DOCUMENT RESUME

ED 061 837 FL 003 093

AUTHOR	Turney, Michael T.; And Others
TITLE	Speech Research: A Report on the Status and Progress of Studies on the Nature of Speech , Instrumentation for Its Investigation, and Practical Applications, 1 October-31 December 1971.
INSTITUTION	Haskins Labs., New Haven, Conn.
SPONS AGENCY	National Institutes of Health (DHEW), Bethesda, Md.; National Inst. of Child Health and Human Development (NIH), Bethesda, Md.; Office of Naval Research, Washington, D.C. Information Systems Research.
REPORT NO	SR-28-71
PUB DATE	71
NOTE	256p.
EDRS PRICE	MF-$0.65 HC-$9.87
DESCRIPTORS	Acoustics; Anatomy; *Articulation (Speech); Auditory Perception; Aural Stimuli; Electromechanical Technology; *Experiments; *Information Processing; Investigations; Japanese; *Language Research; Listening; Phonetics; Physiology; Psycholinguistics; *Speech; Tables (Data); Thai; Visual Perception; Vowels

ABSTRACT

This report on speech research contains papers describing experiments involving both information processing and speech production. The papers concerned with information processing cover such topics as peripheral and central processes in vision, separate speech and nonspeech processing in dichotic listening, and dichotic fusion along an acoustic continuum. Speech production topics covered include the activity of the intrinsic laryngeal muscles in voicing control, velopharyngeal function in oral/nasal articulation and voicing gestures, laryngeal adjustments for vowel devoicing in Japanese, vowel stress and articulatory reorganization, tension in some American English vowels, and word-final stops in Thai. One paper in the collection concerns reading machines for the blind. The final paper discusses the evolution of human speech anatomy. A list of publications and reports is provided along with author and title indexes to the previous status reports on speech research. (VM)

SR-28 (1971)

SPEECH RESEARCH

A Report on
the Status and Progress of Studies on
the Nature of Speech, Instrumentation
for its Investigation, and Practical
Applications

1 October - 31 December 1971

Haskins Laboratories
270 Crown Street
New Haven, Conn. 06510

Distribution of this document is unlimited.

(This document contains no information not freely available to the general public. Haskins Laboratories distributes it primarily for library use.)

ACKNOWLEDGMENTS

The research reported here was made possible in part by support from the following sources:

Information Systems Branch, Office of Naval Research
Contract N00014-67-A-0129-0001

National Institute of Dental Research
Grant DE-01774

National Institute of Child Health and Human Development
Grant HD-01994

Research and Development Division of the Prosthetic and Sensory Aids Service, Veterans Administration
Contract V101(134)P-71

National Institutes of Health
General Research Support Grant RR-5596

National Institute of Child Health and Human Development
Contract NIH-71-2420

CONTENTS

I. Manuscripts and Extended Reports

III. Cumulative Index (SR-1 - SR-28)

On Peripheral and Central Processes in Vision: Inferences from an Information-Processing Analysis of Masking with Patterned Stimuli

Michael T. Turvey
Haskins Laboratories, New Haven*

ABSTRACT

The masking of briefly exposed letter forms by a preceding or succeeding stimulus may originate in either peripheral or central visual mechanisms. The question of how masking varies with origin was examined in a series of experiments which made use of stimuli that masked the target forms only monoptically (or binocularly), or both monoptically and dichoptically. Peripheral forward and backward masking were described by a simple relation between target stimulus energy and the minimal interval between target offset and mask onset permitting evasion of masking: the minimal interval multiplied by the target energy equals a constant. Peripheral forward masking, however, was more sensitive to mask intensity than was peripheral backward masking. On the other hand, central masking, which was primarily backward, was relatively unaffected by stimulus energy and was determined by the interval elapsing between the onsets of the two stimuli. The multiplicative rule and the onset-onset rule characterized, respectively, peripheral and central visual processes. The peripheral processes were viewed as a set of parallel systems or nets signalling crude features of the stimulus, and the central processes as a series of decisions conducted, in part, on these features and resulting in stimulus recognition. The peripheral and central processes were shown to be related in a concurrent and contingent fashion: apparently the two occur in parallel, with the central decisions contingent on the output of the peripheral systems which signal different features at different rates.

INTRODUCTION

Perceptual interference results when two stimuli are delivered to an observer in rapid succession. The term "forward masking" describes the impairment in the perception of the second stimulus induced by the first, and the term "backward masking" describes the interference on the first induced by the second. The phenomena of forward and backward masking are evident

---

*Also University of Connecticut, Storrs.

Acknowledgments: The author wishes to acknowledge the assistance and insights provided by Dr. G. Lukatela, Professor of Electrical Engineering, University of Belgrade, and by Dr. C.J. Darwin, Lecturer in Psychology, University of Sussex, during the course of their stay as visiting research associates. Special acknowledgment is given to Diane Kewley Port for her invaluable help with the theoretical notions contained in the present paper.

in both aural and visual perception (Kahneman, 1968; Raab, 1963; von Békésy, 1971) and they occur, to varying degrees, under conditions where the two stimuli are presented to opposite ears (i.e., dichotically) or eyes (i.e., dichoptically) and under conditions where both stimuli are presented to the same ear or eye. Thus the perceptual interference may originate in the peripheral sense organ or in the more complex structures of the brain. Presumably the rules determining masking differ with the origin of the effect. In the present paper, masking in vision is examined for the purpose of isolating these differences.

Backward Masking and Information Processing

Backward masking of form by visual pattern or visual noise has recently received considerable attention primarily because of the central role it plays in the information-processing approach to visual perception (see Haber, 1969a). In brief, the information-processing analysis represents visual perception as a hierarchically organized temporal sequence of events involving stages of storage and transformation of information. Within this framework backward masking by pattern or noise is proposed as an analytic tool with which to investigate visual perception (Haber, 1969b; Sperling, 1963). The principle argument behind that proposition is that if a pattern mask follows a target stimulus after some delay, processing is assumed to have occurred during that delay but is terminated or interfered with by the mask. This argument is, essentially, the interpretation forwarded by Baxt (1871) for backward masking and, following Kahneman (1968), will be called an interruption hypothesis.

