Dichotic Backward Masking, The "Lag Effect" and Processing Phonetic Features*

S. D. McNabb and D. B. Pisoni
Department of Psychology
Indiana University
Bloomington, Indiana 47401

The "lag effect" has been used to support the argument that speech perception engages distinctive processes that differ from those of nonspeech auditory perception. However, a number of recent experiments have shown that the effect may not be peculiar to speech since it may be obtained with nonspeech timbres, vowels and other stimuli. In fact, the effect appears to be a variation of a more general result obtained in backward masking experiments: a second stimulus may impede the processing of a preceding stimulus. The present study sought to determine the locus of this effect. Under dichotic presentation, one of four syllable targets (/ba/, /da/, /pa/, /ta/) was followed by one of six possible syllable masks (/ga/, /ka/, /gae/, /kae/, /gɛ/, /kɛ/), each 300 msec in duration. The onset time of the mask relative to that of the target was varied over the range 0 to -150 msec. Subjects identified only the target sound in an ear monitoring task. Two findings were obtained which argue that the lag effect has an auditory basis. First, when the target and mask differed on voicing (i.e., b-k, p-g) performance improved with increases in the onset time of the mask; no interference was obtained when the target and mask shared voicing. Secondly, performance varied inversely with the similarity of the vowels in the mask: performance was lowest with /a/ and highest with /ɛ/. Since the interference occurred only for trials that differ on voicing and these trials were also affected by vowel context, we conclude that the locus of the lag effect lies before phonetic analysis and therefore must have an auditory rather than phonetic basis.

*This paper was presented at the 87th meeting of the Acoustical Society of America, April, 1974, New York City, New York. The research was supported in part by PHS grant MH 24027 to Indiana University.

Over the last few years we have heard a large number of papers at the Society which have dealt with the so-called "lag-effect" in dichotic listening experiments.

Insert Figure 1 about here

The effect is shown here in Figure 1 which we have borrowed from Studdert-Kennedy, Shankweiler & Schulman (1970). The second or "lagging" syllable of a dichotic pair of temporally overlapping stimuli is reported or identified more accurately than the "leading" syllable. I am sure you are all familiar with this data.

Our interest in these findings lies in a number of claims that have been made as to the locus of the effect: That is, where does the interaction between the two inputs occur? Also, is the lag effect peculiar to speech perception or does it result from more general perceptual operations? Studdert-Kennedy, Berlin and others have interpreted the lag effect as a form of "interruption of speech processing" occurring at some "central" level of perceptual analysis. For example, Studdert-Kennedy, Shankweiler & Schulman (1970) state that "the lag effect is tied to speech, and, specifically, to those components of the speech stream for which a relatively complex decoding operation is necessary." Indeed, the lag effect has been used to support the general argument that speech perception engages specialized processes that differ from those of non-speech perception--that is that "speech is special."

