A Parallel Between Degree of Encodedness and the Ear Advantage: Evidence from a Temporal Order Judgment Task*

Ruth S. Day[+] and James M. Vigorito[+]
Haskins Laboratories, New Haven

Speech sounds are not speech sounds are not speech sounds. That is, some speech sounds appear to be more highly "encoded" than others (Liberman, Cooper, Shankweiler, and Studdert-Kennedy, 1967). Let us illustrate this notion of encodedness in a very simple way.

Suppose we have a syllable consisting of a stop consonant followed by a vowel. We now break this syllable into two portions. The first segment contains all the information preceding the steady-state portion of the vowel. When we play it in isolation, several times in succession, listeners usually identify it as a coffee pot gurgle, a Model T sputter, or some other nonspeech sound. However, when we play the second portion of the split syllable, listeners have no difficulty in identifying it as the vowel /i/.

The point of this tape splicing demonstration is illustrated in Figure 1. The basic sound units of speech, the phonemes, are not added up like so many beads on a string, as shown on the left side of the display. Instead, there is an overlapping of linguistic segments as shown on the right side of the display. As these segments overlap, they undergo restructuring at the acoustic level.

Some speech sounds undergo more restructuring than others, as shown by the "split /p/" demonstration. The /p/ has a particular acoustic structure, namely one that is appropriate to be in initial position and followed by the vowel /i/. Therefore it cannot be recovered perceptually after the tape splicing procedure. Meanwhile the /i/ has undergone relatively little change as a function of context. Hence it can easily be recovered perceptually despite the fact that it has been spliced out of context.

Those speech sounds that undergo the most restructuring in the sound stream are said to be highly encoded, whereas those that undergo relatively little change as a function of neighboring phonemes are said to be less encoded. In general, stop consonants such as /p/ are highly encoded whereas vowels are relatively unencoded.

---

*Paper presented at the 84th meeting of the Acoustical Society of America, Miami Beach, Fla., December 1972.

[+]Also Department of Psychology, Yale University, New Haven.

[HASKINS LABORATORIES: Status Report on Speech Research SR-31/32 (1972)]

