DOCUMENT RESUME

ED 024 950 48 AL 001 660

By-Kopp, James
A New Test for Categorical Perception.
Michigan Univ., Ann Arbor. Center for Research on Language and Language Behavior.
Spons Agency-Office of Education (DHEW), Washington, D.C. Bureau of Research.
Bureau No-BR-6-1784
Pub Date 1 Sep 68
Contract-OEC-3-6-061784-0508
Note-10p.; Report included in Studies in Language and Language Behavior, Progress Report No. VII.
EDRS Price MF-$0.25 HC-$0.60
Descriptors-*Auditory Perception, Aural Stimuli, Psychoacoustics, *Psycholinguistics
Identifiers-*Categorical Perception, Synthetic Speech Stimuli

Synthetic speech stimuli which varied along a continum defined in terms of the relative onset time of the first formant and which are characteristically associated with the phonemes /do/ and /to/ were presented to two groups of subjects for labelling. The stimuli presented to each group were identical except for a one-step shift on the stimulus continuum. It was found that the response probabilities over the stimuli common to the two stimulus sets were essentially identical for the two groups Subjects did not evidence a shift in labelling judgments concomitant with a shift on the stimulus continuum. Such a shift would occur if subjects were simply partitioning the continuum psychophysically. It is concluded that categorical perception was observed and that the stimulus-shift paradigm may be a new technique for isolating the categorical perception effect. (Author/DO)

A NEW TEST FOR CATEGORICAL PERCEPTION[1]

James Kopp[2]

Center for Research on Language and Language Behavior

The University of Michigan

Synthetic speech stimuli which varied along a continuum defined in terms of the relative onset time of the first formant and which are characteristically associated with the phonemes /do/ and /to/ were presented to 2 groups of subjects for labelling. The stimuli presented to each group were identical except for a one-step shift on the stimulus continuum. It was found that the response probabilities over the stimuli common to the 2 stimulus sets were essentially identical for the 2 groups. Subjects did not evidence a shift in labelling judgments concomitant with a shift on the stimulus continuum. Such a shift would occur if subjects were simply partitioning the continuum psychophysically. It is concluded that categorical perception was observed and that the stimulus-shift paradigm may be a new technique for isolating the categorical perception effect.

It has been shown recently that the perception of certain language sounds is of a peculiarly categorical nature. Identification data collected from human subjects for synthesized speech sounds over a variety of phonemic contrasts reveal that perception is discontinuous, i.e., the stimuli are sorted into one of two categories on pretty much an all-or-none basis (Liberman et al., 1957; 1961.) The data taken to define categorical perception are in the form of mutually overlapping frequency-of-identification gradients which resemble complementary step functions. In addition to the discontinuous nature of these identification functions, it has been shown that: a) there is an abrupt increase in latency of the identification responses at the point where the two categories adjoin, b) the topographies of the verbal identification responses tend to be discrete--are in themselves categorical, and c) stimulus discriminability on a pair-comparison basis is better at the category boundary than any any point within the category (see Figure 1). Liberman has used such data to construct a theory which accounts for the categorical quality of speech perception in

terms of the discreteness of the articulatory movements necessary for the production of certain speech sounds (for a review of the data and the theory see Lane, 1965).

Insert Figure 1 About Here

Others (Cross & Lane, 1962; Cross, Lane, & Sheppard, 1965; Lane, 1965) have observed that the paradigm for categorical perception is essentially a generalization paradigm in which two stimuli on a given continuum exercise a great deal of control over two incompatible responses. When this notion is tested by presenting stimuli intermediate to two well conditioned controlling stimuli, irrespective of the continuum, the finding is that the probability of each response is high and constant over the stimulus range immediate to its respective controlling stimulus, with an abrupt decrement to zero over the stimulus range immediate to the stimulus which controls the other response. Generalization gradients obtained in this way resemble the complementary step functions obtained for the identification of synthetic speech stimuli. In these other studies, the collateral data taken by Liberman to define categorical perception--latency peaks, discrimination enhancement, and discreteness of response topography--have also been shown to occur. These facts have led to the conclusion that categorical perception is not unique to speech stimuli and that an extensive discrimination history involving two mutually incompatible responses is paradigmatic of categorical perception (Lane, 1967).