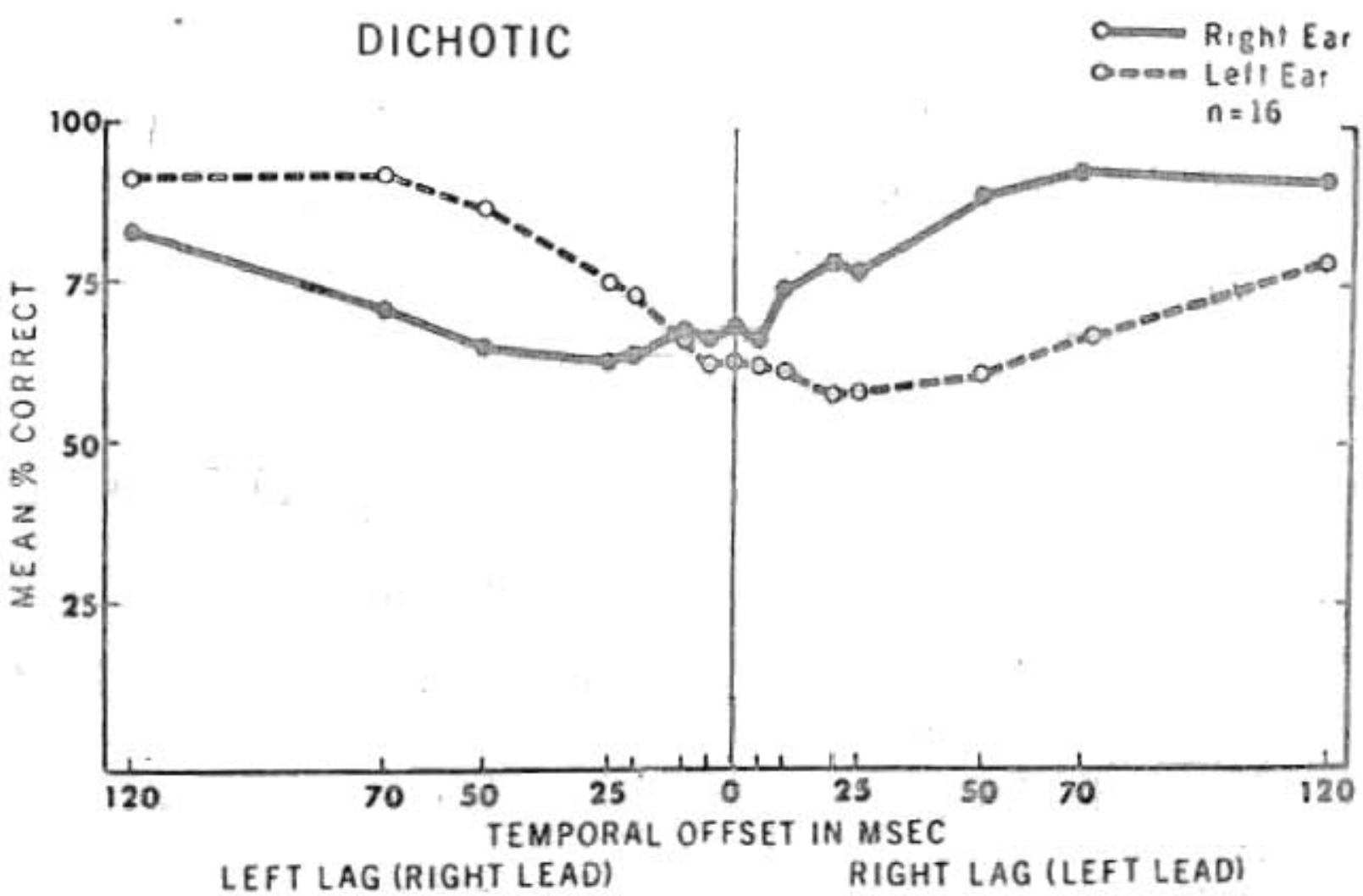
DICHOTIC
Right Ear
Left Ear
n = 16
100
75
50
25
MEAN % CORRECT
120
70
50
25
0
25
50
70
120
TEMPORAL OFFSET IN MSEC
LEFT LAG (RIGHT LEAD)
RIGHT LAG (LEFT LEAD)

Figure 1.

