

A Brief Report on a Study of Imitation of Synthetic Speech Stimuli by Children †

A considerable amount of research in the past ten years has examined the perception of speech sounds by adults. One striking finding that has come out of this research is the categorical nature of the perception of the stops and other consonants. For instance, discrimination and identification studies have shown that discrimination along an acoustic continuum for /d/ and /t/ is noticeably sharper across the phoneme boundary than within either phoneme class (1). This brief report describes a first step towards extending these studies to children under the age of six years.

Some preliminary work suggested that young children would readily imitate speech sounds presented over a loud-speaker. It seemed reasonable, then, to gather more data using this method. Eighteen items from a series of 31 synthetic speech stimuli (CV syllables) made on the Haskins Laboratories Voback synthesizer were mixed in random fashion with 23 real speech CV syllables. The complete tape was then presented a total of three times over a loud-speaker system installed in a double wall IAC booth to children (speakers of American English) between the ages of three and six years. The children were asked to reproduce as best they could each stimulus. Tape recordings were made of their responses. The synthetic speech sampled points along an acoustic continuum (voice onset time). Recent research (2) has shown that voice onset time is an effective measure for sorting the stops into phonemic categories. The synthetic stimuli had

previously been identified by native speakers of American English as /d/ or /t/. The real speech was included to make the task less boring for the children.

At the present time, imitation data has been collected from 12 children between five and six years of age. Four children gave a single response to all synthetic stimuli. The other eight responded categorically, that is, they gave one of two responses to each synthetic stimulus. Four of these last eight responded with /d/ or /t/; the other four gave voiced sounds as one category (/d/ or /z/), and voiceless sounds (/h/) or vowels (/a/) as the other category. The boundary between categories for these last eight children varied slightly around 35 msec. voicing lag, the approximate boundary for the adult listeners.

H.G. Yeni-Komshian

M.S. Preston

J.K. Cullen

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

- (1) Liberman, A.M., Harris, K.S., Kinney, J.A., and Lane, H. The discrimination of relative onset-time of the components of certain speech and nonspeech patterns. J. Exp. Psychol. (1961) 61, 379-388.
- (2) Lisker, L., Abramson, A.S. A cross-language study of voicing in initial stops: acoustical measurements. Word (1964) 20, 384-422.