An alternative interpretation of masking by pattern is also emphasized in the literature. This interpretation, referred to as an integration hypothesis (Kahneman, 1968), stresses the effect that a visual pattern has on the sensory character of the target stimulus representation rather than on the extraction of information from the target representation. The idea is that two stimuli which follow one another in rapid succession are effectively simultaneous within a single "frame" of psychological time, analogous to a double exposure of a photographic plate. In masking by homogeneous flash of light, for example, the outcome of such a process of summation will be a reduced level of contrast between figure and ground (Eriksen and Hoffman, 1963). As Kahneman (1968) has pointed out, this position views masking by pattern as just a special case of temporal summation of heterogeneous stimuli.

Figure 1 summarizes the essential features of a visual information-processing system appropriate to the description of performance in tachistoscopic experiments. It contrasts the interruption and integration hypotheses.

Iconic storage (Neisser, 1967) is seen as a buffer memory system in which the input can be held in a literal form for several hundred milliseconds during the course of conversion to response and/or short-term categorical storage. Although the information in iconic storage is considered to be relatively unanalyzed, preattentive mechanisms have, perhaps, already extracted certain global features of the input (Neisser, 1967). These would include, for example, figure-ground relationships which provide the

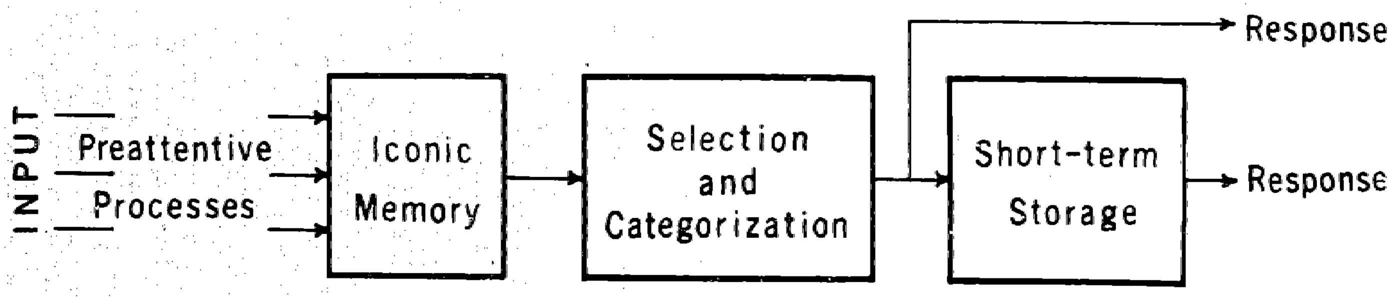


Figure 1: Schematic representation of the visual information-processing system.

raw material for subsequent selective processing of the iconic representation. This selective processing or recoding is demanded by the brevity of the iconic representation (Averbach and Coriell, 1961) and by the limited channel, or processing, capacity of subsequent mechanisms. At all events, the assumption is that the material in iconic storage is, in due course, recoded into categorical form for representation in response and/or short-term storage and that this recoding involves the processes of pattern recognition.

The interruption hypothesis localizes the effect of backward masking by pattern subsequent to iconic storage. It is assumed that a clear icon is established and that an after-coming pattern interferes with the translation into categorical form. The time needed to effect that translation is cut short by the after-coming stimulus. The integration hypothesis, on the other hand, proposes that target material and mask are dealt with as a composite, resulting in an unintelligible icon. For the integration hypothesis the effect of an after-coming pattern is on the formation of the target iconic representation so that it never achieves the acuity, contrast, or clarity that it would have attained in the absence of the mask.

Integration and Interruption as Nonexclusive Hypotheses

Comparisons between an integration story and interruption hypothesis of backward masking are usually made to decide which one is correct. It is, of course, not inconceivable that both are in fact true; they may be descriptions of two different stages in the flow of visual information. This possibility is suggested by the fact that two sorts of independent variables have been used in backward masking experiments. On the one hand, there are the energy properties of target and mask, i.e., duration and intensity; on the other, there is the time elapsing between onset of target and onset of mask. Sometimes backward masking has shown strict dependence on target duration or target-mask intensity (e.g., Eriksen, 1966; Kinsbourne and Warrington, 1962a; Thompson, 1966), yet at others it has shown strict dependence on onset-onset time with stimulus variables such as target duration proving irrelevant (e.g., Haber and Nathanson, 1969; Mewhort, Merikle, and Bryden, 1969). It is possible that when target energy (and/or mask energy, for that matter) is the relevant independent variable, mechanisms underscored by the integration hypothesis are prevailing, but when onset-onset time is the determining variable and target energy properties are irrelevant, interruption is perhaps the more appropriate theory.

The above ideas guided the present series of experiments, for which the experiments of Kinsbourne and Warrington (1962a, 1962b) provided a departure point. Their experiments were interesting in several important respects. First, with a paradigm fundamentally similar to that used, for example, by Sperling (1963), Kinsbourne and Warrington (1962a) found no evidence that in the backward masking situation the number of items reported is a linear function of onset-onset time. That result reported by Sperling (1963) and Allport (1968), among others, may be viewed as evidence for a process of sequential read-out from an intact iconic representation and as support for the interruption hypothesis. Kinsbourne and Warrington by contrast reported that not only did three letters become available at approximately the same onset-onset time as one letter, but also that there was a simple relation between target duration and the minimal interstimulus interval which permitted evasion of the masking action: target duration x interstimulus interval = a constant. That result, as Kahneman (1968) has pointed out, has not subsequently been investigated. It is an important result because the observation that the minimum interval permitting perception varies inversely with target stimulus duration is the very stuff out of which an integration hypothesis is made. Masking is determined by properties of the stimuli, not by the time elapsing between the onsets of the stimuli. The present series of experiments began, therefore, with a partial replication of the experiments of Kinsbourne and Warrington (1962a).

GENERAL METHODOLOGY

What follows is a brief description of some of the terms used in the present communication and some general comments on procedure, apparatus, stimulus materials, and subjects.

Terms

(i) Target (T) refers to the stimulus which S (the subject) is required to identify.

