DOCUMENT RESUME

ED 098 647

CS 500 896

AUTHOR Uselding, Douglas K.; Molfese, Dennis L.

TITLE Adult Speech Perception: Asymmetrical Effects in Categorical Perception.

PUB DATE Sep 74

NOTE 13p.; Paper presented at the Annual Meeting of the American Psychological Association (82nd, New Orleans, Louisiana, Aug. 30-Sept. 3, 1974)

EDRS PRICE MF-$0.75 HC-$1.50 PLUS POSTAGE

DESCRIPTORS *Adults; *Auditory Perception; *Language Research; *Phonemes; *Speech; Speech Compression; Speech Skills

ABSTRACT

To measure the symmetry of adult categorical phoneme perception, 10 adult male undergraduate students enrolled in an introductory psychology class were the subjects for this study as part of their course requirements. The stimuli used in this study were prepared at Haskins Laboratories by means of a parallel resonance synthesizer and computer. The bilabial stop consonants generated had voice onset times of 0, 20, 40, 60, and 80 milliseconds. Each consonant was followed by the vowel "a." Each subject was tested individually in a sound dampened room. Subjects were seated at a table facing two speakers, one located 90 inches from the subject, the other located 6 inches directly behind the first. A response key was positioned directly in front of the subject. The results replicated those of Liberman and other researchers (1967) which demonstrated that although voice onset time is a continuum, adult perception along that continuum appears categorical. (SW)

ED 098647





# Adult Speech Perception: Asymmetrical Effects in Categorical Perception

Douglas K. Uselding and Dennis L. Molfese
Department of Psychology
Southern Illinois University at Carbondale
Carbondale, Illinois 62901



Paper presented at the 82nd Annual Meeting
of the
American Psychological Association,
New Orleans, 1974.

For nearly three decades researchers have attempted to establish which acoustic cues are used by language users to discriminate between speech sounds. One such acoustic cue has been identified by Lisker and Abramson (1964) as voice onset time (VOT)-the time between the release burst and the onset of laryngeal pulsing. VOT has been shown to be a sufficient cue for the discrimination between a number of American English voiced and voiceless stop consonants such as /b/ and /p/, /d/ and /t/, and /g/ and /k/.

Liberman, Cooper, Shankweiler and Studdert-Kennedy (1967) have demonstrated that although VOT is a continuum, adult perception along that continuum appears to be categorical. That is, adults can readily discriminate two sounds that lie across the phoneme boundry of +30 msec VOT (e.g., /ba/ versus /pa/) but perform only at chance level when asked to discriminate between sounds which represent acoustic variations within the same phoneme category (e.g., /ba/ with a 0 msec VOT versus /ba/ with 20 msec VOT). Since adult listeners cannot hear intraphonemic differences, all variations within the phoneme category are assigned the same label by adults.

Consequently, adult perception for speech sounds appears to be qualitatively different from their perception of non-speech sounds such as noises or musical tones. More recently, several researchers have suggested that this ability to categorically perceive speech sounds is already present in the early months of life.

Eimas, Siqueland, Jusczyk and Vigorito (1971) have reported finding categorical perception in one and four-month-old infants in a study which employed non-nutritive sucking as a response measure. Infants, like adults, were able to discriminate between the voiced and voiceless stop consonants /b/ and /p/ but were unable to discriminate between sounds within a phoneme category. Similar findings have been reported by Morse (1971), Trehub and Rabinovitch (1972), and Treub (1973).

In addition to these behavioral data, Molfese (1972), using evoked potential recording procedures reported that infants, children and adults all responded to speech stimuli with larger amplitude electro cortical activity located in different areas of the brain than to non-speech auditory stimuli. The brain appears to be specialized to differentiate speech materials from other sounds early in life.

Thus, some perceptual language skills appear to be available to the infant from birth. Furthermore, these skills are comparable, both behaviorally and neurologically to those of adult language users.