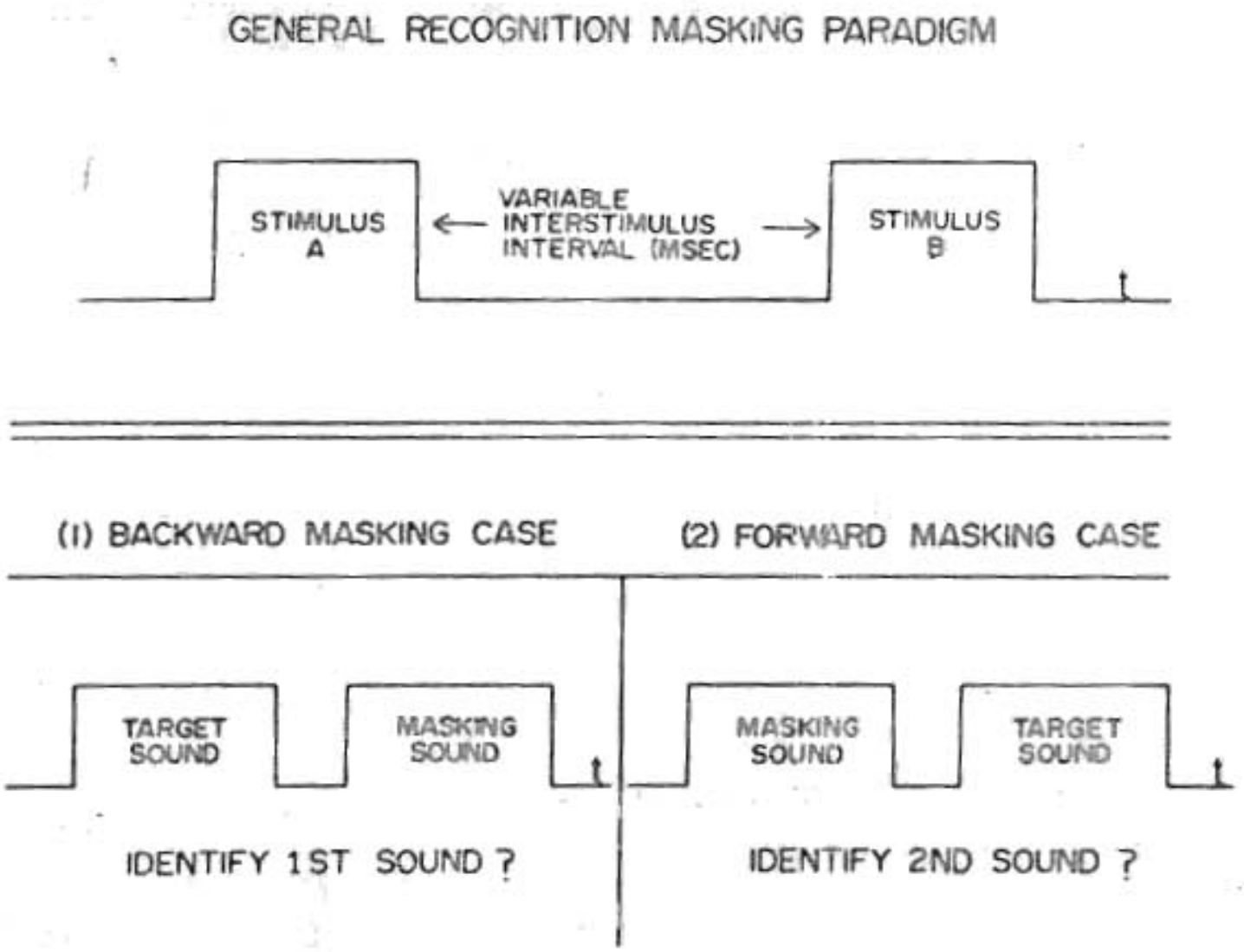
GENERAL RECOGNITION MASKING PARADIGM
STIMULUS A
VARIABLE INTERSTIMULUS INTERVAL (MSEC)
STIMULUS B
(1) BACKWARD MASKING CASE
(2) FORWARD MASKING CASE
TARGET SOUND
MASKING SOUND
MASKING SOUND
TARGET SOUND
IDENTIFY 1ST SOUND ?
IDENTIFY 2ND SOUND ?

Figure 2.