(ii) Random noise (RN) refers to a masking stimulus such as that shown in Figure 2. A mask of this sort had been used in the Kinsbourne and Warrington experiments. Described by Kinsbourne and Warrington as a noninformational stimulus, the mask reproduced in Figure 2 is a section of the random pattern (visual noise) described by Laner, Morris, and Oldfield (1957) --type 80 units/sq. cm. In the present experiments the size of the visual field subtended by RN was 3.5° vertical by 6.5° horizontal.

(iii) The term "pattern mask" (PM) was reserved for masks other than RN.[1] An example of PM is given in Figure 2. An essential feature of PM reproduced in Figure 2 is that the lines comprising the task were of the same thickness as the T letters. All masks classified as PM shared this characteristic with the target material.

(iv) Time elapsing between offset of T and onset of mask field is referred to as interstimulus interval (ISI).

(v) The minimum ISI at which a masking field no longer affects T according to a predetermined performance criterion is referred to as the critical ISI (ISI_c).

(vi) The time elapsing between the onset of T and the onset of the masking field is referred to as stimulus-onset asynchrony (SOA) (see Kahneman, 1968).

(vii) The minimum duration of T that permits evasion of masking (at ISI = 0 msec), according to some criterion, is defined as the critical T duration.

---

[1] The use here of the term "pattern" is, of course, arbitrary; PM like RN is essentially a random arrangement.

Figure 2: Examples of masking and target stimuli.

Procedure

The procedure which was used for most of the present series of experiments *mutatis mutandis* is spelled out below in some detail. It is essentially the procedure used by Kinsbourne and Warrington (1962a, 1962b).

A T stimulus was presented to *S* for a brief period of time followed at varying intervals by the mask stimulus. The task of *S* was to identify the T stimulus. The duration of the T stimulus and the duration of the mask stimulus were held constant and ISI was increased from zero in steps of 2 msec to some value at which *S* correctly identified a T stimulus. At each ISI step, the T stimulus was changed whether *S* was correct or incorrect. The ISI value at which *S* correctly identified a T stimulus was left unchanged for the next T stimulus. If, however, S was incorrect, the ISI was increased by 1 msec for the following T presentation. This procedure was continued until *S* correctly identified four T stimuli in succession. The ISI value at which *S* identified four T stimuli in succession was designated as ISI_c. Kinsbourne and Warrington had defined ISI_c as that ISI at which *S* first reported a T stimulus correctly. However, it had been noted in pilot work that some letters were identified at briefer ISI's than others; that there were slight variations in the transmission values of T slides and thus fluctuations in T stimulus intensity; and that accordingly a criterion of one T stimulus correct did not specify accurately the ISI value at which masking was no longer occurring for all T stimuli.

A number of different T durations were employed, and for each, ISI_c was determined at several values of mask duration. In the course of determining the ISI_c's, several orderings of the T stimuli were used. In all experiments, for all T durations, identification was 100% accurate in the absence of the mask.

Apparatus

A six-channel tachistoscope (Scientific Prototype, Model GB) with automatic stimulus changers was used for the present series of experiments. The two three-channel optical units of the tachistoscope permitted monoptic and dichoptic presentation of stimuli, and one unit could be readily modified for binocular presentation. One of the two separate units was adjustable for interocular distance and convergence angle. The apparent viewing distance was 36 ft and the field of the tachistoscope subtended 3.5° vertical by 6.5° horizontal. Coarse intensity controls were available, but nonlinearities required the use of Kodak neutral density filters for accurate variation of stimulus luminance. Luminance was measured at the eyepiece by a spectra brightness spotmeter (Photo Research).

Stimulus Material

Three 100-slide sets of letter stimuli were constructed. The stimuli were in all cases positives of Stenso Gothic capitals. The positives, which were transparent, were held in 2" x 2", 35-mm. slide mounts. In one set the letters were located singly at the center of the slides. Those letters were the symmetrical letters of the alphabet (A, H, I, M, O, T, U, V, W, X, Y). In a second set those same letters were located singly to the side of the center. A third set consisted of consonant trigrams, i.e., three letters to a slide, selected from all the consonants. No consonant was repeated within a slide.

The letters in all sets subtended .67° vertical by an average .36° horizontal. The thickness of the letters subtended .05° visual angle. In the set of single letters displaced off center, the angular distance between the center of the slide (or fixation point) and the center of a letter was 1.37°. For the consonant trigrams, the separation between the letter edges was on the order of .40°. For all T stimuli, the field of view was 3.5° vertical by 6.5° horizontal.

Subjects

For the most part, Ss were Yale University students who were paid for their services. University of Connecticut graduate students and members of Haskins Laboratories also served as Ss in several experiments. All Ss had normal or corrected to normal vision.

EXPERIMENT I

Method

In Exp. I the T material was the set of centrally located symmetrical letters and the after-coming mask RN. Three durations of T were employed, 2, 4, and 6 msec, which were presented in this order for each S. For each T duration, ISI_c was determined at RN durations of 1, 2, 3, 4, 5, 6, 8, 10, and 50 msec. The luminance of T and that of RN was 15 ft L and the fixation field was .25 ft L.[2] The stimuli were delivered binocularly. Four Ss participated in the experiment; one S was not naive to masking phenomena.

Results and Discussion

The data of the four Ss are shown in Figures 3 and 4. Figure 3 shows the plot of ISI_c as a function of RN duration for each S. Figure 4 shows the ISI_c by RN duration functions averaged across Ss with T duration as the curve parameter.

As can be seen on inspection of the figures, the masking effect of RN varies in a discontinous fashion with its exposure duration. The effect of varying duration of RN achieves its maximum sharply. Increasing the RN duration beyond some value does not augment the masking effect, i.e., it does not extend the interval over which masking can be obtained. All of this concurs with the original observations of Kinsbourne and Warrington (1962a). Inspection of Figure 4 yields further corroboration of Kinsbourne and Warrington in that there exists a simple relation between T duration and ISI_c at asymptote: T duration x ISI_c = a constant. The picture is not as tidy as it might be; T duration x ISI_c does not yield exactly the same value at 2 msec as it does at 4 and 6 msec yet the values are close enough to favor the Kinsbourne and Warrington conclusion.

---

[2] The fixation field, i.e., the field immediately preceding and succeeding the T field, was at this level of luminance for each experiment in the present series.

