ABSTRACT

On the Nature of Categorical Perception of Speech Sounds

David Bob Pisoni
Haskins Laboratories, New Haven

Current theories of speech perception emphasize that the perception of speech sounds may involve processes that are in some way basically different from the processes involved in the perception of other sounds. One of the findings which has been cited as evidence for a special mode of perception is the differences in perception between synthetic stop consonants and steady-state vowels. Stop consonants have been found to be perceived in a categorical mode, unlike other auditory stimuli. Discrimination is limited by absolute identification. Listeners are able to discriminate stimuli drawn from different phonetic categories but cannot discriminate stimuli drawn from the same phonetic category, even though the acoustic distance between stimuli is comparable. On the other hand, steady-state vowels have been found to be perceived continuously. Discrimination is independent of category assignment. Listeners are able to discriminate many more differences than would be predicted on the basis of absolute identification.

The primary goal of the present investigation was to examine the differences between categorical and continuous perception and to evaluate three different explanations for the phenomena of categorical perception. Six experiments dealing with the identification and discrimination of synthetic speech sounds were conducted to determine the nature of categorical perception. The first experiment replicated the original findings on the differences in perception between consonants and vowels reported by investigators at Haskins Laboratories. Perception of stop consonants was found to be "nearly categorical" in the sense that listeners tend to discriminate pairs of stimuli only to the extent that they identify them as different. Perception of steady-state vowels was found to be more "nearly continuous" in the sense that the same listeners discriminate many more intraphonemic differences than they identify absolutely.

The second experiment attempted to assess the effects of discrimination training with non-speech stimuli on categorical perception. The results indicated that there were large individual differences among Ss and that no definite conclusions could be drawn about the effects of discrimination training in producing categorical perception with non-speech stimuli.

The third experiment considered an explanation of categorical perception in terms of the auditory and phonetic processes involved in speech discrimination tasks. It was found that steady-state vowels tend to be perceived more categorically at brief stimulus durations. The results also confirmed predictions derived

*Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy, The University of Michigan.
+Now at Indiana University, Bloomington.
from a model proposed by Fujisaki which suggested that auditory short-term memory may be involved in speech discrimination.

The fourth and fifth experiments were concerned with comparing a new discrimination procedure, the four-interval test of paired similarity (4IAX), with the traditional ABX discrimination test. It was found that substantial differences in discrimination may be obtained with the 4IAX procedure as compared with the ABX for vowels, while less marked differences in discrimination may be obtained with consonants.

The sixth experiment tested the hypothesis that consonants and vowels differ in the degree to which auditory short-term memory is employed in their discrimination. The results of a delayed comparison recognition memory task indicated that accuracy of discrimination for vowels both within and between phoneme boundaries was related to the magnitude of the comparison interval. In contrast, discrimination of stop consonants remained relatively stable both within and between phoneme boundaries.

The results of this investigation suggested that the major differences between categorically and continuously perceived speech stimuli are related to the differential availability of auditory short-term memory for the acoustic cues distinguishing different classes of speech sounds. For highly encoded speech sounds such as stop consonants, within-category discrimination is so poor as to suggest that information other than a binding phonetic categorization is unavailable to the listener for use in discrimination.