This position has been questioned by Butterfield and Cairns (1974). They note that infant categorical perception of phonemes is asymmetrical. Infant discrimination across the phoneme boundry is better if the VOT values increase from one stimulus to the next (e.g., 20 to 40 msec) than if they decrease (e.g., 40 to 20 msec). If infant perception is asymmetrical while adult perception is not, this behavioral difference may indicate that infant and adult speech perception skills are not as comparable as is currently thought.

A study was conducted by the present authors to measure the symmetry of adult categorical phoneme perception.

BEST COPY AVAILABLE

## Method

Subjects:

Ten adult male undergraduate students enrolled in an Introductory Psychology class participated in this study as part of their course requirements.

Stimuli:

The stimuli employed in the present study were prepared at Haskins Laboratories by means of a parallel resonance synthesizer and computer. The bilabial stop consonants generated had voice onset times of 0, 20, 40, 60, and 80 milliseconds (the interval between release burst and laryngeal pulsing or voicing). Each consonant was followed by the vowel /a/. All stimuli were 300 msec in duration.

Thirty-nine randomly ordered sets of ten stimuli were constructed. Within each set the stimuli were presented at a rate of one per second. A seven second gap separated each set.

Fifteen of the sets contained a series of ten identical stimuli while the remaining 24 sets contained VOT changes of 20 msec in either the fourth, sixth or eighth stimulus. For example, the first five stimulus sounds in one series of ten stimuli might be bilabial stop consonants with 20 msec VOTs while the second five sounds of the same series would be bilabial stop consonants with 40 msec VOTs. Twelve of these "VOT change" sets were characterized by an increase in VOT (e.g., 0 to 20, 20 to 40, etc.) and 12 contained a decrease (e.g., 20 to 0, 40 to 20, etc.).

Procedure:

Each S was tested individually in a sound dampened room. Subjects were seated at a table facing two speakers: one located 90 inches from the subject, the other located 6 inches directly behind the first. A response key was positioned directly in front of the subject.

The speech stimuli were delivered through the farthest speaker at 83 db (SPL) as measured at the subject's ear. White noise was presented continously at 65 db (SPL) through the front speaker to mask background ambient noise. The subjects were instructed to listen to each set and to press a key to indicate the position of a perceived change. The subjects were told that some sets would contain obvious, subtle or no changes.

## Results

To insure the quality of the stimuli synthesis, 15 additional adult subjects were administered the labeling and ABX discrimination tasks of Liberman et al (1967). Figure one presents the labeling data that were thus obtained.

<u>Insert Figure One</u>

As can be seen from this figure, the stimuli were identified as /ba/ when the VOT values were 0 or 20 msec. When VOT values were 40, 60 or 80 msec, the stimuli were identified as /pa/. This finding replicates those reported by Liberman et al (1967).

Figure two presents the discrimination data obtained using an ABX paradigm.

<u>Insert Figure Two</u>

This figure indicates the percentage of correct responses illicited by various stimuli changes. The abscissa values present the VOT values for the different stimuli.

The abscissa values are the starting points of ascending VOT changes (e.g., 0 is the starting point of a 0 to 20 VOT change). Each abscissa value is also the terminal point of a descending VOT change (e.g., 0 is the terminal point of a 20 to 0 VOT change). Thus, each abscissa has two ordinate values, one for ascending and one for descending VOT changes.

The ABX data clearly indicate categorical perception. Correct responding drastically increased when VOT changes transgressed the +30 msec boundry. Perception within phoneme categories was at chance level. In addition no asymmetry was found. Perception was equal for ascending and descending VOT changes.

In general these findings replicate those reported earlier by Liberman et al (1967).

The data generated by the ten-item 'change detection" paradigm described earlier were analyzed using a repeated measures ANOVA. Alpha was set at .01 for all statistical analyses.

The analysis indicated significantly better discrimination for changes across the phoneme boundry than within adult phoneme phoneme categories (F= 21.036). There was also a significant interaction between place of VOT change (across or within boundry) and direction of change (asending or descending) (F= 5.375).

Figure three presents the discrimination function obtained from the change detection paradigm.

Insert Figure Three

This figure is read in the same way as was figure two. This figure indicates that the significant effect of place of VOT change derives from the fact that perception is significantly better across the phoneme boundry than within categories.