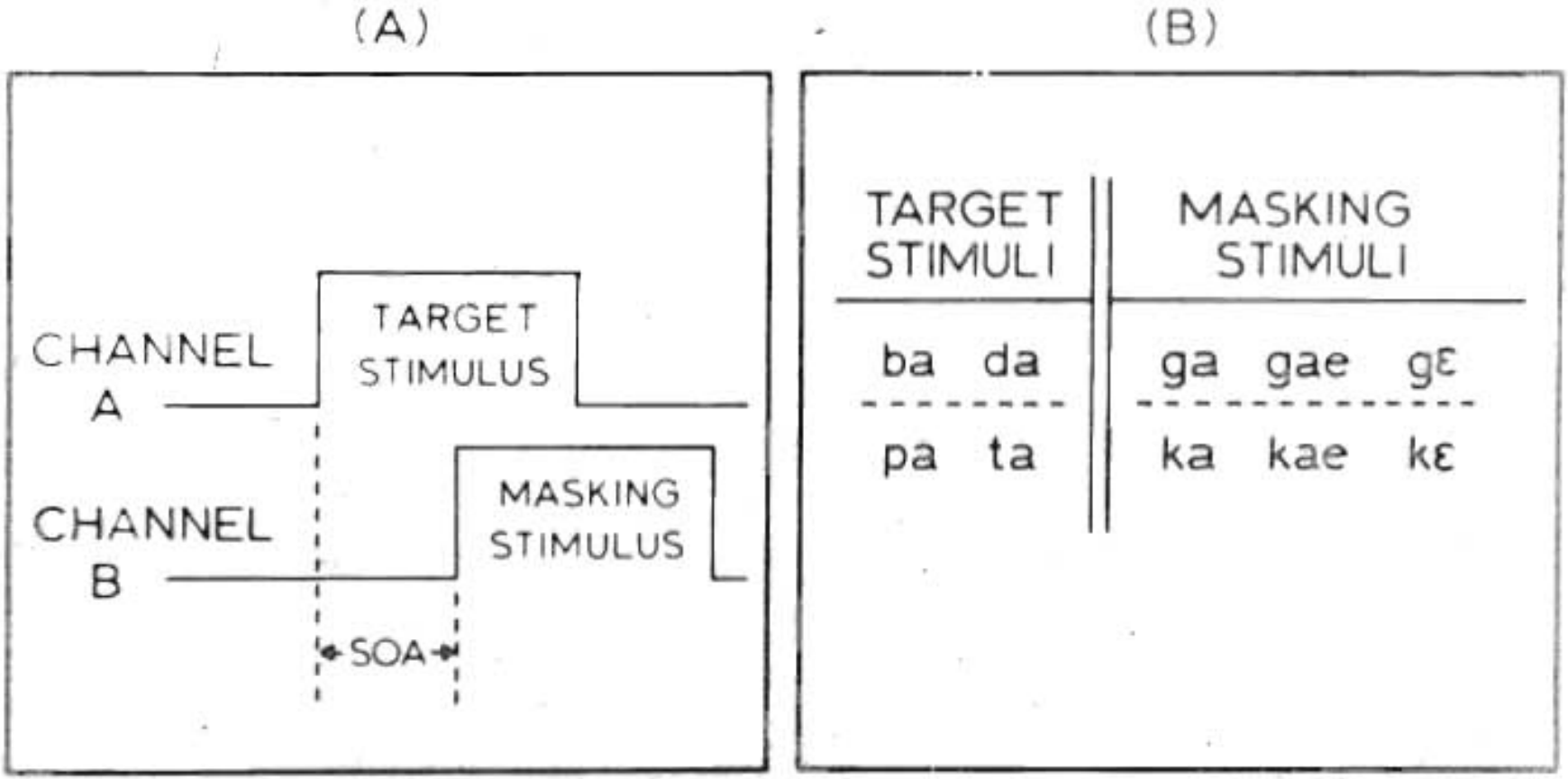
(A)
CHANNEL A
TARGET STIMULUS
CHANNEL B
MASKING STIMULUS
SOA
(B)
TARGET STIMULI
MASKING STIMULI
ba da
ga gae gɛ
pa ta
ka kae kɛ

Figure 3.

