The development of perceptual cue-weighting in children aged 6 to 12

Valerie Hazan and Sarah Barrett

Department of Phonetics and Linguistics, UCL, 4, Stephenson Way, London NW1 2HE.
http://www.phon.ucl.ac.uk

ABSTRACT

Development in the ability to categorise phonemic contrasts was examined in 85 children aged 6 to 12 and thirteen adult controls using synthetic stimulus continua presented in identification tests. The range of contrasts included a stop voicing contrast (/g/-/k/), a stop place contrast (/d/-/g/), a fricative voicing contrast (/s/-/z/), a fricative place contrast (/s/-/ʃ/), and a fricative-affricate manner contrast (/ʃ/-/tʃ/). This study aimed to test: (1) whether there would be age effects in the sharpness of categorisation; (2) whether the sharpness of categorisation would differ across phonemic contrasts acquired at different stages of development and (3) whether children would give greater perceptual weighting than adults to cues marked by rapid spectral change.

Shallower identification function gradients were obtained for the 6;0-7;6 year old group than for the 9 to 12 year old groups, and all groups of children showed decreased sharpness of categorisation relative to adults. As predicted, the stop place and voicing contrasts were categorised more sharply than the fricative place and voicing contrasts. There was no conclusive support for the view that children aged 6 or more still placed greater reliance on dynamic cues than older children or adults.

1. INTRODUCTION

Early studies of phonological development which examined labelling ability in children aged 3 and above using identification tests [e.g. 1,2] showed evidence of gradual development over a period of years. It was suggested that, with age, there was an increase in the sharpness of categorisation and increased proficiency in using the full range of acoustic cues marking a phonemic contrast. Although such studies tested children in different age groups, they have typically each been limited to the evaluation of a single phonemic contrast so that it has not been possible to look at the relative development of labelling ability across a range of phonemic contrasts. In order to examine this issue, this study has tested children from a wide age-range on a set of phonemic contrasts thought to be acquired at different stages of phonological development.

More recently, a number of studies by Nittrouer and others [e.g. 3] have focussed on the difference in perceptual weighting given by young children and adults to cues such as formant transitions that involve rapid spectral change ('dynamic' cues) relative to cues such as friction regions which show slower variations ('static' cues). Their findings led to the Developmental Weighting Shift (DWS) model [3] which proposes that at an early stage of their acquisition of speech and language, children assign greater weight to dynamic cues, which signal syllable structure. It is suggested that children gradually shift to the weightings given by adults to acoustic cues in order to retrieve more fine-grained phonetic information as their lexicon increases. It has been suggested that the shift away from a strong reliance on dynamic cues occurs by the age of five [e.g.4] but few of these studies have tested children aged above 7 or 8. It is therefore of interest to examine the development of labelling and cue-weighting in school-age children between the ages of 6 to 12 to evaluate whether there is still evidence of a differential use of acoustic cues between children and adults.

In summary, the objectives of this study were to examine: (1) whether sharpness of categorisation continues to develop between the ages of 6 and 12, and in comparison with adults and (2) whether there is still evidence in 6 to 12 year olds of an increased reliance on formant transition cues relative to adults.

2. EXPERIMENT

2.1 Subjects

Listeners included 85 children (22 boys and 63 girls) aged between 6 and 12. They were divided into 6 age groups: 6;0-7;6 (16 children), 7;6-8;6 (9 children), 8;6-9;6 (7 children), 9;6-10;6 (19 children), 10;6-11;6 (16 children),11;6-12;6 (18 children). There were 13 adult controls.

Inclusion criteria for children were that they were: (1) native speakers of English (non-bilingual), with (2) no documented history of chronic middle-ear infection, (3) normal hearing thresholds and (4) no documented history of Specific Learning Difficulties. A great majority of children (80/85) were pupils at six independent (private) schools in the London area. All children passed the following screening tests: (a) a test of language competence (Recalling Sentences subtest of the CELF-R language test), (b) a pure tone audiometry screening test at 20 dB HL at frequencies between 0.25 and 4 kHz; (c) an identification test using naturally-produced versions of the minimal pairs used in the experiment.

