

Perception of Voiced and Voiceless Stop Consonants in Three Year Old Children †

Little is known of the process by which children learn to perceive the phonological structure of speech in the manner of adults. Most 3 year old English-speaking children are able to produce /b/, /d/ and /g/, and /p/, /t/, /k/ and to understand some words in which these sounds occur. But this alone does not demonstrate that the children can utilize the acoustic properties of stop consonants as adults do, because of contextual cues (environmental or verbal) and gross perception of syllables or words containing the sounds.

Studies of adults by Lisker and Abramson, (1964) have shown that a major acoustic cue used to distinguish between the voiced and voiceless members of the alveolar stop consonantal pair /d/-/t/ in word initial position is the time relation between the burst of noise accompanying release of the occlusion and the onset of vocal cord vibration. (This relationship is hereafter referred to as voice onset time and abbreviated VOT). Experiments showed that if adults are presented with a series of synthetic consonant-vowel speech stimuli in which the vowel remains constant (/a/) and VOT is systematically varied, they tended to identify almost all stimuli in which vocal cord vibration began more than 30 msec. after release of occlusion (VOT = +30) as a /ta/, and almost all stimuli with VOT values of +20 or less as /da/. The purpose of the experiment under discussion was to determine to what extent 3 year old children are capable of adult perceptual performance

in identifying stop consonants occurring in these synthetic consonant-vowel stimuli.

Method

The subjects were 5 children aged 2-9, 2-10, 2-11, 3-0 and 3-6. The stimuli for the experiment consisted of two sets of synthetic alveolar stop consonant-vowel syllables recorded on tape. The stimuli used were constructed by Lisker and Abramson. The first or training tape consisted of a random ordering of 15 examples of each of two stimuli, one acceptable as /da/ in which vocal cord vibration onset preceded the release of occlusion by 30 msec. (VOT=-30) and one acceptable as /ta/ (VOT = +100). The second or test tape began with a randomly ordered set of 6 stimuli, consisting of three of each of the types just described. The remainder of the tape was a randomly ordered list consisting of 6 examples of each of 6 stimuli (VOT = -30, -5, +20, +45, +70 and +100). The stimuli were spaced at 11 second intervals on the tape.

The experiment was first designed to establish a different motor response to the two training stimuli /da/ and /ta/ by playing each to one of two spatially separated loudspeakers, and rewarding S for pulling a lever next to the speaker from which the sound came. The speakers were about 9' apart behind a cardboard room divider on which were painted two animal faces. A lever was situated in the "arm position" of each figure, and a reward chute in the mouth position. After the response to the spatial localization of the acoustic stimuli was established, the stimuli were moved to a central speaker. If S continued going to one lever on hearing /da/ and to the other on hearing /ta/, then the test tape containing VOT characteristics intermediate between those of the training stimuli would be played. Whether the S

heard them as /da/ or /ta/ could be determined by which lever he went to on hearing the stimulus.

After several sessions, this experimental plan was abandoned -- largely because the task was not easily assimilated by the children and the reinforcement contingencies, however varied, were not able to produce orderly participation in the task.

A second experimental procedure was then tried which proved considerably more effective. E showed the child a watch and said "this is a time" and then showed the child a dime and said "this is a dime." Then E had the child repeat the words, and made a game of having the child either name the object or point to the object E had named. When the S played this game correctly, E began to change her verbal labels from "dime" to "da-da-da-dime", and from "time" to "ta-ta-ta-time", and eventually to just "da-da-da" and "ta-ta-ta". When the S could appropriately verbalize these labels and point to their significates, the training tape was played and the child was asked to say whether he heard a /da/ or a /ta/. The training tape was repeated till the child responded without error. The test tape was then played immediately, and the child's responses recorded.

Results

Four of the five subjects gave almost "errorless" repetitions of the test tape. They responded to practically all of the stimuli with VOT values of -30, -5, and +20 as /da/ and to almost all other stimuli as /ta/, just as adult subjects do. Furthermore, their repetitions were made immediately after stimulus presentation. None of these subjects made more than three errors and the errors were not systematically distributed. The errors which did occur

were judged to be due to inattention.

The fifth child (2-9) showed a more variable response pattern which is summarized in Table 1. The S refused to respond to the last 4 stimuli on the tape and made a number of errors, all in the form of responding /ta/ to stimuli judged by adults to be /da/. However, the difficulty experienced in controlling the subject during the experimental session suggested that the main problem was one of motivation.

Discussion

The experiment leaves little doubt that children of this age are capable of perceiving one of a pair of alveolar stop consonantal stimuli 25 msec. apart in VOT as /da/ and the other as /ta/, as adults would. It remains to be seen whether children of this age can distinguish between stimuli only 10 msec. apart, at the "phoneme boundary" between /t/ and /d/ as adults do. However, the routine way in which the subjects made correct responses, and the typical short latencies of response, suggest that the children may be able to match more closely adult performance than could be demonstrated with the stimuli chosen for this experiment. Their performance also suggests that considerably younger children may be able to use small differences in VOT values to discriminate between voiced and voiceless stop consonants when other cues to their identity are not available.

Thus the three year old children categorize one class of linguistically relevant nonsense syllables approximately the way adults do. It would seem, therefore, that perception of the critical acoustic properties of the message contributes little to the comprehension difficulties shown by children of this age.

J. M. S. Winterkorn, R. F. MacNeilage, M. S. Preston

Table 1. Responses of one child (age 2-9) to the first 32 test stimuli. The variability of responses by this child was atypical.

Stimulus (VOT value)	Response		
	da	ta	total
-30	2	4	6
- 5	5	1	6
+20	2	3	5
+45	0	6	6
+70	0	5	5
+100	0	4	<u>4</u>
			<u>32</u>

References

Lisker, L. and Abramson, A. S. A cross-language study of voicing in initial stops: acoustic measurements. Word (1964) 20, 384-422.

Acknowledgments

We wish to thank Drs. Lisker and Abramson for making their synthetic stimuli available to us, and Dr. Richard P. Youtz for making the facilities of the Barnard College Psychology Department available to us for testing the subjects.

† The research reported here is part of a cooperative effort with the Neurocommunications Laboratory, the Johns Hopkins University Medical School.