Categorical perception is of interest to a student of psychophysics: although it resembles simple stimulus equisection, it has additional properties. Of special significance is the enhancement of discrimination at category boundaries. This discontinuity is not characteristic of discrimination with such stimulus attributes as amplitude, duration, or frequency. In fact, the relative monotonicity of changes in discrimination probabilities on these latter continua is the basis for the construction of Fechnerian scales for the growth of sensation.

Now, another property of equisection is the independence of equisection judgments from the absolute values of the stimuli to be categorized. Bisection of a 10 db interval of sound intensity will be at a point about 6 db above the lower intensity, whether the values to be bisected are 60 to 70 db, 65 to 75 db, or 70 to 80 db. This observation leads to an interesting question. What are the properties of equisection on linguistic continua? Will the bisection function for synthetic speech stimuli ranging over some distance, x, on a continuum be the same as or different from the bisection function for stimuli ranging over the same distance but at a slightly different locus on the continuum? In other words, will the positioning of the perceptual midpoint on a synthetic speech continuum be independent of the absolute stimulus values, as it would be in the case of simple bisection, or will it stay relatively fixed and thus reflect the categorical nature of the perception of phonemes?

Method

Synthetic speech stimuli which are typically identified as continuous with respect to the two English phonemes /do/ and /to/ were presented to two groups of subjects for identification. Subjects served individually in sessions lasting approximately 30 minutes. The stimuli were synthesized at Haskins Laboratories on a device known as the Pattern Playback (Liberman et al., 1961). Each group was presented with one of two sets of stimuli. The sets were identical except for a one-step shift of the relative onset time of the energy in the first and second formants of the synthetic speech signal. There were six stimuli in each stimulus set. Each stimulus was presented an equal number of times in a triad, sandwiched, as it were, between the two stimuli defining the range of the stimulus set. The subjects rated each stimulus as identical to either the first or third member of the triad. Instructions to the subject were as follows:

> In this experiment you will judge language sounds. You will hear two kinds of sounds--/do/ and /to/. They

> will be presented to you in triads. In each triad, the first stimulus will always be /do/ and the third stimulus will always be /to/. The middle stimulus will be either /do/ or /to/. After each triad depress the left hand switch if you think the middle stimulus is more like the first stimulus (/do/) or the right hand switch if you think the middle stimulus is more like the third stimulus (/to/). Be sure to respond each time but make only one response after each triad. Do not try to make an equal number of responses on each switch as /do/ and /to/ do not necessarily appear with equal frequency in the middle position. Have you any questions?

The tape recorded stimuli were presented in random order to the subject through a set of audiometer headphones (Sharpe, Type A, model HA-10). Responses were recorded automatically on a multiple channel event recorder when the subject depressed one of the two switches to indicate his response. Each stimulus triad was presented 10 times. There were 10 Ss in each group.

Results and Discussion

From Figure 2 it can be seen that the two group identification functions, when plotted over the range of stimuli used in the total study, are essentially congruent; response probabilities are about the same for both groups at any given absolute cutback value. The difference between the intersection points for the two sets of data is on the order of 1 msec., a difference not at all close to the 10 msec. shift by which the two stimulus sets differed. So close a matching would not be expected if the stimuli had been tone intensities, lengths of lines, etc.

Insert Figure 2 About Here

The abrupt disparity observed at one point on the continuum might be interpreted as indicative of a shift in the identification function in accord with the shift in the stimulus set. This interpretation is weakened, however, when the data from only the most reliable subjects, those subjects for whom there were no abrupt discontinuities in the identification function, are analyzed (see Figure 3). These subjects yielded gradients which tended to be

superimposed even more than those of the groups as a whole. Yet one would expect any shift due to context to be most readily apparent with the reliable subjects, since it seems reasonable to assume that their behavior was most under stimulus control.