2.2 Stimuli

Five consonantal phonemic contrasts were employed. They included a stop voicing contrast (/g/-/k/), a stop place contrast (/d/-/g/), a fricative voicing contrast (/s/-/z/), a fricative place contrast (/s/-/ʃ/) and a fricative-affricate manner contrast (/ʃ/-/tʃ/) in word-initial position. The minimal pairs included words taken from age-appropriate vocabulary lists that can be represented pictorially. Because of this requirement, the

vocalic context could not be kept constant over all pairs. The phonemic contrasts are presented below, in expected order of acquisition given our knowledge of phonological development.

Phonemic Contrast	Minimal pair	Acoustic cues varied
/g/-/k/ Stop voicing	Goat-Coat *	VOT, F1 transition
/d/-/g/ Stop place	Date-Gate *	Burst frequency, F2/F3 transitions
/s/-/z/ Fricative voicing	Sue-Zoo	Friction duration, voicing amplitude
/s/-/ʃ/ Fricative place	Sue-Shoe *	Friction frequency, F2 transition
/tʃ/-/ʃ/ Fric/Affricate manner	Chew-Shoe	Friction duration, burst intensity

Table 1: List of phonemic contrasts and minimal pairs. The pairs that are asterisked are those that were presented in both combined-cue and single-cue conditions.

Copy-synthesised versions of these words were prepared, each based on an utterance produced by a phonetically-trained female speaker. A six-stimulus continuum was then created in which the acoustic patterns marking the phonemic contrast were varied in equal logarithmic steps between the values appropriate for the two endpoints. In the test situation, these tokens were presented using a two-alternative forced-choice identification procedure. An adaptive procedure was used in order to tailor the duration of the test to a child's ability to identify the contrast. Typically, each test consisted of about 65 presentations.

All minimal pairs were presented in a 'combined-cue' condition in which two cues to the contrast were varying simultaneously. In order to look at the perceptual-weighting of acoustic cues, three of these pairs (asterisked in the table above) were presented in both 'combined-cue', and 'single-cue' conditions. In the single-cue conditions, acoustic patterns marking one of the cues varied as in the 'combined-cue' condition whilst the others were kept at a constant value. Details of the parameters that were varied and spectrograms of the endpoint stimuli for the different continua can be found at the following web site: www.phon.ucl.ac.uk/project/spedev.htm

2.3. Test procedure

Children were tested individually in a quiet room in their own school. Testing, which included a presentation of the ten test conditions and various screening tests, was accomplished in two sessions of approximately 30 minutes each. The speech perception tests were presented via a laptop computer with the child hearing the stimuli via headphones. For each listening test, pictures representing the two possible choices were placed on the laptop computer and the children indicated which word they had heard by pressing the mouse button closest to the appropriate picture. At the end of each testing session, the first listening test was repeated to evaluate test-retest reliability. Each child was randomly presented with one of five possible orders of presentation of the ten listening tests.

2.4 Results

The output of each test is an identification function that plots the percentage of one of the two alternative responses across the continuum. A maximum likelihood estimation (MLE) procedure was used to fit a cumulative normal function to each child's set of data per continuum. Two parameters were extracted to characterise each identification function: (1) the gradient of the fitted curve (expressed as probit units divided by the number of stimuli in the continuum) and (2) the phoneme boundary, calculated as the 50% point of the fitted labelling curve. The identification function gradient may be used as an indication of labelling consistency. The phoneme boundary is an important measure for assessing effects of the various cue adjustments. The mean gradients for the combined-cue conditions of the five phonemic contrasts are presented in Figure 1. The gradient and boundary measures were used in general linear model analyses (repeated measures) to examine the effects of the factors of age group and test condition.

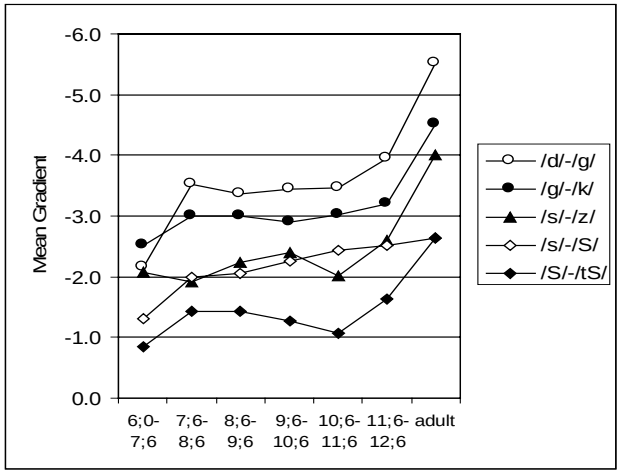

Figure 1: Mean gradient per age-group for the five phonemic contrasts presented in 'combined-cue' conditions.