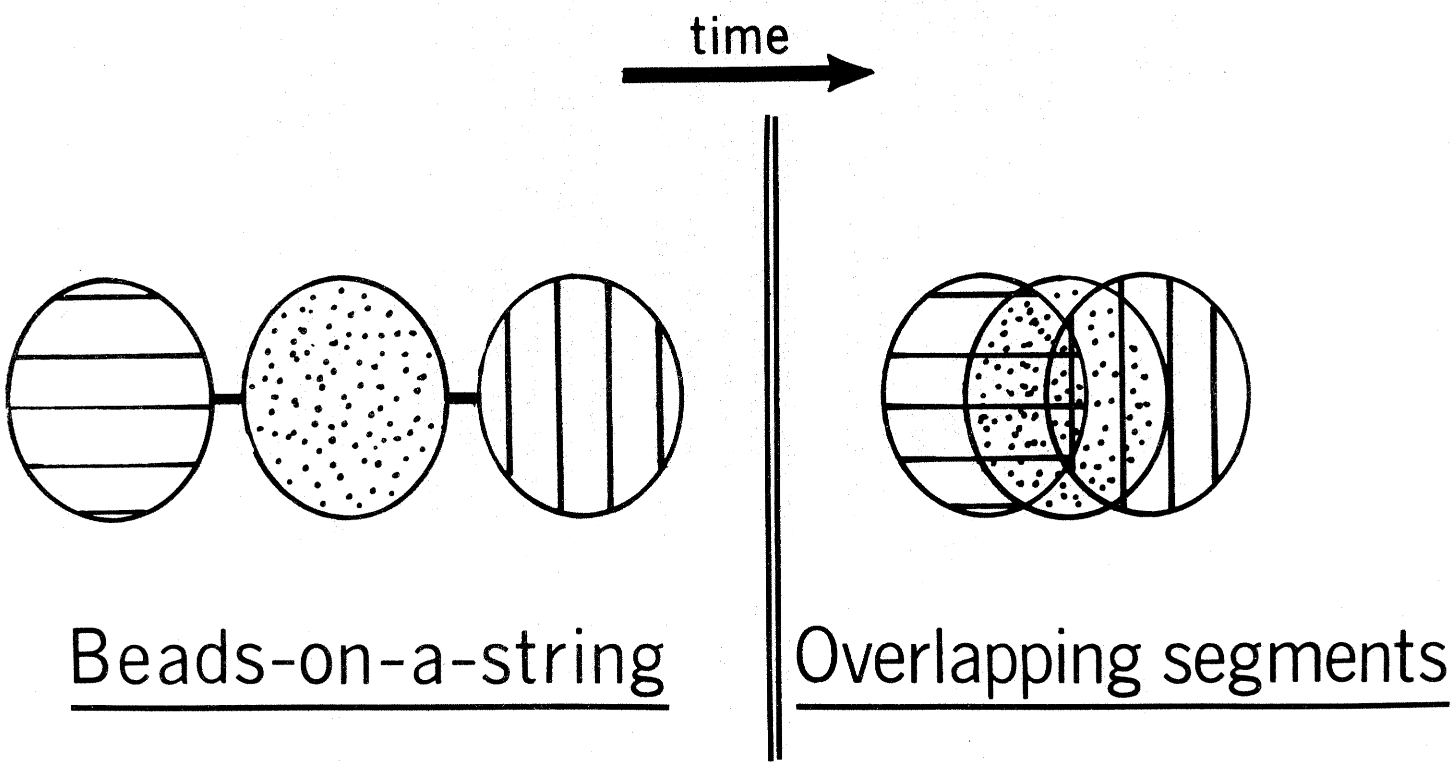

Fig. 1

Figure 1: Two views of the arrangement of phonetic units in time.

A summary of tape splicing results is shown in Figure 2. We have already considered stop-vowel syllables such as /pi/. What happens when we apply the same tape splicing procedure to other classes of speech sounds? Consider a fricative-vowel syllable such as /si/ as shown on the right side of the display. When listeners are asked to identify the first segment, they report hearing /s/ most of the time. This finding suggests that /s/ is not as highly encoded as the stop consonants. Nevertheless, it is not recovered perceptually as often as the vowel; therefore fricatives appear to be more highly encoded than vowels.

It may well be that there is an encodedness continuum for classes of speech sounds, with stop consonants at one end, vowels at the other end, and remaining classes such as fricatives and liquids falling in the middle.

At this point a word of caution is in order. There are several different ways to locate classes of speech sounds on such an encodedness continuum. Tape splicing experiments have been discussed here since they can be presented quickly and clearly. Other types of supporting evidence can also be presented, for example categorical perception for consonants versus vowels (Studdert-Kennedy, Liberman, Harris, and Cooper, 1970).

What happens when we pit phonemes within a given class against each other in a dichotic listening task? A right-ear advantage in dichotic listening is thought to reflect the participation of a special speech processing mechanism. Do the highly encoded speech sounds engage this mechanism to a greater extent than the less encoded sounds? Indeed, can we predict the magnitude of the ear advantage by placing classes of speech sounds along an encodedness continuum determined on independent grounds? Other investigators have shown that dichotic stimuli differing only in the initial stop consonant yield highly reliable right-ear advantages. Vowel contrast pairs, however, yield inconsistent results, with some studies obtaining a small ear advantage and others no ear advantage at all. Such results suggest that vowels can be perceived as speech or as nonspeech. For a recent review of this literature, see Studdert-Kennedy and Shankweiler (1970).

The present experiment compared the ear advantages of stops, liquids, and vowels. Note that we are interested in a <u>rank ordering</u> of these stimulus classes in terms of the ear advantage: stops should have the largest right-ear advantage, liquids less of a right-ear advantage, and vowels the least right-ear advantage.