Insert Figure 3 About Here

A more likely explanation for the slight disparity between the two gradients is in terms of response alternation. Assuming that stimulus control dictated that the response probabilities be identical for the two groups, then one of the groups would find itself emitting many more of one kind of response than of the other. The tendency of this group would be then occasionally to emit a few responses of the second kind, irrespective of which stimulus was presented--a tendency to balance things up, so to speak. The instructions were worded so as to minimize such a possibility, but evidently to no avail.

Conclusions

The present study suggests that a test using identification probabilities and a shift in stimulus context might point to differences in categorical responding in situations involving speech stimuli as against situations in which some simple stimulus continuum is employed. The data obtained show that identification functions from two groups of subjects for two overlapping sets of synthetic speech stimuli tend to be superimposed, although the effect was contaminated by a tendency toward response alternation on the part of the subjects. A study is now underway with the stimulus triads arranged in both ascending and descending order. Eliminating the tendency to alternation should lead to better congruence of the two functions.

References

Cross, D. V., & Lane, H. L. On the discriminative control of concurrent responses: The relations among response frequency, latency, and topography in auditory generalization. Journal of the Experimental Analysis of Behavior, 1962, 5, 487-496.

Cross, D. V., Lane, H. L., & Sheppard, W. C. Identification and discrimination functions for a visual continuum and their relation to the motor theory of speech perception. Journal of Experimental Psychology, 1965, 70, 63-74.

Lane, H. L. The motor theory of speech perception: A critical review. Psychological Review, 1965, 72, 275-309.

Lane, H. L. A behavioral basis for the polarity principle in linguistics. In K. Salzinger (Ed.), Research in verbal behavior and some neurophysiological implications. New York: Academic Press, 1967. Pp. 79-98.

Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. The discrimination of speech sounds within and across phoneme boundaries. Journal of Experimental Psychology, 1957, 54, 358-368.

Liberman, A. M., Harris, K. S , Kinney, J. H., & Lane, H. L. The discrimination of relative onset time of the components of certain speech and non-speech patterns. Journal of Experimental Psychology, 1961, 61, 379-388.

Footnotes

[1]The research reported herein was performed in part pursuant to Contract OEC-3-6-061784-0508 with the U. S. Department of Health, Education, and Welfare, Office of Education, under the provisions of P. L. 83-531, Cooperative Research, and the provisions of Title VII, P. L. 85-864, as amended. This research report is one of several which have been submitted to the Office of Education as Studies in Language and Language Behavior, Progress Report VII, September 1, 1968.

[2]Now at Claremont Graduate School.

[3]An abrupt discontinuity was defined as any non-monotonic change in the identification function in the region of transition from 100% to 0% response probabilities.

Figure Captions

Figure 1. Observed distributions of discrimination accuracy and identification probability and latency for the /do/-/to/ continuum. From Lane (1967).

Figure 2. Identification functions for two groups of Ss for two overlapping sets of synthetic speech stimuli. Each point is the arithmetic mean percentage across all Ss for /do/ responses only. The /to/ response percentages are complementary.

Figure 3. Identification functions for the best discriminating Ss in each of two groups of Ss for two overlapping sets of synthetic speech stimuli. The criterion for "best discriminating" was a lack of any abrupt discontinuity in a S's identification function. Each point is the arithmetic mean percentage across all selected Ss for /do/ responses only. The /to/ response percentages are complementary.