Effect of age. The effect of age group on function gradient was strongly significant [$F(6,91)=16.896$; $p<0.0001$) and pairwise comparisons showed that steeper identification function gradients were obtained for adults than for all child groups, and that within the groups of children, gradients obtained for the youngest group (6;0-7;6) differed significantly from those for groups of children aged between 9;6 and 12;6. There was therefore evidence of development over the ages of 6 to 12;6, but by the age of 12;6, children did still not appear to have reached the same sharpness of categorisation as young adults.

Effect of phonemic contrast. Next, the data was analysed to look for evidence of significant difference in sharpness of categorisation across phonemic contrasts. For this, data from the five combined-cue conditions were used. The effect of phonemic

contrast was strongly significant [F(4,364)=42.247; p<0.0001] as was the effect of age group [F(6,91)=12.598;p<0.0001]. The contrast by group interaction did not reach significance. Pairwise comparisons showed that all contrasts differed from each other in terms of their gradient except that /s/-/z/ did not differ in gradient from /s/-/ʃ/. As expected, the stop contrasts were labelled more sharply than the fricative contrasts which were themselves labelled more sharply than the fricative-affricate contrast.

Evidence of significant differences between age groups for individual contrasts was examined using Tukey's HSD post-hoc analyses. For the /d/-/g/ test (combined-cue condition), adults labelled the stimulus continuum more sharply than any of the groups of children and the youngest group differed significantly from the oldest group of children and adults. For the /g/-/k/ and the /ʃ/-/tʃ/ continua, there was no significant increase in gradient for children aged between 6 and 12, but a strong difference between children aged 6 to 11 and adults. For the /s/-/z/ test, the pattern was more complex with gradients seen in some of the child groups not differing from the adult group. For the /s/-/ʃ/ test, the gradient did not differ significantly between any groups.

Perceptual weighting of cues. Perceptual weighting was examined in three contrasts: /g/-/k/, /d/-g/ and /s/-/ʃ/. The aim was to discover whether adults weighted certain acoustic cues differently to children aged between 6 and 12, and if so, whether it would still be possible to see evidence of greater weighting given by children to the dynamic formant transition cues. Results are presented here for the two stop contrasts /k/-/g/ and /d/-/g/.

The use of voice onset time and first formant transition information in cueing the /g/-/k/ contrast has been investigated in a number of studies [e.g. 1,2]. In our stimulus continuum, VOT ranged from 5 ms at the /g/ endpoint to 55 ms at the /k/ endpoint; F1 onset frequency varied from 380 Hz at the /g/ endpoint to 716 Hz at the /k/ endpoint. In the 'no F1 transition' condition, F1 onset frequency was kept constant at 716 Hz. According to the DWS model, it would be expected that the F1 transition cue would have greater weighting in young children than in adults. It is of interest here whether such an effect would still be evident in older children aged between 6 and 12. An effect of the F1 transition cue would be seen as a shift in phoneme boundary between the two test conditions (i.e. less /g/ responses in the absence of a rising F1 transition).

Analyses of variance were carried out on the boundary data for each of the age groups separately. Pairwise comparisons between the phoneme boundaries for the combined-cue and 'no F1 transition' conditions were significant for each of the age groups. There was no evidence of the extent of the phoneme boundary shift being greater in children than in adults. Looking at the mean identification function for the youngest and oldest children and adults (Figure 2), it does appear that, even though the shift in boundary occurs in all groups, a difference is seen between groups in the labelling of stimuli with a voice onset time of less than 20 ms. Adults consistently perceived such stimuli as voiced whether or not the F1 transition was present whereas the children gave fewer /g/ responses in the absence of the transition. There was however no conclusive evidence that children aged 6 or over were giving more weight than adults to the F1 transition cue.