METHOD

On each dichotic trial one of the items began 50 msec before the other. The subject's task was to determine which syllable began first. Thus he had to make a temporal order judgment (TOJ). There were three tests which differed only in their vocabulary: the stop test used /bæ, dæ, gæ/, the liquid test /ræ, læ, wæ/, and the vowel test /i, a, u/. All stimuli were prepared on the parallel resonance synthesizer at the Haskins Laboratories, then arranged into dichotic tapes using the pulse code modulation system. The syllables were highly identifiable, as determined by a binaural pretest.

The 16 subjects received all three dichotic tests. The listeners were right-handed, native English speakers, and had no history of hearing trouble.

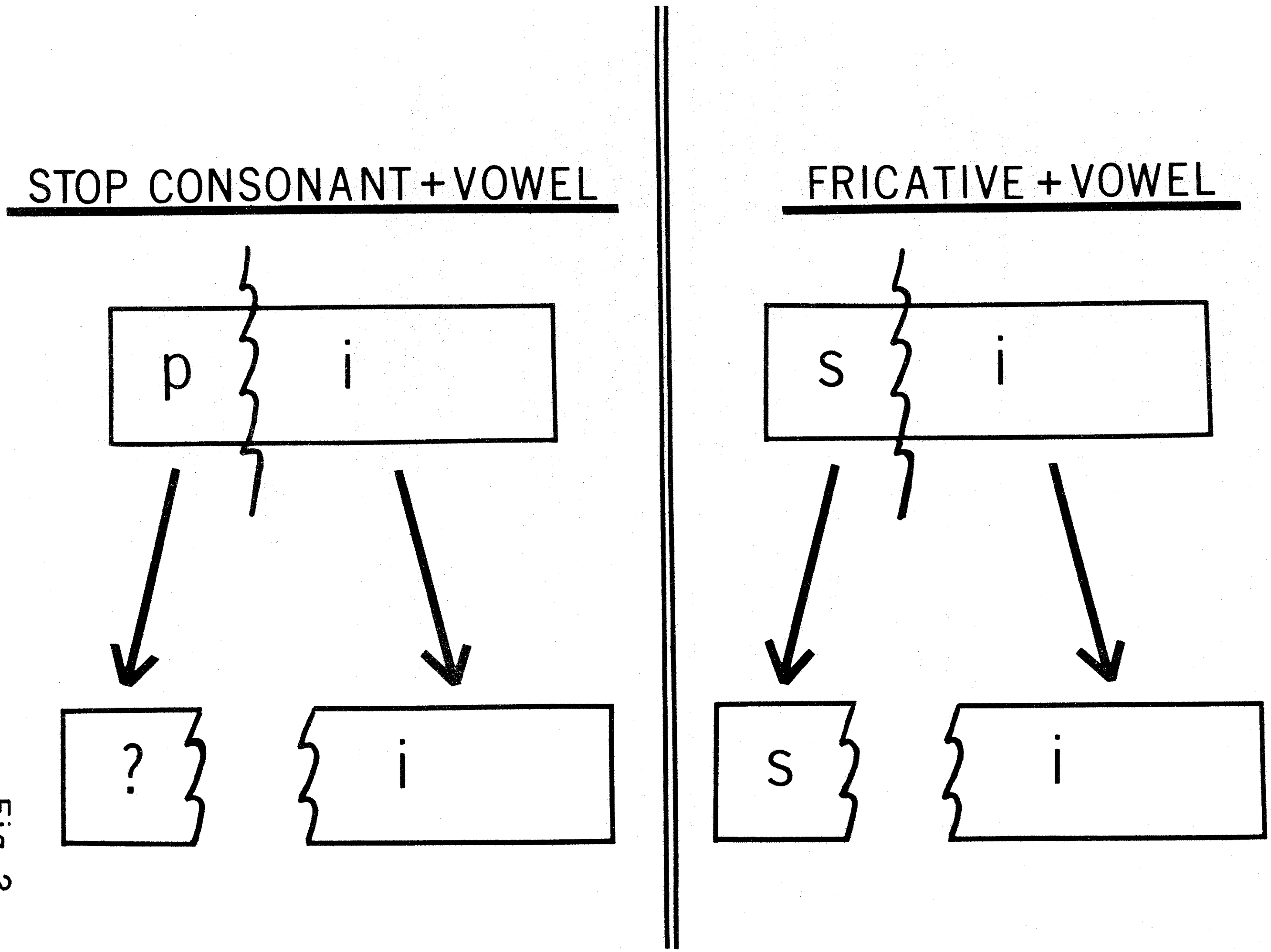

Fig. 2

Figure 2: Summary of tape splicing experiments.