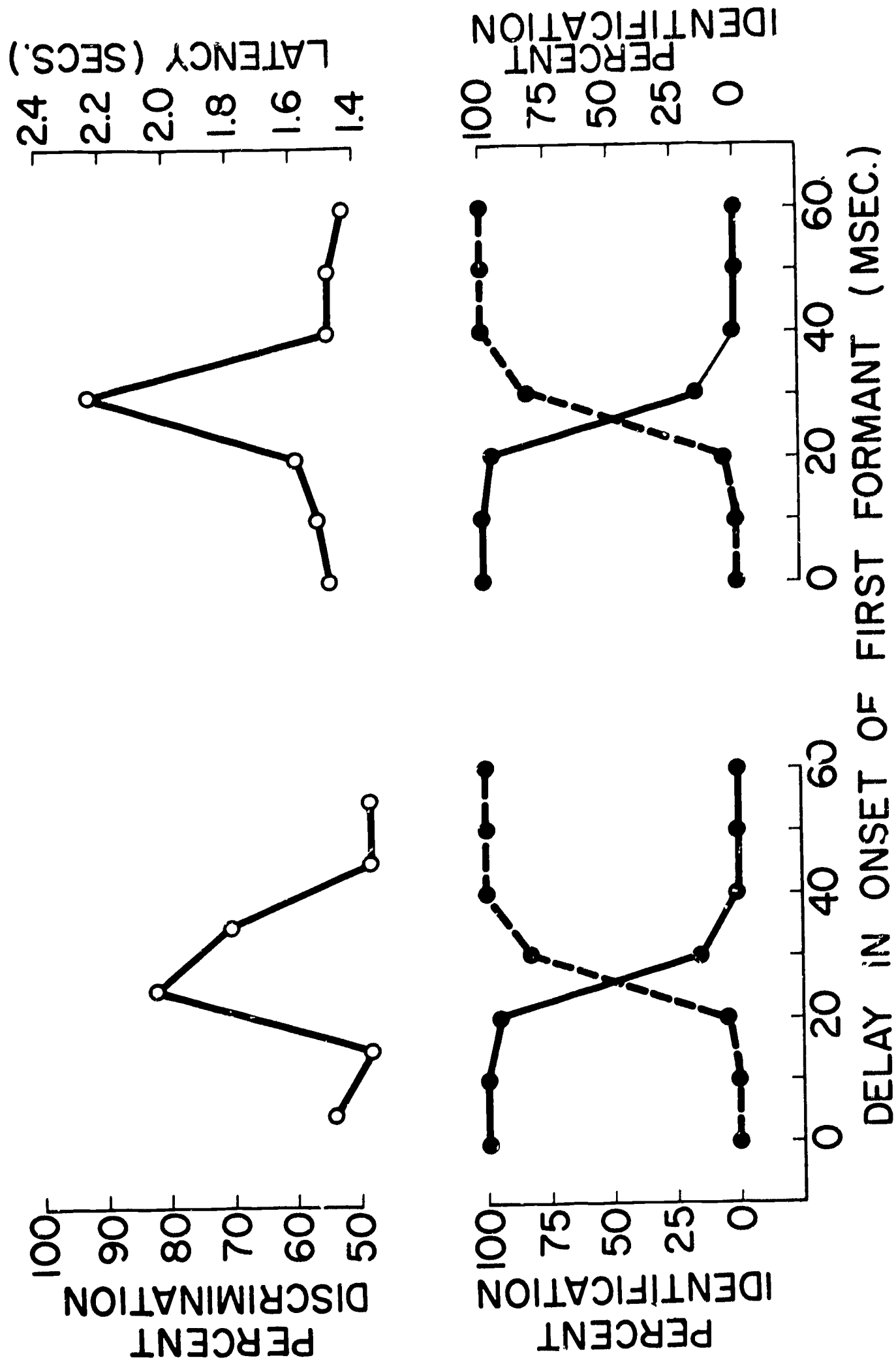
LATENCY (SECS.)
2.4
2.2
2.0
1.8
1.6
1.4
PERCENT IDENTIFICATION
100
75
50
25
0
0
20
40
60
FIRST FORMANT (MSEC.)
PERCENT DISCRIMINATION
100
90
80
70
60
50
PERCENT IDENTIFICATION
100
75
50
25
0
0
20
40
60
DELAY IN ONSET OF