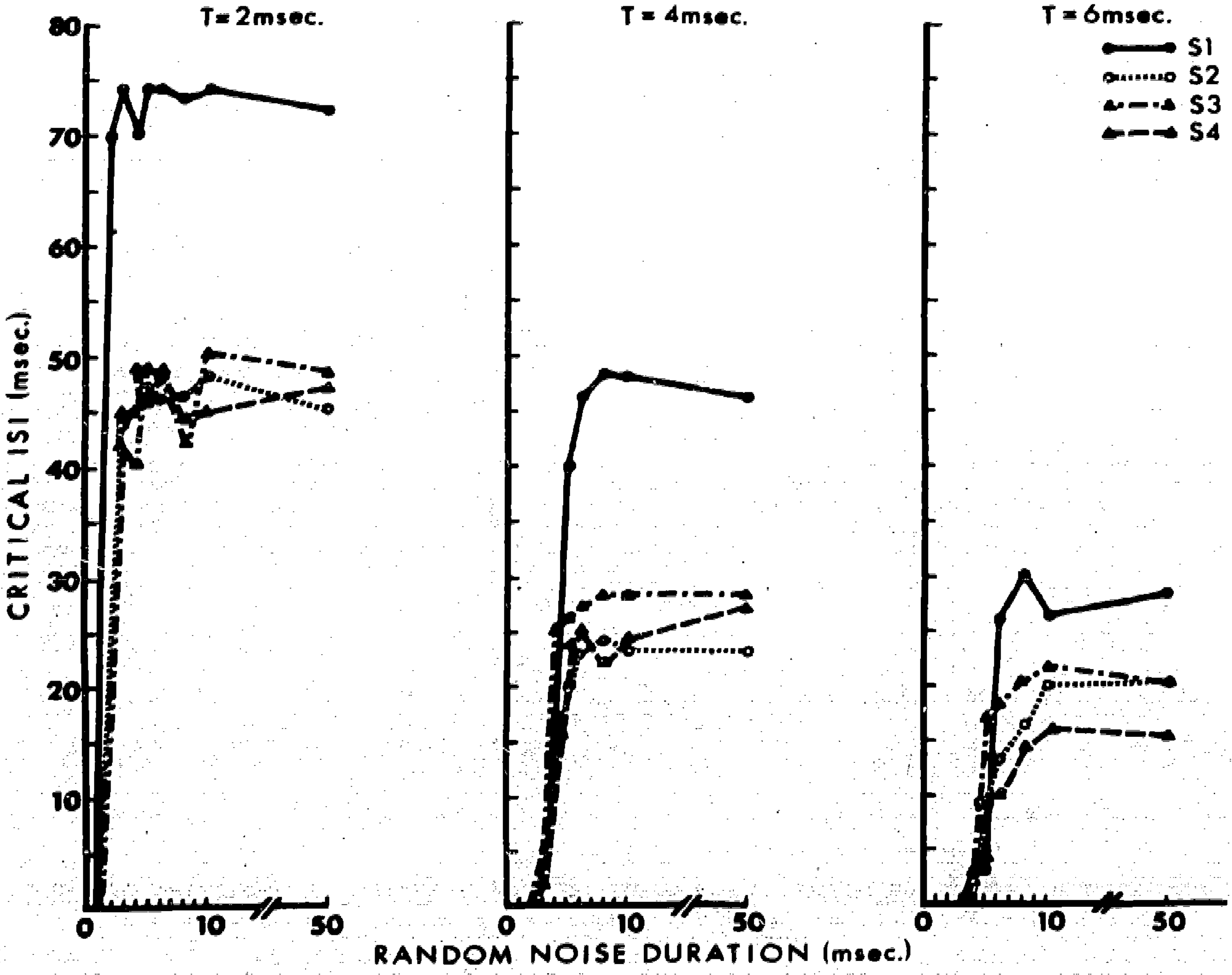


Figure 3: Relation between RN duration and ISI_c for binocular masking at three values of T duration for each S in Exp. I.