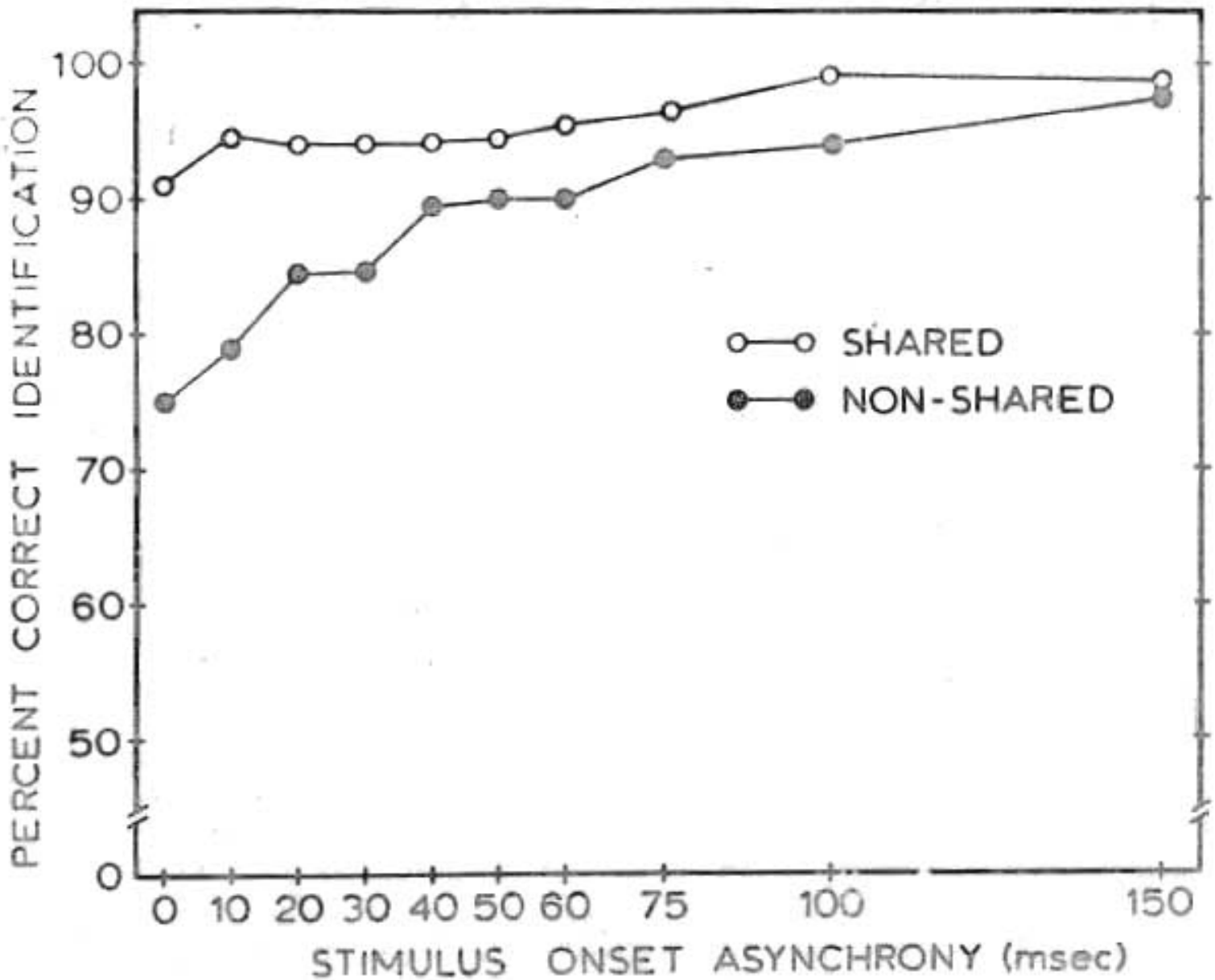
PERCENT CORRECT IDENTIFICATION
100
90
80
70
60
50
0
SHARED
NON-SHARED
0 10 20 30 40 50 60 75 100 150
STIMULUS ONSET ASYNCHRONY (msec)

Figure 4.

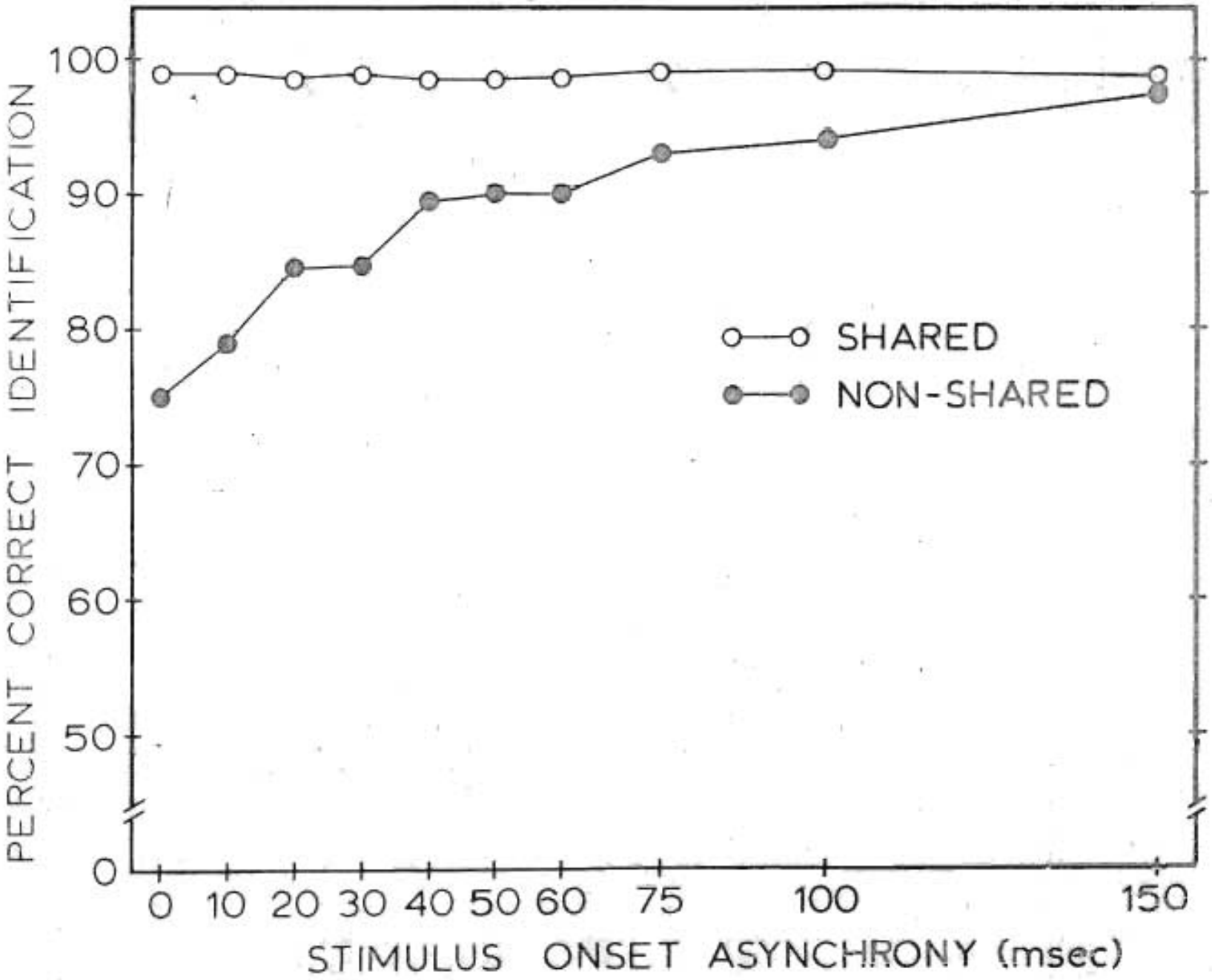
PERCENT CORRECT IDENTIFICATION
100
90
80
70
60
50
0
SHARED
NON-SHARED
0 10 20 30 40 50 60 75 100 150
STIMULUS ONSET ASYNCHRONY (msec)

