Perceptual Competition Between Speech and Nonspeech

Ruth S. Day† and James E. Cutting†
Haskins Laboratories, New Haven

Ear advantages are reported throughout the dichotic listening literature. When different messages are presented to each ear at the same time, the information presented to one ear is identified more accurately than the information presented to the other ear. In order to get an overview of the ear advantage results, consider the shorthand summary shown in Figure 1. When both inputs are speech (S/S), there is a right-ear advantage: stimuli presented to the right ear are identified more accurately than those presented to the left ear. This result has been obtained for a wide variety of speech stimuli, including digits (Kimura, 1961), words (Borokowski, Spreen, and Stutz, 1965), and consonant-vowel syllables (Shankweiler and Studdert-Kennedy, 1967). When both inputs are nonspeech (NS/NS), there is a left-ear advantage: stimuli presented to the left ear are identified more accurately. This result has been shown for various nonspeech stimuli, including melodies (Kimura, 1964), sonar signals (Chaney and Webster, 1966), and environmental noises (Curry, 1967).

Given that S/S yields a right-ear advantage and NS/NS yields a left-ear advantage, what happens when we present speech to one ear and nonspeech to the other ear at the same time? Will there be any ear advantage? One plausible prediction is that performance will be best when each stimulus is presented to its "proper" ear, that is, when nonspeech goes to the left ear and speech goes to the right ear (NS_L/S_R). This is a reasonable prediction if indeed the right-ear/left-hemisphere system processes speech and the left-ear/right-hemisphere system processes nonspeech, as suggested by Kimura (1967). The present study was designed to study the perception of speech and nonspeech pairs in the dichotic listening situation (S/NS).

**EXPERIMENT I - IDENTIFICATION TASK**

**Method**

Stimuli. The speech stimuli were the syllables /ba, da, ga/. They were real speech samples produced by one of the authors (RSD). The nonspeech stimuli were sine wave tones of 500, 700, and 1000 Hz. All stimuli were 300

---

*Paper presented at the eightieth meeting of the Acoustical Society of America, Houston, 3-6 November 1970.

+Also, Department of Psychology, Yale University, New Haven.
Summary of Ear Advantage Results from Previous Dichotic Listening Studies That Used Two Speech Stimuli (S/S) or Two Nonspeech Stimuli (NS/NS); Situation Studied in the Present Study (S/NS)

Fig. 1
msec in duration. The intensity envelopes of the tone stimuli were shaped to match those of the speech stimuli.

**Paradigm.** On each trial, one of the speech stimuli was paired with one of the tone stimuli. The assignment of stimulus type to ear varied randomly over trials, and all of the appropriate counterbalancing measures were taken into account. The onsets of the dichotic pairs were aligned using the pulse code modulation system at the Haskins Laboratories (Mattingly and Cooper, 1969). On some trials, both the speech and tone began at the same time. A sample pair with this 0-lead time is shown in Figure 2. On the remaining trials, one of the stimuli began first, by 25, 50, 75, 100, 150, or 200 msec. This paradigm enabled us to assess the role of temporal cues in perceiving the dichotic S/NS stimuli.

**Subjects.** Ten Yale University students served as subjects. All were right handed, were native English speakers, and had no history of hearing trouble. Each was tested individually.

**Pretest.** Before beginning the main experiment, each subject successfully completed a binaural identification test for the six stimuli.

**Procedure.** The subject was told that a speech stimulus would be presented to one ear and a tone stimulus to the other ear on each trial. His task was simply to write down which two stimuli he heard: (/ba/, /da/, or /ga/) and (high, medium, or low tone).

**Results**

The overall level of performance was excellent: the mean percent correct over all subjects and trials was 97 percent. The lowest identification score for any subject was 95 percent. This performance level is higher than those reported in the previous dichotic listening literature. Since there were virtually no errors, there was no opportunity for an ear advantage to occur: neither ear yielded superior performance, nor did it make any difference whether the stimuli were presented to the "proper" or "improper" ears.

In the present S/NS situation, then, we have two findings that conflict with the previous literature: essentially perfect identification accuracy and no ear effect.

**Discussion**

In light of these findings, perhaps we must reconsider what we mean by dichotic competition. Perceptual competition does not occur simply by presenting different messages to each ear at the same time, since S/NS pairs yielded virtually error-free performance. The S/NS condition does not appear to create a situation of information overload. The data fit the notion that there are two processors for auditory stimuli: a speech processor and a non-speech processor. An ear advantage occurs only when a given processor is
Oscilloscope Photograph of a Sample Dichotic Trial Where Speech and Tone Began at the Same Time
overloaded, as when two speech stimuli are sent to the speech processor or when two nonspeech stimuli are sent to the nonspeech processor. In the present S/NS situation, the stimuli appear to be sent to different processors, each of which can perform its job without interference from the other.

