.U. & Miller, T.K. (1993). Auditory perception, phonological processing, and reading ability/disability. *Journal of Speech and Hearing Research, 36*, 850–863.

Werker, J.F. (1993). The contribution of the relation between vocal production and perception to a developing phonological system. *Journal of Phonetics, 21*, 177–180.

Werker, J.F. (1994). Cross-language speech perception: Development change does not involve loss. In J.C. Goodman & H.C. Nusbaum (Eds.), *The development of speech perception: The transition from speech sounds to spoken words* (pp. 95–120). Cambridge, Massachusetts: MIT Press.

Werker, J.F. & Lalonde, C.E. (1988). Cross-language speech perception: Initial capabilities and developmental change. *Developmental Psychology, 24*, 672–683.

Werker, J.F. & Logan, J.S. (1985). Cross-language evidence for three factors in speech perception. *Perception and Psychophysics, 37*, 35–44.

Werker, J.F. & Tees, R.C. (1983). Developmental changes across childhood in the perception of non-native speech sounds. *Canadian Journal of Psychology, 37*, 278–286.

Werker, J.F. & Tees (1984a). Cross language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant and Behavior and Development, 7*, 49–63.

Werker, J.F. & Tees (1984b). Phonemic and phonetic factors in adult cross-language speech perception. *Journal of the Acoustical Society of America, 75*, 1866–1878.

Werker, J.F. & Tees, R.C. (1987). Speech perception in severely disabled and average reading children. *Canadian Journal of Psychology, 41*, 48–61.

Werker, J.F. & Tees (1992). The organization and reorganization of human speech perception. *Annual Review of Neuroscience, 15*, 377–402.

Werker, J.F., Gilbert, J.H.V., Humphrey, K. & Tees, R.C. (1981). Developmental aspects of cross-language speech perception. *Child Development, 52*, 349–355.

Wolf, M. (1984). Naming, reading, and the dyslexias. *Annals of Dyslexia, 34*, 87–115.

Zlatin, M.A. & Koenigsknecht, R.A. (1976). Development of the voicing contrast: A comparison of voice onset time in stop perception and production. *Journal of Speech and Hearing Research, 19*, 93–111.

Address for correspondence: Denis Burnham, Director, MARCS Auditory Laboratories, University of Western Sydney, Sydney, NSW 1797, Australia
Phone: +61-2-9772-6681; Fax: +61-2-9772-6736; E-mail: d.burnham@uws.edu.au