Figure 5.

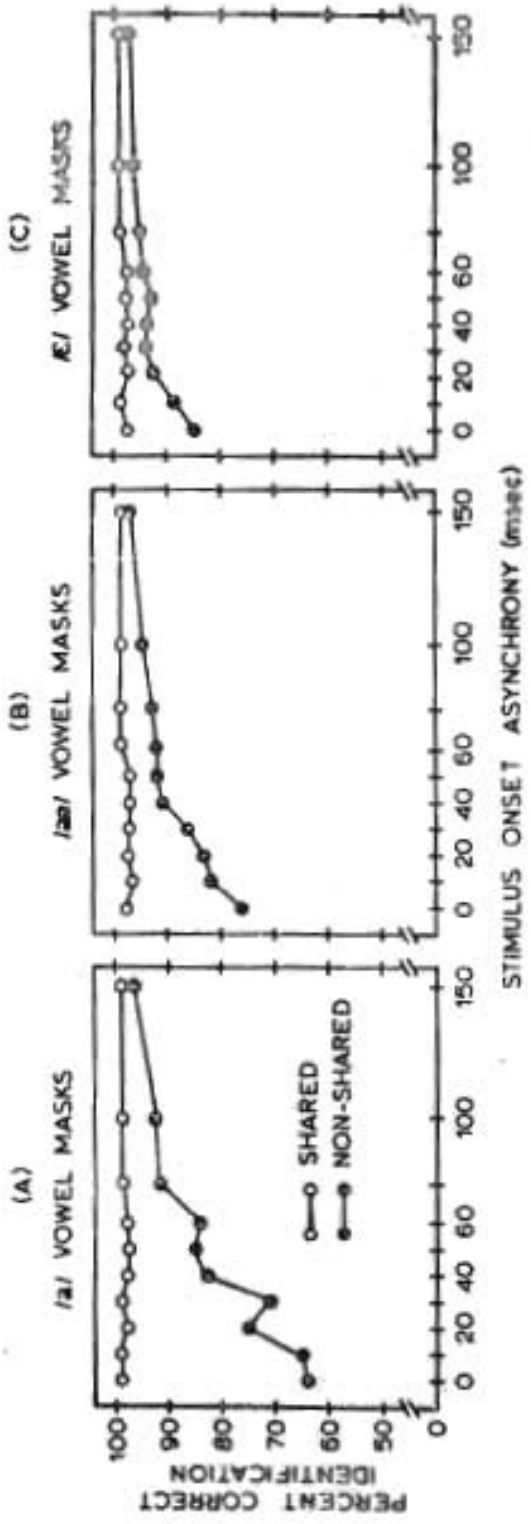
(A)
/a/ VOWEL MASKS
(B)
/æ/ VOWEL MASKS
(C)
/ɛ/ VOWEL MASKS
SHARED
NON-SHARED
PERCENT CORRECT IDENTIFICATION
100
90
80
70
60
50
0
0 20 40 60 100 150
STIMULUS ONSET ASYNCHRONY (msec)

Figure 6.