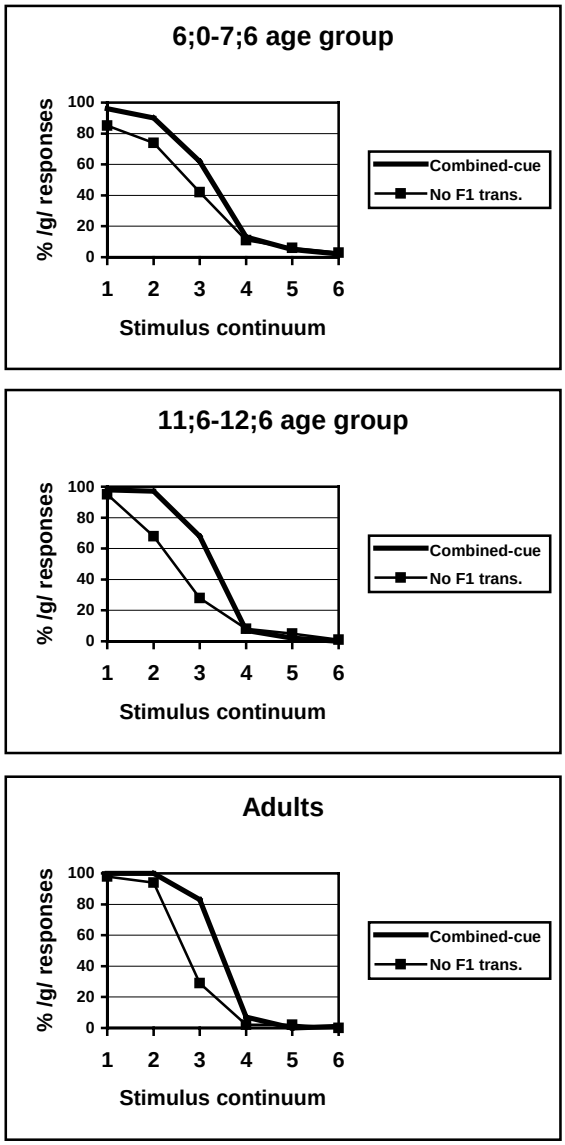

Figure 2: Mean identification functions for the 6;0-7;6, 11;6-12;6 and adult groups obtained by averaging percentages of /g/ responses obtained by each subject for each test condition. Step1 of the continuum corresponds to the /g/ endpoint.

In the /d/-/g/ test, the cues that were varied were the spectral characteristics of the initial burst transient and the F2/F3 transitions into the following vowel. The F2 onset frequency varied between 1800 Hz in /d/ and 2522 Hz in /g/. F3 frequency varied between 2910Hz in /d/ and 2925 Hz in /g/. The burst transient was synthesised using the parallel configuration of the synthesiser and differed in the frequency and amplitude of two poles. The main change was in the frequency of F5 which varied between 5950 Hz in /d/ and 3700 Hz in /g/ with a concomitant

reduction in amplitude. In the burst-cue condition, F2 onset frequency was fixed at 2060 Hz and F3 frequency at 2916 Hz. In the transition-cue condition, a burst was present but did not vary in frequency and amplitude across the continuum. Its parameters were set at an intermediate value to those used for the endpoint stimuli.