The significant place x direction interaction indicates an asymmetry. Scheffe's test was computed to test for significant differences between the 8 points on figure three. This ana. s indicated that perception of VOT changes that descended across the boundry (from 40 to 20) was significantly better than any other type of VOT change including changes ascending across the boundry (from 20 to 40 msec). The only other significant differences were between ascending across boundry changes (from 20 to 40 msec) and descending within category changes (from 40 to 60 and from 60 to 80 msec). The responding to the ascending VOT change from 40 to 60 msec was at 10% which is precisely at chance level. It is noted that the ascending across boundry VOT change (from 20 to 40 msec) approached but did not reach a level of statistically significant difference when compared to this chance level.

## Discussion

A task was employed to assess categorical perception different from the standard ABX paradigm Chance responding in the standard ABX task is at the 50% level as compared to only 10% in the present task. Using this task, we were able to replicate the adult categorical discrimination effect. However, the data indicate that the effect is asymmetrical. Perception is better if a decrease is made in the VOT across the boundry than if it increases.

Butterfield and Cairns (1974) report an asymmetry of categorical perception for infants in the opposite direction to that reported here with adults. The data of those researchers indicate superior perception for ascending VOT changes from -20 msec to +80 msec. The lack of agreement between the present study and that of Butterfield and Cairns may stem from several sources. First, the present study employed 20 msec VOT changes as opposed to a 100 msec change. Whether different steps in the VOT change will affect the asymmetry is an emperical question that deserves further study. A second difference is that the Butterfield and Cairns study measures discriminations between prevoicing and delayed voicing. It is not clear to the present authors that this procedure is necessarily comparable to comparing different voicing lags. The former discrimination may assess different abilities. A third possible source of the difference between results may be the use of different response measures. Infant discriminations were inferred from changes in high amplitude sucking. The adult subjects in the present study were required to respond as if operating in a signal detection situation. Consequently, the grossly different response measures may account for the discrepancy in the data. A final point is that the data may reflect actual developmental differences in the perception of speech cues.

In any event, the existence of asymmetry in categorical perception by adults and children may indicate that the skills employed by each are at some level comparable to each other. The data tend to support a position for an innate basis for language.

A crucial qualification needs to be expressed. The phenomenon of categorical perception may be dependent on acoustic and physiological variables as well as on linguistic significance. This is indicated by the asymmetry which has been found. The distinction between two phonemes, if solely determined by the linguistic significance of each, would be equal regardless of which phoneme is presented first Thus, hearing /ba/ or /pa/ first should not significantly alter the discrimination between them.

## References

Butterfield, E.C., and Cairns, G.F. Whether infants perceive linguistically is uncertain, and if they did, its importance would be equivocal. In R.L. Schiefelbusch and L.L. Lloyd (Eds.) Language Perspectives -- Acquisition, Retardation, and Intervention. Baltimore: University Park Press, 1974.

Eimas, P.D., Siqueland, E.R., Jusczyk, P. & Vigorito, J. Speech perception in infants. Science, 1971, 171, 303-306.

Liberman, A.M., Cooper, F.S., Shankweiler, D.D. & Studdert-Kennedy, M. Perception of the Speech Code. Psychological Review, 1967, 74, 431-461.

Lisker, L. and Abramson, A.M. A cross-language study of voicing in initial stops: acoustical measurements. WORD, 1964, 20, 384-422.

Molfese, D.L. Cerebral asymmetry in infants, children and adults: auditory evoked responses to speech and non-speech stimuli. Paper presented to the 84th meeting of the Acoustic Society of America, Miami, December 1,1972.

Morse, P.A. Speech perception in six week old infants. Paper presented to meeting of Society for Research in Child Development, Minneapolis, 1971.

Trehub, S.E. and Rabinovitch, M.S. Auditory Linguistic Sensitivity in early infancy. Developmental Psychology, 1972, 6, 74-77.

Trehub, S.E. Infants Sensitivity to vowel and tonel contrasts. Developmental Psychology, 1973, 9, 91-96.