But this is not the whole story. There was another task.

EXPERIMENT II - TEMPORAL ORDER JUDGMENT (TOJ) TASK

Method

Fifteen subjects (the ten from Exp. I plus five from the same pool) listened to the same stimulus tape. This time, each subject was asked to determine which stimulus began first on every trial, be it /ba/, /da/, /ga/, high, medium, or low tone. Therefore we are interested in percent correct temporal order judgment (TOJ) as a function of the lead-time conditions. Will performance again be equally accurate for both ears and will overall performance stay virtually error-free?

Results

First, let us examine the results for a single subject (MS) as shown in Figure 3. Here we have percent correct TOJ plotted as a function of ear and magnitude of lead. When the stimulus led in the left ear, performance was very good. This subject gave correct temporal order judgments over the entire range of left-ear leads. However, when the stimulus led in the right ear, performance was poor. The subject continued to report that the left-ear stimulus led. This effect was most striking at the short leads. Compare performance for the two ears at the 25-msec lead condition. MS was 94 percent correct for stimuli leading in the left ear and only 22 percent correct for stimuli leading in the right ear. Therefore, there was a net 72 percent advantage for the left ear over the right at 25 msec. This is a very large effect and contrasts with the typical ear advantage of approximately 10 percent reported in the previous literature (obtained for S/S identification trials). Similarly, at 50 msec, there was a net 55 percent left-ear advantage. Even at 75 msec, the left ear still had a 19 percent advantage. However, from 100 msec on outward, both ears performed equally well. Note that there is no "correct" response at the 0-lead point. Here we have simply entered the proportion of left- and right-ear responses, and attached the open circles to the appropriate sides of the continuum.

At the long leads MS performed very well; only at the short leads did he show an ear asymmetry. At what point did he judge the left-ear stimulus to lead 50 percent of the time and the right-ear stimulus 50 percent of the time—what we might call the "point of subjective equality"? The inputs for each dichotic trial have been aligned very precisely according to computer-controlled methods. So we know that the O-lead items began at indeed the same point in time, down to 500 microsec accuracy. But it appears that MS does not know where the O-lead cases are in the array of physical stimuli.
Sample Subject: Temporal Order Judgment (TOJ) Performance as a Function of Ear and Magnitude of Lead

Subject M.S.

Fig. 3
In fact, the point of subjective equality appears to be out as far as about 60 msec for this subject. That is, the stimulus in the right ear must lead by at least 60 msec in order for him to respond at chance level.

We have seen this general pattern of results many times now, for most of the fifteen subjects in the present experiment and for about fifty others in variants of the task. Subjects differed in the magnitude of the left-ear advantage and in the place on the continuum where the right ear recovered. No subjects showed a significant right-ear advantage.

The results for all fifteen subjects are shown in Figure 4. The data are shown in histogram form to facilitate comparison of performance for the two ears. There is a left-ear advantage at each point on the continuum, with effects statistically significant out through the 100 msec lead condition. These pooled data represent 540 observations per ear for each lead case and 270 observations per ear for the 0-lead case. This display does no great disservice to any individual subject. In contrast with the identification task, then, the TOJ task yielded a large left-ear advantage and a decrease in overall level of performance, especially at the shorter leads.

The left-ear advantage for the S/NS TOJ task is a very robust phenomenon. It appears to be immune to the effects of practice: performance remained stable over the course of the experiment, and subjects tested on several subsequent occasions gave comparable results in each session. None of the subjects was aware of the left-ear advantage when questioned in the postsession interview. When prompted to guess which ear yielded better performance, most guessed that they performed best on right-ear leads because they are right handed.

The effect carries with it some powerful phenomenological consequences. A brief anecdote concerning its discovery will illustrate this point. We were testing some new equipment designed to measure reaction times in another experiment. Briefly, the speech stimulus presented to the subject's earphones was also recorded on one channel of a response tape. The subject made one of two button-press responses on each trial. Each button activated an oscillator, and the resulting high or low tone was recorded on the second channel of the response tape. After a series of trials, the response tape was ready to be run through a data processor that would measure the time from onset of a given stimulus (speech) on one channel to the onset of the response (tone) on the other channel. One of the authors served as "subject" while the other monitored the response tape over dichotic earphones. On the first trial, the "monitor" heard a tone, then speech. The "subject" was asked to wait until he heard the onset of the speech stimulus before pressing the button. He replied that he had. Another trial was run. Again, the "monitor" heard tone-speech and warned the "subject" not to give false start. Again, the "subject" replied that he had heard the speech stimulus before he pressed the button. After a rather frustrating interchange, the "monitor" reversed the earphones and another trial was run. The effect went away. Under the original earphone configuration, the tone had gone to the left ear.