Figure 1

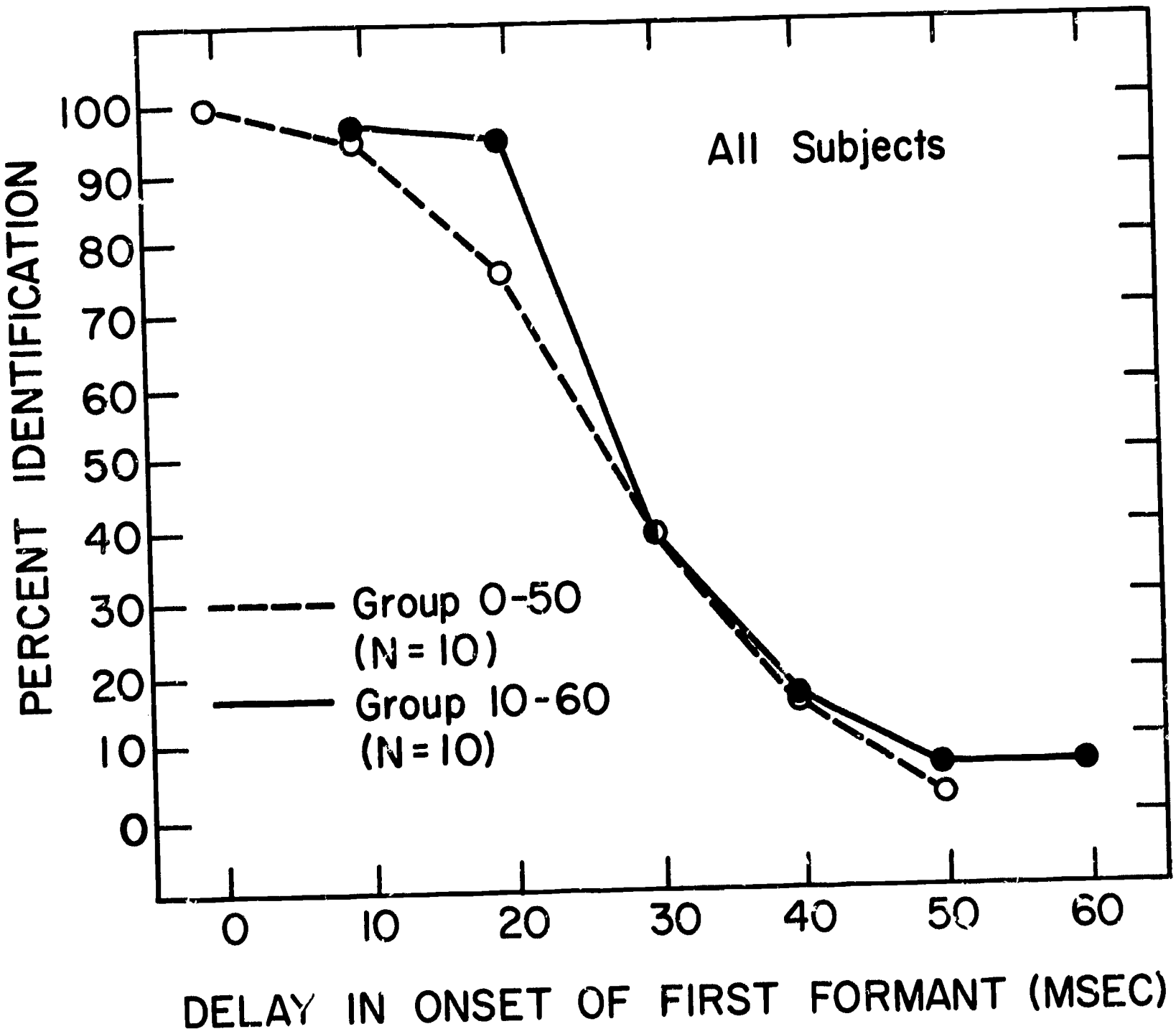
All Subjects
PERCENT IDENTIFICATION
100
90
80
70
60
50
40
30
20
10
0
Group 0-50
(N = 10)
Group 10-60
(N = 10)
0
10
20
30
40
50
60
DELAY IN ONSET OF FIRST FORMANT (MSEC)