Figure 1

Identification of Synthetic Speech Stimuli

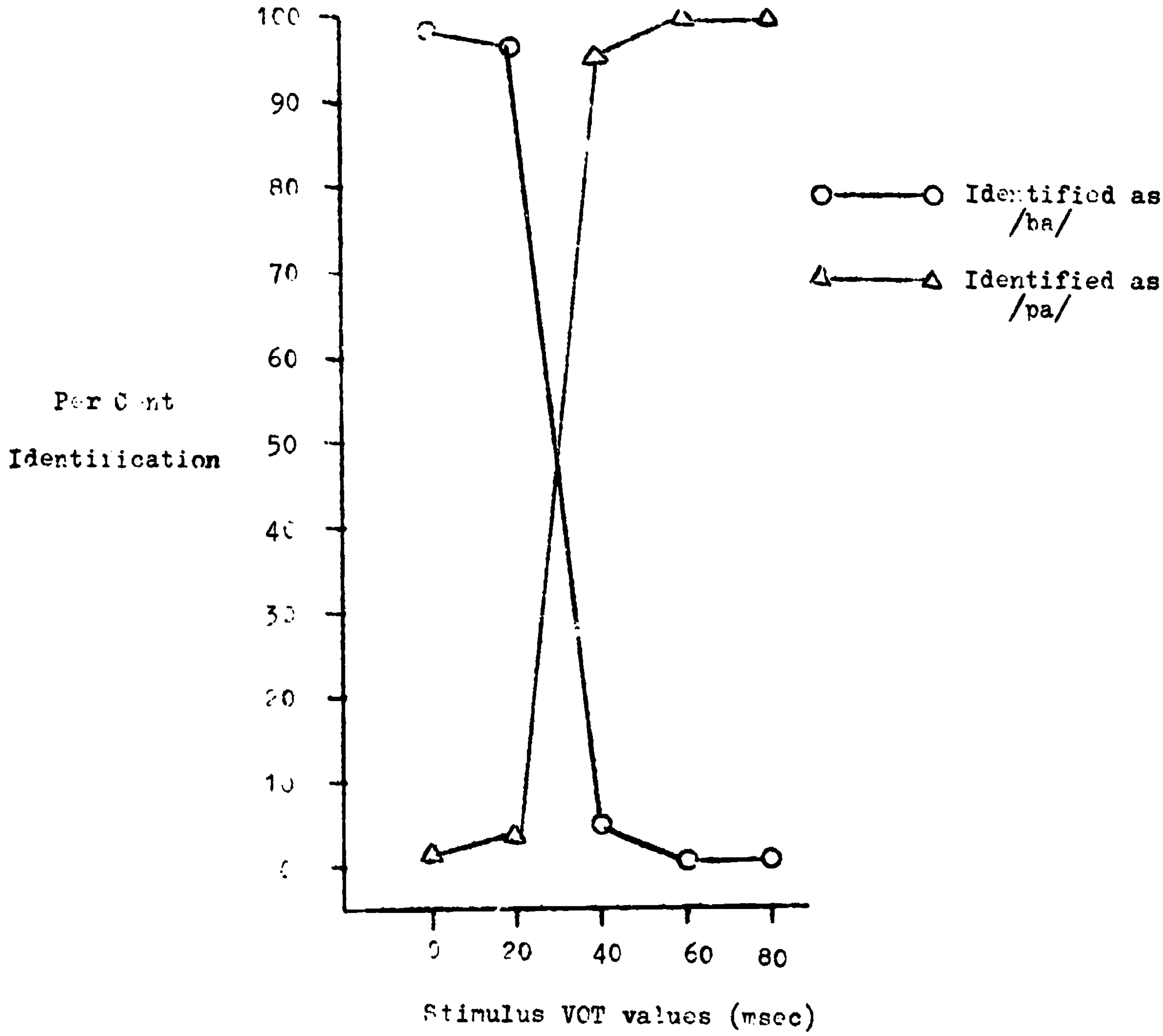

Figure 2

Discrimination of Stimuli