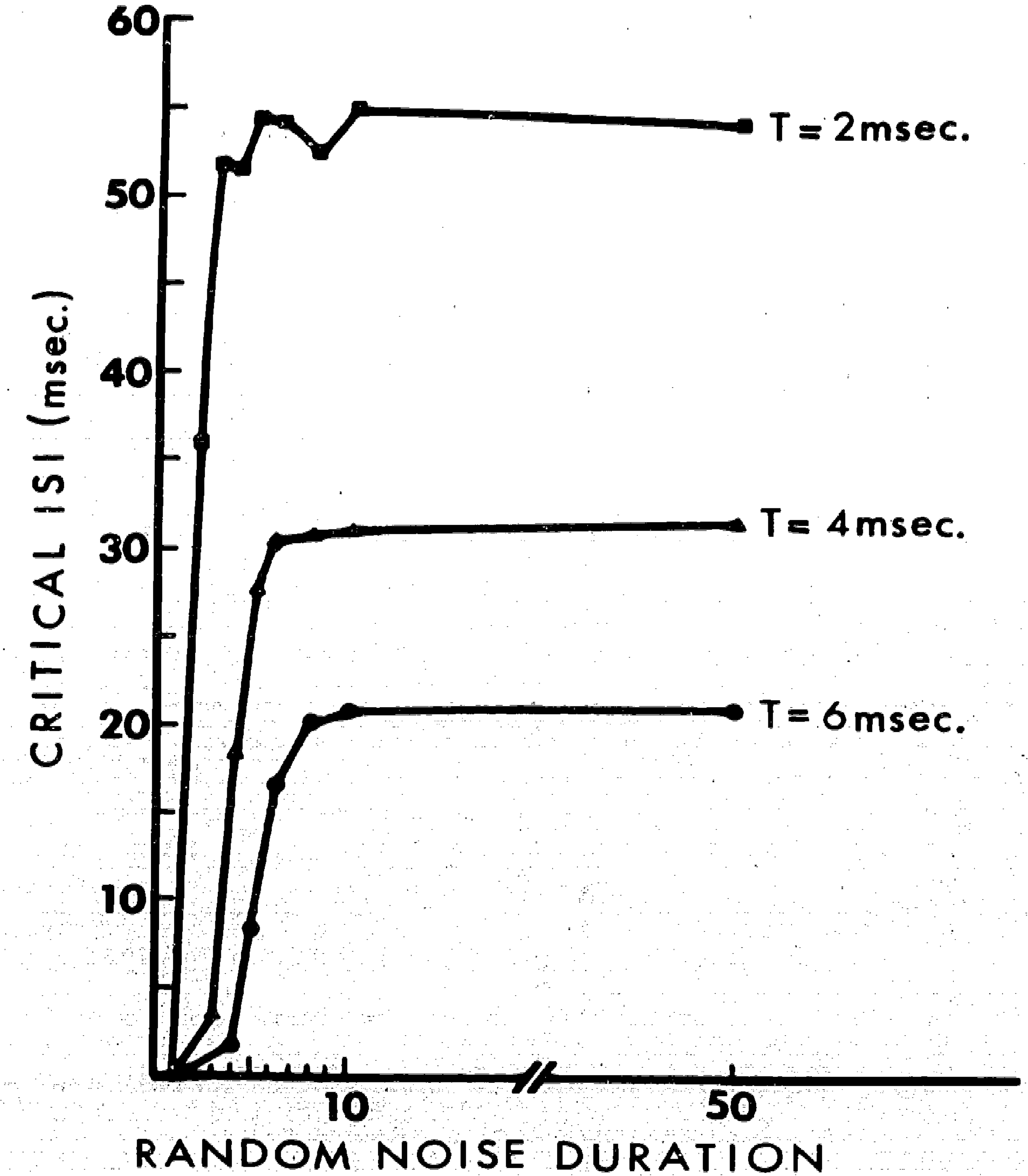


Figure 4: Relation between RN duration and mean ISI_c for binocular masking at three values of T duration in Exp. I.

Kinsbourne and Warrington (1962b) interpret this result as reflecting the fact that ISI_c is "the time which permits the perceptual process to deal with the two stimuli separately in succession, rather than simultaneously as a composite, and therefore unintelligible stimulus" (p. 235). It is quite evident that the formulation, T duration x ISI_c = a constant, argues strongly against onset-onset time and for T duration as the relevant parameter in masking by noise. There is, of course, the question of whether it is T exposure duration per se or the quantity of light in the stimulus that is important. The second experiment examined this question.

EXPERIMENT II

The time-intensity reciprocity law, known for human vision as Bloch's Law, says that a given effect can be achieved by the reciprocal manipulation of luminance and duration of a light flash. In the second experiment luminance of T was manipulated so as to produce a constant energy value for different exposure durations. If T energy rather than T duration was the important independent variable, then varying T exposure duration with energy held constant should not produce the inverse relation between T duration and ISI_c obtained in Exp. I; rather, ISI_c should remain constant. Such an outcome would indicate that the formulation, T duration x ISI_c = a constant, should be written: T energy x ISI_c = a constant.

Method

Experiment II was conducted in two parts. In Part 1, stimulus presentation was binocular. Two naive Ss were tested in the paradigm described in Exp. I. For both Ss, ISI_c was determined at several values of RN duration for two duration-intensity values of T: 2 msec, 20 ft L and 8 msec, 5 ft L. The T stimuli were the set of centrally located symmetrical letters. The luminance of RN was 15 ft L. In Part 2, presentation of stimuli was monocular. The stimuli were presented at the right eye. Two different naive Ss were tested in the manner described in Exp. I and Part 1 above. The values of T were 2 msec, 4 ft L and 4 msec, 2 ft L. The stimuli were the set of consonant trigrams. The definition of ISI_c in this case was four trigrams reported correctly in succession. The S had to report all three letters to be correct; correct order of letters, however, was not required. The intensity of RN was 15 ft L.

In both parts 1 and 2, order to T values was counterbalanced across the two Ss.

Results and Discussion

The data of the two Ss in Part 1 are given in the upper panels of Figure 5. The data of the two Ss in Part 2 are given in the lower panels of the same figure.

Comparison of the functions reproduced in Figure 5 with those in Figure 3, which give the data of Exp. I, suggests that T energy, not T duration, is the proper independent variable. In the upper panels of Figure 5, for example, ISI_c for both Ss was unchanged from a 2-msec to an 8-msec exposure duration of T. As Figure 3 from Exp. I shows, ISI_c for the exposure duration of 6 msec is significantly lower than ISI_c for the exposure duration of 2 msec. In Exp. I energy increased with increase in exposure duration;