So far we have been discussing ear effects. What about stimulus effects? Subjects did show a bias or preference for reporting one class of stimuli over the other. For most, this bias was in favor of the speech stimulus. The stimulus effect may be related to labelling processes. The name for a speech
Group Data: Temporal Order Judgment (TOJ) Performance as a Function of Ear and Magnitude of Lead

![Bar Chart]

- **LEFT EAR LED**
- **RIGHT EAR LED**

% CORRECT TOJ

LEAD TIME (MSEC.)

0 25 50 75 100 150 200
stimulus is, in a sense, already in some sort of echoic memory before the subject makes his response, whereas he must search for the name to give to a tone stimulus, e.g., "medium" or "low." In any event, it is important to emphasize that subjects showed a wide range of stimulus preferences. On the other hand, the ear effect was unidirectional and was maintained no matter what type of preference an individual had for the stimuli.

Discussion

How are we to explain the left-ear advantage on the TOJ task? Unfortunately the present experiment does not provide enough data for a clear answer. Nevertheless, there are several approaches one might take to begin thinking about this robust and puzzling phenomenon.

Left-hemisphere function. Let us assume that the outlines of current dichotic listening models are reasonable (e.g., Kimura, 1967). As shown in Figure 5, we will assume 1) that the contralateral (crossed) connections from ears to hemispheres are prepotent in the dichotic listening situation, and 2) that the left hemisphere has a highly specialized processor, one that is designed to handle speech stimuli. If indeed the left hemisphere has some extra, very complicated machinery, then perhaps this machinery takes longer to "warm up" once stimuli are fed into it. Put crudely, the "turn around time" for the special processor might be on the order of 50-75 msec. However, this model needs amplification in order to deal with the entire dichotic listening literature. In the temporary absence of some appropriate follow-up data, discussion of the revised model would be merely speculative.

Right-hemisphere function. Many investigators have suggested that space is processed primarily in the right hemisphere (e.g., Carmon and Bechtoldt, 1969; Dorff et al., 1965; Kimura, 1963; Milner, 1958). The present S/NS data suggest that perhaps time is also processed more efficiently in the right hemisphere. Such a view would go counter to that of Efron (1965) who has suggested that time is handled back in the left hemisphere. But Efron was dealing with very simple situations such as light flashes and finger shock. Such situations may well involve different time perception mechanisms than those involved in speech. The view that time perception is handled primarily in a single hemisphere may well be an over-simplification. Day and Cutting (1971) recently found that TOJ accuracy yielded different ear advantages depending on the nature of the stimuli: S/S yielded a right-ear advantage while both NS/NS and S/NS yielded a left-ear advantage.

Modes of neural organization. So far we have been looking for specific functions that each hemisphere might perform: speech in the left hemisphere and time in the right hemisphere. There is an alternative approach with a different emphasis. Semmes (1968) has suggested that the two hemispheres represent contrasting modes of neural organization. The left hemisphere involves focal representation of functions, such as that for speech, whereas the right hemisphere involves diffuse representation of functions. To quote Semmes, "diffuse representation of elementary functions in the right hemisphere may lead to integration of dissimilar units." Therefore perhaps it is the comparison between dissimilar units in the S/NS TOJ task that produces the left-ear/right-hemisphere effect that we see so clearly in the data.
Schematic Diagram of the Connections from Ears to Hemispheres

LEFT

RIGHT

Fig. 5
More work is needed in order to derive a satisfactory explanation for the S/NS data. No doubt various aspects of the above approaches will be helpful in solving this puzzle.

**SUMMARY**

On the identification task, there was no ear advantage, since performance was essentially perfect. On the TOJ task, there was a large left-ear advantage. Performance levels were near chance for O-lead cases, and improved gradually to virtually perfect performance at the long leads.

There appear to be two necessary conditions for the large left-ear advantage to occur: 1) both speech and nonspeech must be put into the system, and 2) the subject must make a judgment that depends on relative time perception.

We have an interesting puzzle on our hands. We have a large and robust phenomenon that is not predicted by existing models of dichotic listening. Perhaps, then, the existing models only appear to explain the results now in the literature, and what is needed is a revised model that will handle both the old and the new phenomena.

**REFERENCES**