using an ABX paradigm

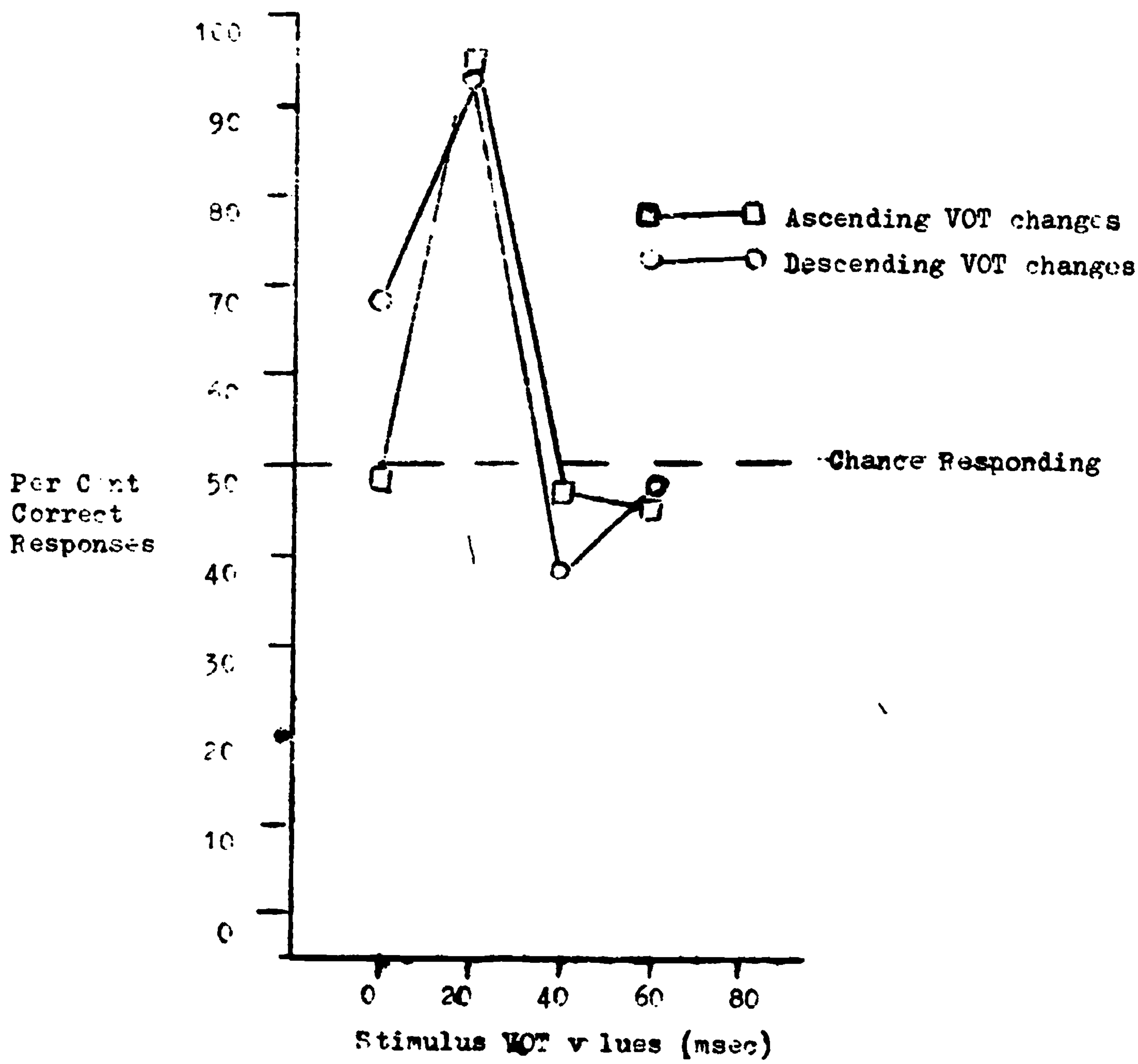

Figure 3

Discrimination of Stimuli using a 10-item series paradigm