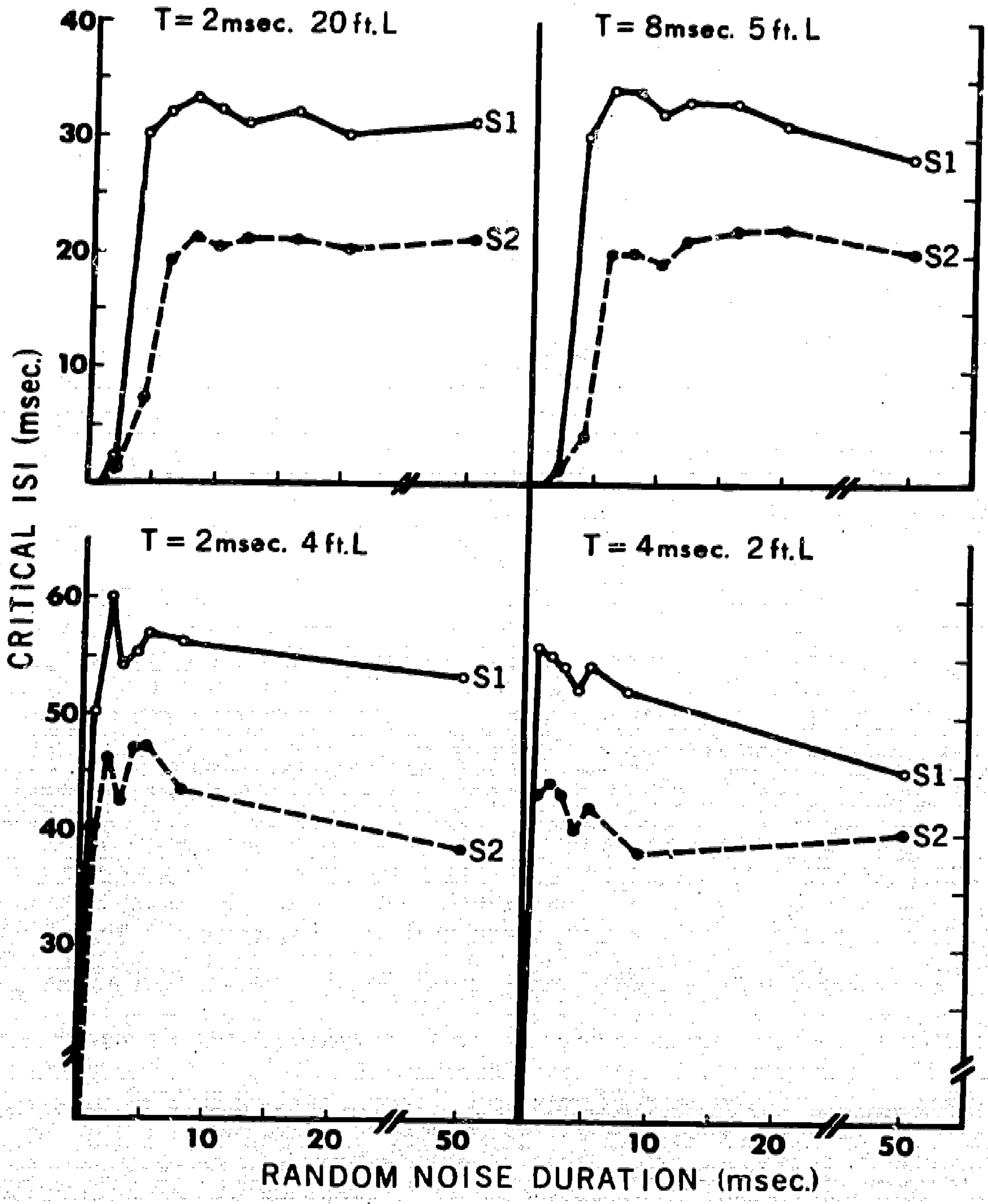


Figure 5: Relation between RN duration and ISI_c at various T duration-intensity combinations in Exp. II. Upper panel shows functions of the two Ss in Part 1; lower panel shows functions of the two Ss in Part 2.

in Exp. II total quantity of light was held constant across the exposure durations. Masking by RN, therefore, seemed to be very much a matter of stimulus energies. Further evidence to this effect was given in the observation that the minimal RN energy needed to mask a T stimulus was directly related to the T energy as can be seen by inspection of Figure 4.

The masking reported in Exps. I and II is, perhaps, explicable in terms of the lack of fine temporal resolution in the visual system. That is to say, stimuli presented in succession with too brief an interval elapsing between them are perceived as essentially simultaneous, an interpretation of masking proposed by Kinsbourne and Warrington and championed in a multitude of articles by Eriksen (e.g., Eriksen, 1966; Eriksen and Collins, 1965; Eriksen and Hoffman, 1963). The best-known example of this lack of temporal resolution is Bloch's Law: within some critical period, usually of the order of 100 msec, time and intensity of stimulation can be reciprocally interchanged without changing the visual effect. Research by Davy (1952) has shown, at least for the periphery of the retina, that such integration of energy over time by the visual system is independent of the distribution of energy within this period. The reciprocity between luminance and duration in the rule, T energy x ISI_c = a constant, suggests that the masking demonstrated in Exps. I and II may be another manifestation of the visual processes underlying Bloch's Law.

The explanation of masking by RN as due to the lack of fine temporal resolution in the visual system implies that T and RN are treated as a single package presumably at some later stage in the processing of visual information. The impairment in the perception of T may be attributed to a confusion of features or contours or to a change in the minimum acuity requirements (Eriksen and Collins, 1965; Purcell, Stewart, and Dember, 1968). It may also be due to summation of T luminance with RN luminance. Luminance summation would reduce the contrast between the T form and its background, thereby impairing detection and identification (e.g., Thompson, 1966). In any event, the argument is that the resulting representation of T is degraded.

It should be noted that an explanation closely related to the integration hypothesis described above may also account for the masking observed in Exps. I and II. This explanation assumes that the masking stimulus overtakes or smears the discriminability of T by "catching up" with T somewhere in the transmission channel (Crawford, 1947; Fry, 1934; Leibovic, 1968; Stigler, 1910). Essential to "overtake" hypotheses is the requirement that the masking stimulus be more intense than the T stimulus. The latency of retinal and cortical responses to stimulation is inversely related to stimulus intensity (Monnier, 1952). Thus a mask will travel at a greater speed between receptor and cortex than a T of less intensity. Elegant data favoring an overtake hypothesis for backward masking by a flash of light have recently been reported by Schiller (1968). In single-cell recordings in the lateral geniculate nucleus of the cat, Schiller (1968) observed that cells which respond at their maximal level to the mask stimulus fail to register the earlier display of the T stimulus. In certain ways, as pointed out by Kahneman (1968), the overtake conception may be described as an integration theory: it assumes a nonlinear summation of response rather than a linear summation of stimuli.

At all events, there must be serious reservations about the utility of "overtake" as an explanation of the masking by RN. That T and RN in Exp. I were of equal intensity suggests that "overtake" is not appropriate to that experiment. What of Exp. II? In three of the conditions of the two parts of Exp. II, RN intensity was greater than T intensity. In one condition RN intensity was less than T intensity. Inspection of Figure 5 does not reveal any difference between condition T = 2 msec, 20 ft L, and the other conditions to suggest that mask intensity was the crucial variable.

EXPERIMENT III

The third experiment primarily compares the severity of backward masking by RN under monoptic and dichoptic presentation. It also looks for differences in ISI_c as a function of the hemisphere receiving the stimuli. In the monoptic conditions, T and RN were delivered to the same hemiretina. In the dichoptic conditions, T and RN were presented to different hemiretinas but to the same hemisphere.

