PRODUCTION-PERCEPTION RELATIONSHIP IN THE VOICING CONTRAST FOR MEDIAL STOPS IN CHILDREN AND ADULTS.

Cecile T. L. Kuijpers
Max-Planck-Institut für Psycholinguistik, Nijmegen, The Netherlands

ABSTRACT

A phoneme identification experiment was carried out to investigate perception of the word-medial voicing contrast by Dutch four-year-old children, six-year-old children, twelve-year-old children, and adults. The data of this experiment are related to an earlier production study in which we studied word-medial closure durations. Production and perception of the voicing contrast are considered to develop synchronously in young children.

INTRODUCTION

The perception and production of the voicing contrast by children and adults has been investigated in numerous phonetic studies. Several acoustic cues play an important role depending on the position of the phoneme in the word. This paper will be concerned with one of the durational features related to the distinction, namely closure duration. In English, the Voice Onset Time (VOT) is the major acoustic cue to distinguish initial voiced and voiceless stops. In final position vowel duration differences cue the distinction. In Dutch, the intial contrast is characterized by a negative VOT (of approximately -70 ms) for voiced and a positive VOT (of 10-20 ms) for voiceless stops [1]. In final position, the contrast is neutralized because of a devoicing rule.

In many languages the word-medial voicing contrast is characterized by a difference in closure duration; a short closure duration for voiced and a long closure duration for voiceless stops. In some classical studies on perception of the medial voicing contrast the closure duration was manipulated (that is, a silent interval). Here, the same tendency showed up: short intervals were identified as a voiced stop, and long intervals as a voiceless stop [1], [2].

Experimental study of production and perception of the medial voicing contrast in young children is scarce. In [3],[4] we discuss a number of English production studies, and we report on our own production data of Dutch 4-, 6-, 12-year-old children and adults. In our production study the young children displayed a large variability and, consequently, voiced and voiceless stops were less clearly differentiated in the younger age groups than in the older age groups [4]. Those results will be integrated in the discussion section of this paper where we will describe shortly the relationship between the perception and production data of the Dutch children.

Nearly all developmental studies investigated perception of the *initial* and/or *final* voicing contrast. Perception of *intervocalic* stops by 6-year-olds and adults was examined by [5]. They used the naturally produced words 'petal' vs. 'pedal' and examined the interaction of VOT, closure duration, and preceding vowel duration. With respect to the role of closure duration *per se* the data indicated a difference between children and adults: in the children's data the phoneme boundary (50% crossover) was situated at approximately 110 ms and in the adult data at 130 ms.

The identification experiment presented in this paper deals with the perception of intervocalic stops by Dutch children and adults. The silent interval of naturally produced words was manipulated. On the basis of our production data, and on the basis of the literature on initial and final stops, we expect that the perceptual differentation of voiced and voiceless stops develops only gradually with age.

METHOD

Participants

Four age groups participated in the experiment: 4-year-olds (mean age 4;6), 6-year-olds (mean age 6;4), twelve-year-olds (12;3), and adults (age range 22-55). In each age group 15 listeners were tested, male and female, and all were monolingual speakers of Dutch. There was no known hearing deficiency in any of the listeners.

Material

Two minimal pairs of bisyllabic nonsense words were used, viz. the names 'Táppi'-'Tábbi' and 'Pátto'-'Páddo'. The sound structure of these names did not evoke any association with existing words or names. The initial vowel was always /A/, and the final vowels were /i/ and /o/ since Dutch names are often characterized by these endings. The four words were used to construct four different continua, henceforth the P-, B-, T-, and D-set of words.

Task

An identification task was set up as a game with two pairs of large puppets. The resemblence of the names was reflected in the resemblence of the puppets. The stimuli were put into carrier phrases in which the stimulus name was repeated at the end, such as 'Give the ball to Tappi..to Tappi'. By actually giving a ball to a puppet the child manifested its response in a two-alternatives-forced-choice labelling paradigm.