Responses were scored in the following way. Given that a particular ear received the leading stimulus, what percent of the time was the subject correct in determining that that item did indeed lead? Ear difference scores were then obtained by subtracting percent correct TOJ for one ear from percent correct TOJ for the other ear.

RESULTS

The results are shown in Figure 3. The horizontal line indicates no ear advantage; here performance for the two ears is comparable. The region above this line indicates a right-ear advantage; that is, the right-ear score surpassed the left-ear score by the percent shown on the ordinate. The region below the horizontal line indicates a left-ear advantage in a comparable fashion.

Stops yielded a right-ear advantage, liquids a slight right-ear advantage, and vowels a left-ear advantage. Thus, as we moved along the encodedness continuum from stops to liquids to vowels, the right-ear advantage became reduced and finally disappeared. This is exactly the rank ordering predicted.

DISCUSSION

There appears to be a parallel between encodedness and the ear advantage. If indeed the right-ear advantage in dichotic listening reflects the operation of a special processing mechanism, then the present data suggest that the highly encoded speech sounds require the services of this mechanism to a greater extent than do the less encoded sounds.

The present data are compatible with those of Cutting (1972) who used consonant-consonant-vowel (CCV) syllables. The first consonant was always a stop consonant (/g/ or /k/), the second was a liquid (/l/ or /r/), and the following vowel was either /æ/ or /ɛ/, yielding eight syllables in all. A different syllable was presented to each ear at the same time and, for a given block of trials, the subject was asked to monitor one ear and report only the syllable presented to it. Ear advantage scores were obtained by subtracting percent correct for one ear from the percent correct for the other ear; this analysis was performed separately for each phoneme class. Again the rank ordering of phonemes in terms of the right-ear advantage was stops > liquids > vowels. However, the liquids yielded a sizable right-ear advantage while the vowels yielded no ear advantage in Cutting's experiment. Thus the data in the present TOJ experiment have in a sense been shifted "leftward" by comparison. There are several possible reasons for this shift. Cutting used CCV syllables; hence the speech processor may well have been engaged early in stimulus presentation, such that the liquids were clearly perceived in a "speech mode" while vowels were so perceived roughly half the time. Subjects in the present experiment heard liquid-vowel syllables and vowels in isolation; hence there were no stop consonants to engage the speech mechanism before the target contrasts were presented. Another possibly relevant factor is the type of dichotic tasks used. Cutting's ear-monitoring task required only identification of a syllable and not a judgment about relative timing. It could be that judgments about relative onset time are best handled in the nonspeech hemisphere, which would tend to depress the overall level of right-ear advantage scores. Finally, the different ear advantage levels obtained in the two experiments may have occurred simply because different subjects were used. Individuals differ in the extent