Method

The procedure was identical in most respects to that described in Exp. I. For monoptic and dichoptic delivery of inputs, both of the three-channel units of the tachistoscope were used. The two fields of view, one for each eye, both contained a centrally located fixation point and were set at the same luminance, .25 ft L.[3] Whether S was receiving a monoptic or dichoptic sequence, he was required to view with both eyes. Both T and RN appeared on the same side of the fixation point: in the monoptic condition they came to the same eye, and in the dichoptic, to different eyes. The T stimuli were the set of symmetrical letters displaced off center 1.37°. The RN subtended 2.25° horizontal by 1.5° vertical with its inner edge bisecting vertically the fixation point.

Four Ss participated in the experiment. Two Ss were not naive. Two Ss received dichoptic conditions followed by monoptic; the other two Ss received monoptic followed by dichoptic. Each S received one of the four orders of the dichoptic conditions in a partially counterbalanced design in which each condition was tested once across Ss in each test-order position. Stimuli presentations were not mixed; S always knew that within a condition, T and RN would always appear in the same half of the visual field, say the left.

For each S, the interocular distance of the two eye pieces was adjusted to facilitate convergence of the two fixation points. The Ss were required to converge the two fixation points prior to presentation of stimuli. The Ss were told to indicate to the experimenter any occasion on which

---

[3] This was the case for each dichoptic presentation condition described in the present paper.

they were aware of their eyes moving off the converged fixation point prior to stimulus presentation. Involuntary eye movements do occur during fixation; however, the work of Riggs, Armington, and Ratliff (1954) indicates that during a 10-msec exposure the typical excursion is less than 5 sec of arc.

For all conditions the exposure duration of T was 4 msec. The exposure durations of RN were 1, 2, 3, 4, 5, 6, 8, 10, and 50 msec. Critical ISI was determined by the usual procedure at each RN duration in the order shown. Throughout, T and RN were of equal luminance, 10 ft L.

Results and Discussion

The most important feature of Exp. III was the failure to obtain masking in any of the dichoptic conditions. In all the dichoptic conditions, S was able to identify the T letters at any ISI value in the range 0 msec to 300 msec, at SOA = 0 msec, and at any exposure duration of RN ranging from 1 msec to 500 msec. (All these results were confirmed subsequently with several other Ss.) The data for monoptic presentation are given in Figure 6. All functions are in accord with that observed in Exp. I for T = 4 msec.

The mean of the ISI_c's at RN durations of 5, 6, 8, 10, and 50 msec were computed for each S in each of the four conditions. These means were submitted to a Treatment x Ss analysis of variance. The main effect of transmission line was significant, $F(3,9) = 6.39$, $p < .05$.

Inspection of Fig. 6 suggests that (i) ISI_c for the nasal transmission lines was less than that for the temporal transmission lines and (ii) ISI_c was less for stimuli presented in the right visual field, i.e., to the left hemisphere. Both suggestions are in agreement with the general body of data on laterality differences with unilateral presentation conditions (see White, 1969). No further comment will be made on these laterality data. More important for present purposes is an examination of possible reasons underlying the absence of dichoptic masking by RN.

Although S in dichoptic conditions could identify T stimuli without difficulty, T was not completely unaffected by RN. Further investigation revealed that at exposure duration and luminance close to threshold, identification and/or appearance of T stimuli could be impaired by RN at SOA = 0 msec. An increase in T duration of the order of several milliseconds would be enough, however, to overcome that effect of RN. The effect of RN in the dichoptic mode was at best a very modest one. Since in dichoptic presentation T and RN can only interact centrally, the conclusion must be drawn that the locus of masking by RN as observed in Exps. I and II was primarily, if not solely, in the peripheral visual system.

Backward masking of forms in the dichoptic mode has been reported in the literature (e.g., Schiller, 1965; Schiller and Wiener, 1963; Smith and Schiller, 1966). The effect, however, is restricted to masks which contain contours; a homogeneous flash of light does not mask forms dichoptically (e.g., Mowbray and Durr, 1964; Smith and Schiller, 1966; Schiller and Wiener, 1963).[4] Thus, in the present experiments RN is operating like a

---

[4] Flashes of light may yield slight dichoptic effects, but they depend for the most part on the use of near-threshold T stimuli and the relatively close proximity of T and mask borders (Battersby and Wagman, 1962; Boynton, 1961).