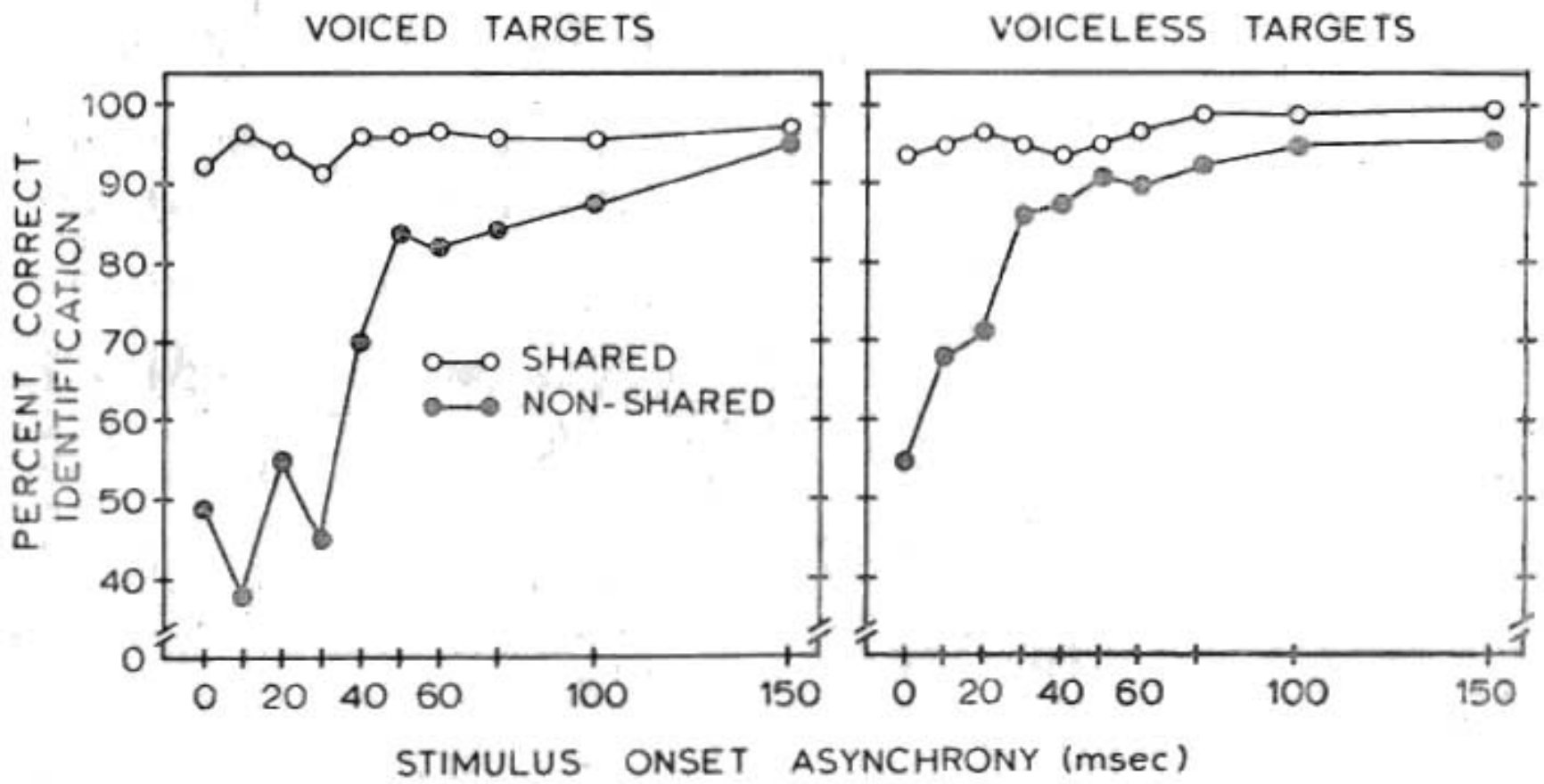
VOICED TARGETS
VOICELESS TARGETS
PERCENT CORRECT IDENTIFICATION
100
90
80
70
60
50
40
0
SHARED
NON-SHARED
0
20
40
60
100
150
STIMULUS ONSET ASYNCHRONY (msec)

Figure 7.