Manipulation

The four tokens were digitized on a Digital microVAX II computer using a 20 kHz sampling frequency. The initial vowel was set at a neutral value so that listeners would not be biased by its duration. The formant transitions were always unimpaired. The stop closure was replaced by different silent intervals with endpoints at 10 ms and 130 ms. We used 20 ms increments and created 7 different versions of each test word. The prominence of the voiceless burst was modified in order to yield consistent voiced and voiceless responses at the end-points; duration and intensity were modified on the basis of independent perceptual judgments.

Procedure

First, an auditory picture discrimination pre-test (ADIT) was carried out to ensure that the young children could discriminate the selected minimal pairs. Next, all children were carefully trained to learn the test words and to relate the names to the puppets. The final training set comprised 8 randomized trials ('Now show me..Patto) in which the children had to give at least 7 correct responses out of 8. In the main experiment each stimulus occurred twice and the children were tested in two sessions within two weeks; in each session 28 stimuli were presented. The stimuli were randomized and counter-balanced across subjects. The sentenced were separated by an inter-stimulus interval of 6.5 s. The 12-year-olds and adults were tested individually and they listened to the same sentences as the young children.

RESULTS

Four identification functions were calculated for each age group. The data were processed by a statistical analysis which considers the identification function as a cumulative probability distribution on Z-scores. Next, a linear regression analysis was carried out to fit the original data. Phoneme boundaries between groups were compared by means of a t-test. The slope of the functions (phoneme boundary width) were compared by means of a t-test on the coefficients of regression. We refer to [4] for a detailed description of the analyses.

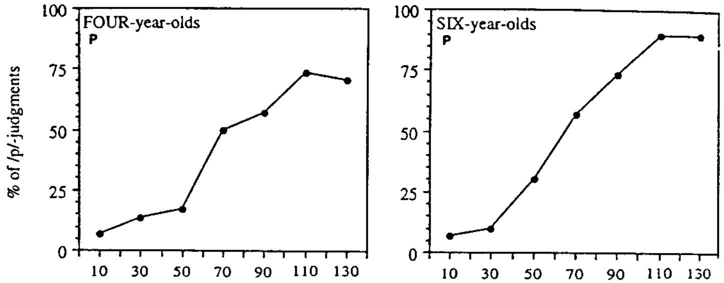

Figure 1. Identification functions of the /b/-/p/ continuum from the P-set of words for the 4-year-olds and the adults.

Phoneme boundary

The group identification functions were calculated for the four separate continua. Phoneme boundary corresponds to the 50% crossover, and to Z=0 in the transformed data. The percentage of voiceless jugments are given as a function of the silent interval. In Figure 1 we illustrate the identification functions of the 4-year-olds and the adults (P-set of words). In the P- and B-set of words no significant differences were found between any age group. In the T-set of words a significant difference was found between the 4-year-olds and 12-year-olds ($t=3.09$; $p<0.008$). No significant differences were found in the D-set of words. The data do show that all listeners need relatively long silent interval durations in the B/D-set of words in order to perceive a voiceless stop (see Figure 2). This results from the differences in formant transition information. Moreover, the young children's judgments mostly range between 10%-80% voiceless responses probably due to their inaccurate awareness of the durational characteristics of medial voiced and voiceless stops.

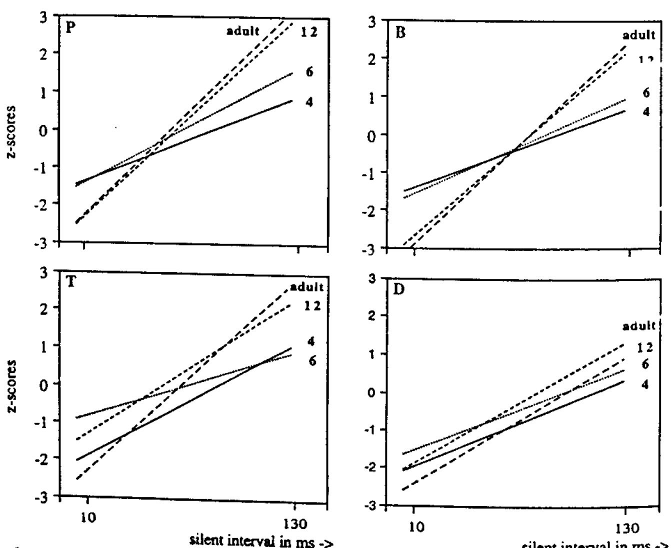

Figure 2. Modelled identification functions of Z-scores on stimuli with different silent intervals for the four age groups.