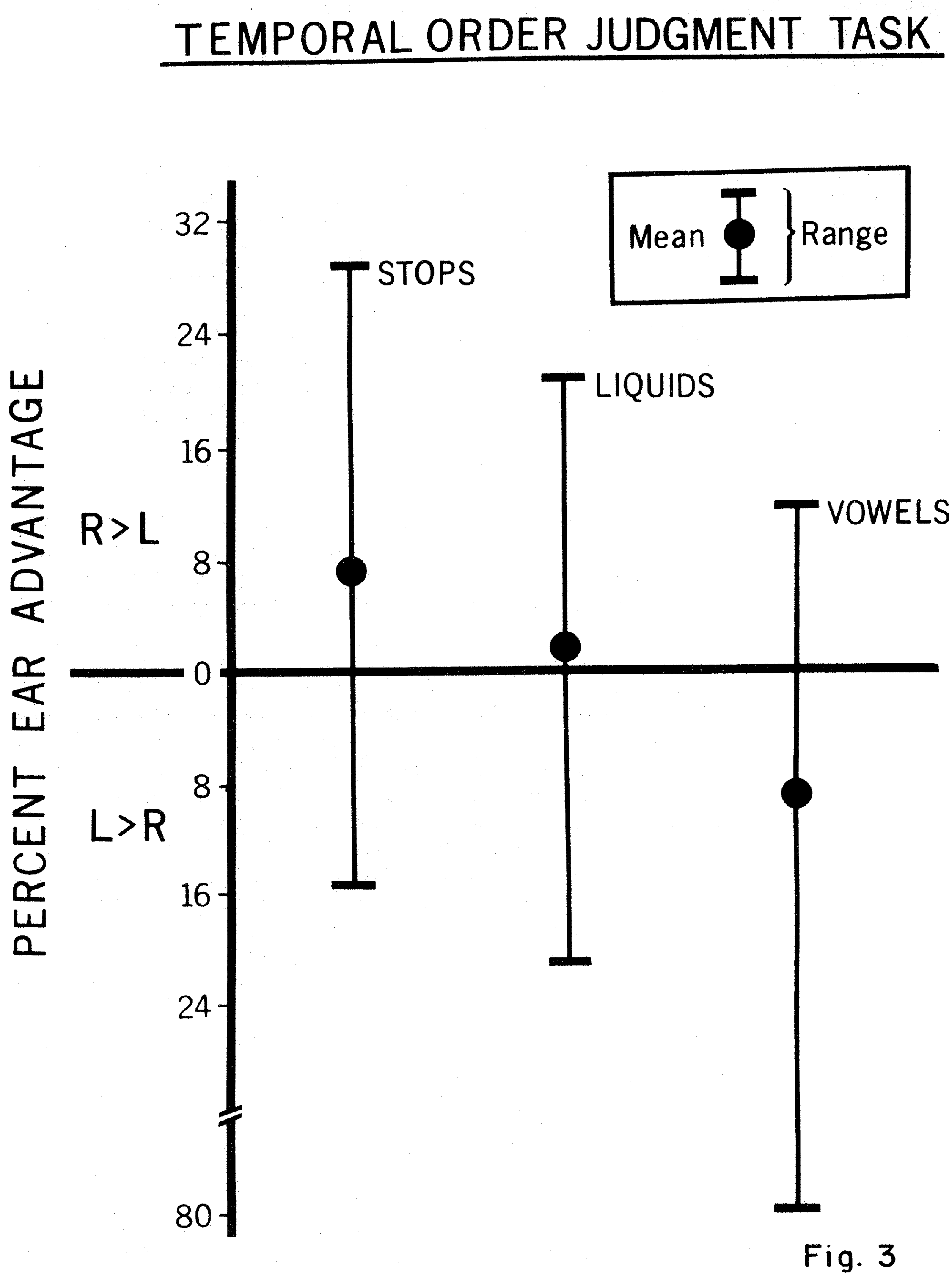

Figure 3: Ear advantages for stops, liquids, and vowels.

to which they show a right-ear advantage for dichotic speech items. Nevertheless, both experiments lend support to the notion that there is a parallel between the encodedness of speech sounds and the ear advantage.

REFERENCES

Cutting, J. E. (1972) A parallel between degree of encodedness and the ear advantage: Evidence from an ear-monitoring task. Paper presented at the 84th meeting of the Acoustical Society of America, Miami Beach, Fla. (Also in Haskins Laboratories Status Report on Speech Research SR-29/30, 61-68, as: A parallel between encodedness and the magnitude of the right ear effect.)

Liberman, A. M., F. S. Cooper, D. P. Shankweiler, and M. Studdert-Kennedy. (1967) Perception of the speech code. Psychol. Rev. 74, 431-461.

Studdert-Kennedy, M., A. M. Liberman, K. S. Harris, and F. S. Cooper. (1970) Motor theory of speech perception: A reply to Lane's critical review. Psychol. Rev. 77, 234-249.

Studdert-Kennedy, M. and D. P. Shankweiler. (1970) Hemispheric specialization for speech perception. J. Acoust. Soc. Amer. 48, 579-594.