We do not believe that the lag effect is mysterious or peculiar and we hope to show you why. The effect seems to be a variation of a more general result obtained in backward recognition masking experiments: A second stimulus can impede the processing of a preceeding stimulus (Massaro, 1972).

Insert Figure 2 about here

Strictly speaking, the lag effect refers to the relative difference between dichotic forward and backward masking. Under the conditions usually studied there is more dichotic backward masking than forward masking (see also Repp, 1973).

In this study we were concerned with the "locus" of the interaction between the dichotic inputs. Where in the flow of information does the effect arise? Put another way, does the effect have an auditory or a phonetic basis? To study this problem we looked at the effect of three variables on the identification of a known set of "target" CV syllables.

Insert Figure 3 about here

These can be best shown in Figure 3. The first variable is stimulus onset asynchrony (SOA) as shown in panel (A). This was varied over a ten step range from 0 msec. to -150 msec. The second

variable was the feature composition of the target and mask pair. On half the trials the target stimulus and mask shared voicing; on the other half they differed on voicing. For example, /ba/ followed by /ga/ is a shared pair whereas /ba/ followed by /pa/ is a non-shared pair or double contrast trial. The last variable we consider is the vowel in the masking stimulus. As shown in panel (B) of this figure we have 3 vowel contexts for our masking stimuli: /a/, /ae/ and /ɛ/.

If the interference in the lag effect is due to "interruption" of processing, we should expect any mask to be equivalent with regard to its masking effectiveness. On the other hand, if part of the features in the first stimulus have been recognized then whether the target and mask share voicing should affect recognition performance.

If the interference is due to interactions on the phonetic feature level, we should not expect the vowel context of the mask to affect recognition of the target since the consonant would have already been abstracted from the syllable. On the other hand, if the interaction has an auditory basis at an earlier stage of analysis where the syllables interact we would anticipate a systematic vowel effect according to vowel similarity. That is, performance should be lowest with the /a/ vowel masks, highest with /ɛ/ and midway for /ae/.

Twenty-four Ss were run in a dichotic ear monitoring task where on each trial one of the 4 targets was followed by one of the 6 possible masks. These 24 pairs of stimuli were presented at 10 different SOA values in a random order. Ss identified only the targets /ba/, /da/,

/pa/, /ta/ and ignored the other stimuli.

Insert Figure 4 about here

Figure 4 shows the percent correct identification for shared vs. non-shared pairs averaged over all vowel contexts. Performance is relatively good for shared pairs and is not affected very much by SOA. The non-shared condition is lower and is affected by increases in SOA. Performance improves as SOA becomes larger.

Insert Figure 5 about here

This figure shows the same data but scored now for the voicing feature. Performance is perfect for voicing when target and mask share the voicing feature. Performance steadily improves as a function of SOA for non-shared pairs.

Insert Figure 6 about here

This slide shows the effect of the vowel context of the mask. Performance on voicing is lowest for non-shared pairs when the vowel in the mask and target are the same (i.e., /a/). Performance improves as vowel similarity decreases from /a/ to /ae/ to /ɛ/.

Thus, we think we have good grounds for arguing that the interference

obtained in the lag effect has an auditory rather than phonetic basis since the vowel in the mask interacts with recognition of the target stimulus.

Insert Figure 7 about here

This figure shows the trials broken down by voiced and voiceless targets for shared and non-shared trials. The feature effect shows up again. Shared pairs are recognized better than non-shared pairs.

To summarize, we have found that interference in a backward recognition masking experiment does not occur for all stimulus contrasts but only those differing on voicing (i.e., double contrasts). Hence, the interference in the lag effect cannot be due to "interruption" of processing as suggested in earlier reports. We also found that the vowel of the mask systematically effects recognition of the target. We interpret this to mean that the interaction between the inputs (that is, the locus of interaction) occurs before "phonetic" analysis.

We think these data argue strongly that the interference obtained in the lag effect has an auditory rather than phonetic basis and that there is nothing mysterious or peculiar about the lag effect results obtained in previous dichotic listening experiments.