Phoneme boundary width

The phoneme boundary width corresponds to the 25%-75% interval on the fitted regression line. The comparison of the slopes of the regression lines in the P- and B-set of words indicated significant differences at $p=0.008$ between the groups 4-12, 4-adults and between the groups 6-12, 6-adults. In the T-set of words significant differences were found between the groups 4-6, 4-adults, 6-12, 6-adults. In the D-set of words, although not significant, there is a remarkable parallellism between the functions of the age groups 4 and 6 on the one hand, and 6 and 12 on the other hand.

We also examined the individual variability within each group. In short, the older age groups mainly displayed a continuous response mode, that is an upward ordering of voiceless judgments to stimuli with an ever increasing silent interval. The younger age groups displayed numerous discontinuous response modes, that is a less consistent perceptual behaviour. The upward ordering of voiceless judgments was frequently interrupted by voiced judgments.

DISCUSSION

We have indicated that the perception of the intervocalic stop voicing contrast by 4- and 6-year-olds deviates from the perception of that contrasts by 12-year-olds and adults. In comparison with the two latter groups, the young children need a relatively large difference in silent interval to perceive the distinction, and they have difficulty in determining which allophones come to be grouped in the same phonemic category. They do not categorize as consistently as the older listeners.

These findings concur with those reported by [6] who likewise concluded that the categorical perception of 6-year-olds is still unlike the adult norm. Surely, we presented stmuli in which the medial stop contrast was induced by only one acoustic parameter (silent interval) neglecting other parameters such as voicing. It is possible that younger children rely more than adults on multiple acoustic cues.

The findings reported in this paper are in agreement with our production data on closure duration differences in the medial voicing contrast. As in perception, the distinction becomes more and more attuned with age [4]. Both developments (production and perception) can be expressed in terms of a 'distincivity' which gradually increases with age. We assume that a parallel can be drawn between children's perception and production of the phonemic contrast.

ACKNOWLEDGEMENTS

This research was part of my Ph.D study financially supported by the University of Amsterdam and carried out at the Institute of Phonetic Sciences Amsterdam. I would like to thank L. Pols and F. Koopmans-van Beinum for their many helpful suggestions and discussions on the dissertation.

REFERENCES

[1] Slis, I. & Cohen, A. (1969a,b). "On the complex regulating the voiced-voiceless distinction I, II". *Language & Speech*, vol. 12, pp. 80-102, pp. 137-155.

[2] Lisker. L. (1978)."Rapid vs. Rabid. A catalogue of acoustic features that may cue the distinction". *Haskins Laboratory Status Report on Speech Reserach*, vol 54, pp. 127-132.

[3] Kuijpers, C. (1993a). "Temporal aspects of the voiced-voiceless distinction in speech development of young Dutch children". *Journal of Phonetics, vol.* 21, pp. 313-327.

[4] Kuijpers, C. (1993b). *Temporal co-ordination in speech development.* Un-published Doctoral dissertation, University of Amsterdam, The Netherlands.

[5] Allen, G. & Norwood, J. (1988). "Cues for the intervocalic /t/ and /d/ in children and adults. *Journal of the Acoustical Society of America* , vol. 84, pp. 868-875.

[6] Burnham, D., Earnshaw, L. & Clark, J. (1991). "Development of categorical identification of native and non-native stops: infants, children and adults". *Journal of Child Language*, vol. 18, 231-260.