Figure 2

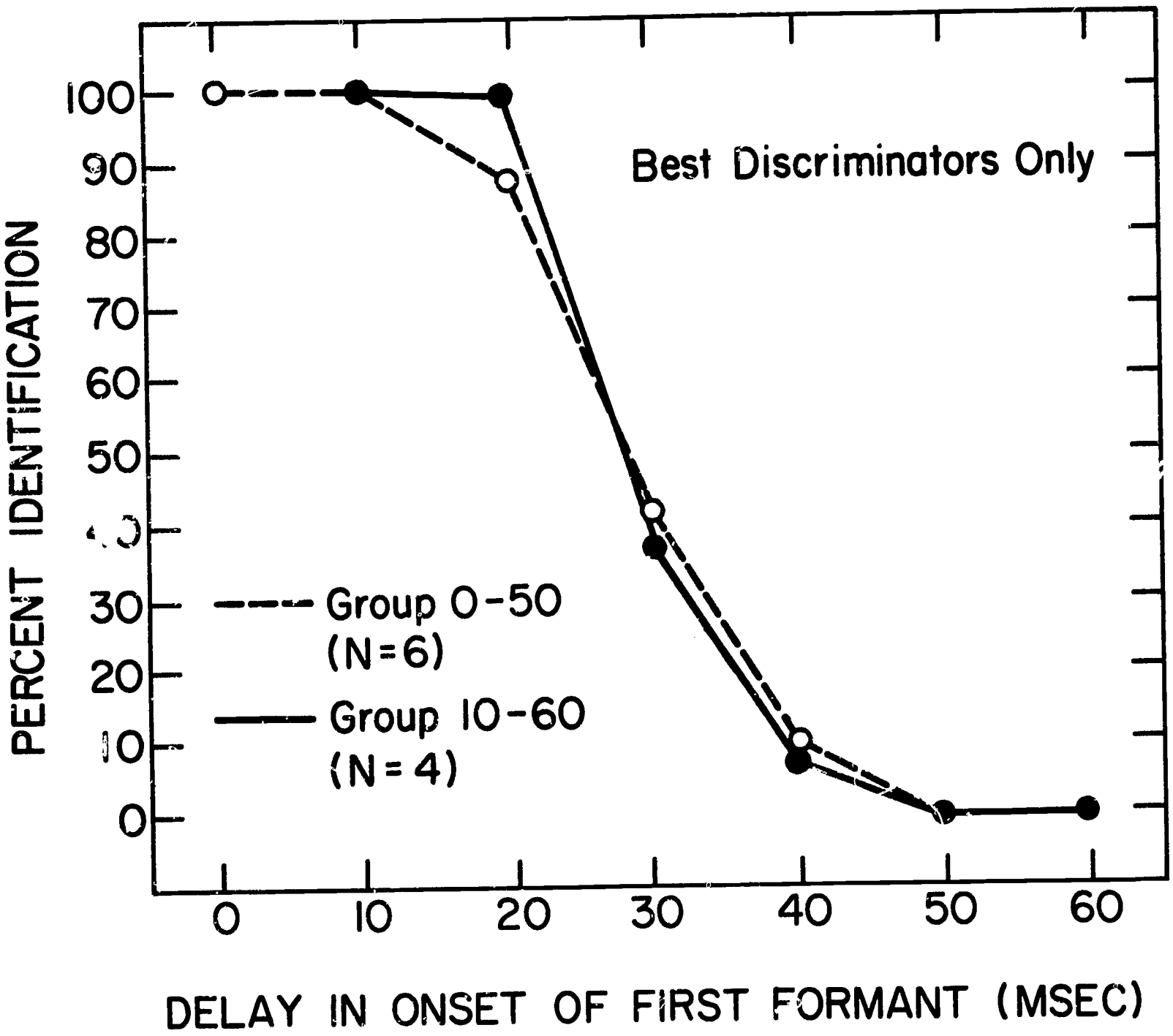
Best Discriminators Only
PERCENT IDENTIFICATION
100
90
80
70
60
50
30
20
10
0
Group 0-50
(N=6)
Group 10-60
(N=4)
0
10
20
30
40
50
60
DELAY IN ONSET OF FIRST FORMANT (MSEC)

Figure 3