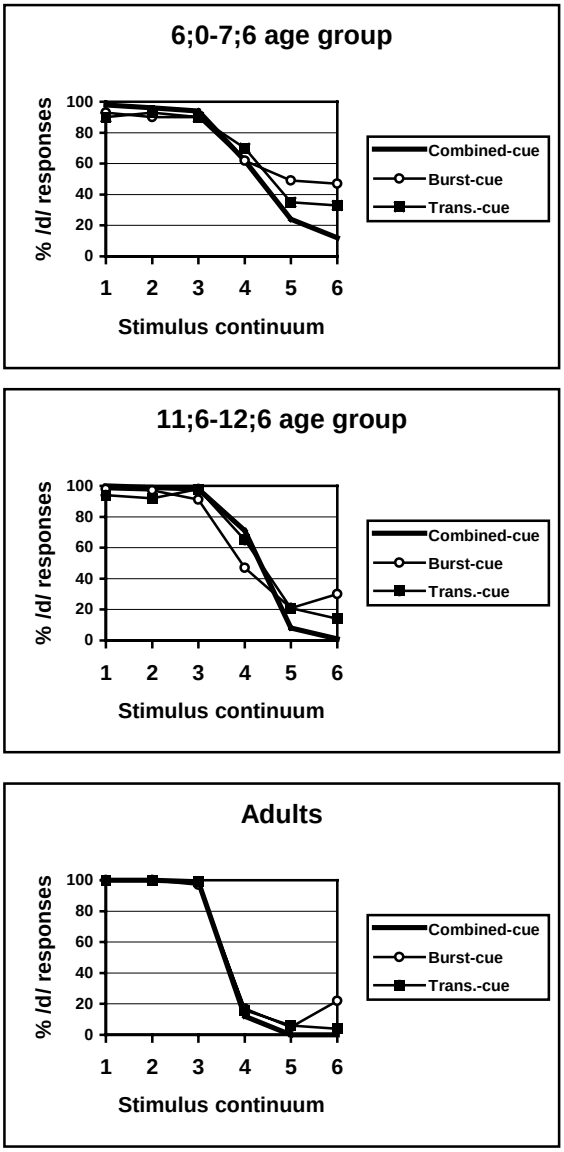

Figure 3: Mean identification functions for the 6;0-7;6, 11;6-12;6 and adult groups obtained by averaging percentages of /d/ responses obtained by each listener for each test condition.

For both the transitions-cue [$F(6,91)=10.169$; $p<0.0001$]and the burst-cue [$F(6,91)=4.950$; $p<0.0001$] conditions, Tukey's HSD post-hoc analyses revealed that identification function gradients were shallower for all children than for adults. Even though the difference between child groups was not significant, there was a trend of increasing gradient with age. In the youngest group of children, the difference in gradient between conditions was significant [$F(2,30)=11.774$;$p<0.0001$]; both single-cue conditions were less sharply labelled than the combined-cue condition but the burst-condition was not less sharply labelled than the transition-condition (Figure 3). In summary, by age 6 or older, children did not show any evidence of being able to categorise the /d/-/g/ contrast any better when it was cued by a change in formant transitions than when it was solely cued by a change in the burst; adults generally were more proficient than children in labelling both single-cue stimuli. This finding that, by age 6, children do not show greater reliance on formant transition information than adults, is in accordance of Ohde et al. [4] who investigated the perception of cues to stop place contrasts using very different stimuli and methodology.

4. DISCUSSION

It was found that sharpness of categorisation continued to increase between the ages of 6 and 12 but that, by that age, children were still not labelling the five phonemic contrasts under investigation as consistently as adults. This therefore provides further evidence [e.g. 5] that sharpness of categorisation, and hence labelling consistency of phonemic contrasts continues to increase until adulthood. Also, there was evidence in children as well as in adults of sharper categorisation of stop contrasts relative to fricative or fricative-affricate contrasts. The second objective of this study was to examine the relative use of formant transition cues in children and adults. By the age of six, there was no longer any conclusive evidence of a greater use than adults of dynamic cues such as formant transitions. This does not contradict the claims of the Developmental Weighting Shift model as the shift away from dynamic cues is proposed to have arisen before this age.

5. ACKNOWLEDGMENTS

This work was funded under the small-grants scheme of the Medical Research Council.

6. REFERENCES

1. Simon, C., and Fourcin, A. "Cross-language study of speech-pattern learning", J. Acoust. Soc. Am., vol. 63, 1978, pp.925-935.

2. Burnham, D., Earnshaw, L., Quinn, M "The development of categorical identification of speech". in McKenzie and Day, (eds*) Perceptual development in early infancy: problems and issues*. Laurence Erlbaum Associates, London, 1987.

3. Nittrouer, S. and Miller, M.E. "Developmental weighting shifts for noise components of fricative-vowel syllables", *J. Acous. Soc. Am.*, vol. 101, 1997, pp. 572-580.

4. Ohde, R.N., Haley, K.L., Vorperian, H.K. and McMahon, C.W. "A developmental study of the perception of onset spectra for stop consonants in different vowel environments", *J. Acoust. Soc. Am.*, vol. 97, 1995, pp. 3800-3812.

5. Parnell, M.M. and Amerman, J.D. "Maturational influences on the perception of coarticulatory effects", Journal of Speech and Hearing Research, vol. 21, 1978, pp 682-701.