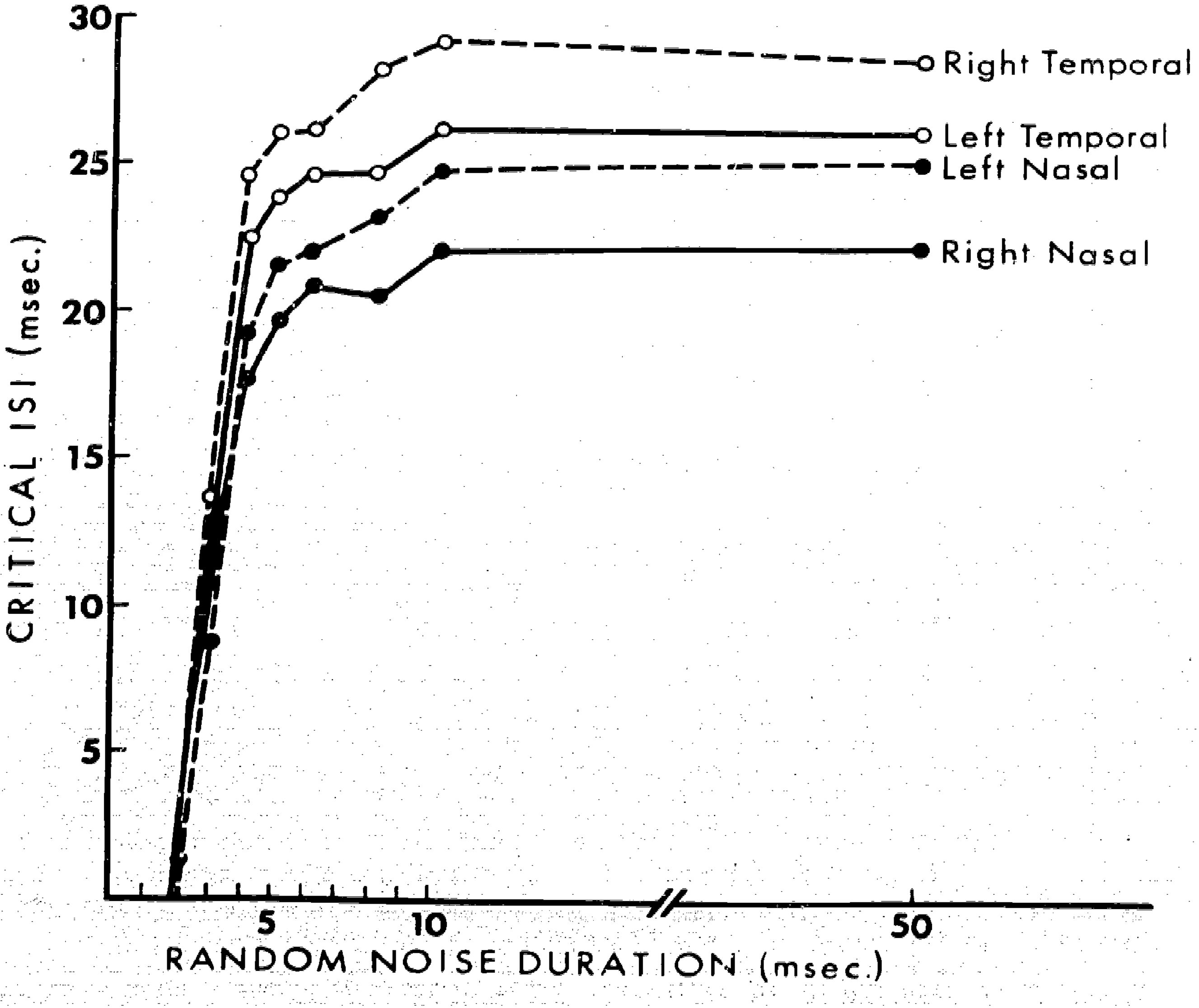


Figure 6: Relation between RN duration and mean ISI_c as a function of the hemiretina receiving the stimuli in Exp. III.

homogeneous field of light. A first guess, therefore, was that failure to confirm masking in dichoptic presentation in the present experiment was in some part due to the relation between T and RN. Several investigators have commented on the fact that in many instances masking is highly form-specific (e.g., Buchsbaum and Mayzner, 1968; Fitzgerald and Kirkham, 1966; Houlihan and Sekuler, 1968; Parlee, 1969; Schiller, 1965; Sekuler, 1965). There was little, if any, formal similarity between T and RN in the present experiment. It is, however, important to note that with almost the same type of stimuli (T and RN), Kinsbourne and Warrington (1962b) did obtain masking with dichoptic presentation. For these reasons, various stimuli were examined in the dichoptic mode, including other random noise displays. An initial observation was that the inverse of the RN mask did produce a fairly significant effect dichoptically. Figure 7 probably shows why. The inverse of RN has much larger dark regions which approximate the thickness of the T letters. However, the inverse of RN was not considered to be as effective a mask as some other stimulus patterns, one of which is also shown in Figure 7.

A pattern of lines of the same thickness as T letters located only in the region of the display field occupied by a T letter was eventually selected for further investigation of masking in dichoptic regard. The pattern mask (PM) is shown in Figure 2. Casual investigation revealed that such a pattern was an effective dichoptic mask.

EXPERIMENT IV

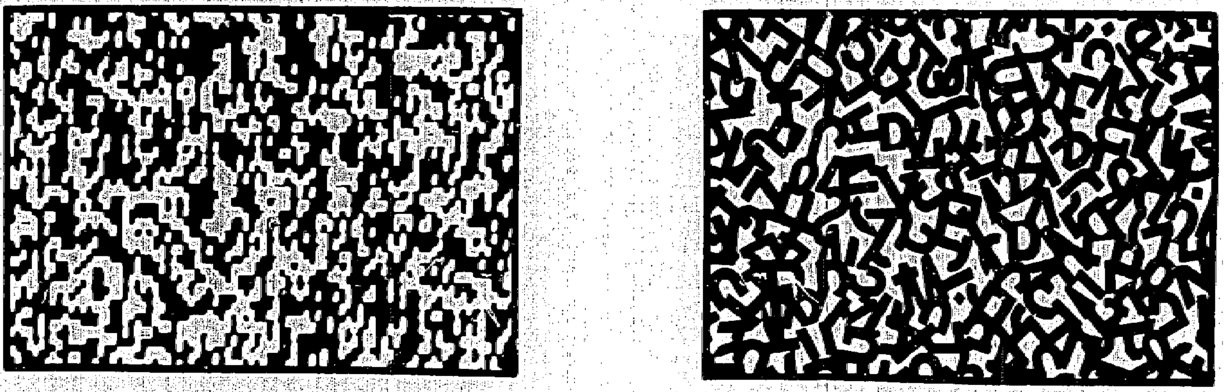
Experiment IV looked at the relation between RN and PM. Specifically it asked whether it was correct to assume that RN influenced a stimulus only if it followed on the same transmission line and, therefore, differed from PM, which could have a central influence.

Method

As in Exp. III, T and mask stimuli were presented to hemiretinas. The T stimuli were presented in the right visual field of the left eye, i.e., on the left temporal transmission line. The RN and PM stimuli described above were used as masks. The RN stimulus was presented in either the right visual field of the right eye or the right visual field of the left eye, i.e., on the right nasal or left temporal transmission line. The PM stimulus was presented in the right visual field of the right eye (see Figure 8).

Exposure durations of T, PM, and RN were 4 msec, 4 msec, and 10 msec, respectively. Pilot work had shown that a 4-msec exposure of PM could effectively mask dichoptically a 4-msec exposure of T within a relatively large ISI range. Previous experiments, Exps. I and III for example, had already shown that RN of 10-msec duration effectively masked a preceding 4-msec exposure of the same intensity on the same transmission line if the two stimuli were separated by an ISI of less than about 30 msec. The intensities of T, PM, and RN were equal at 15 ft L. The pattern of lines constituting PM were displaced off-center on a slide so as to cover the area in the field occupied by the set of off-center symmetrical letters which were the T stimuli.

There were five conditions, which are reproduced in Figure 8.



INVERSE OF RN

LETTER FRAGMENTS
(approximately 50% density)

Figure 7: Further examples of masking stimuli.

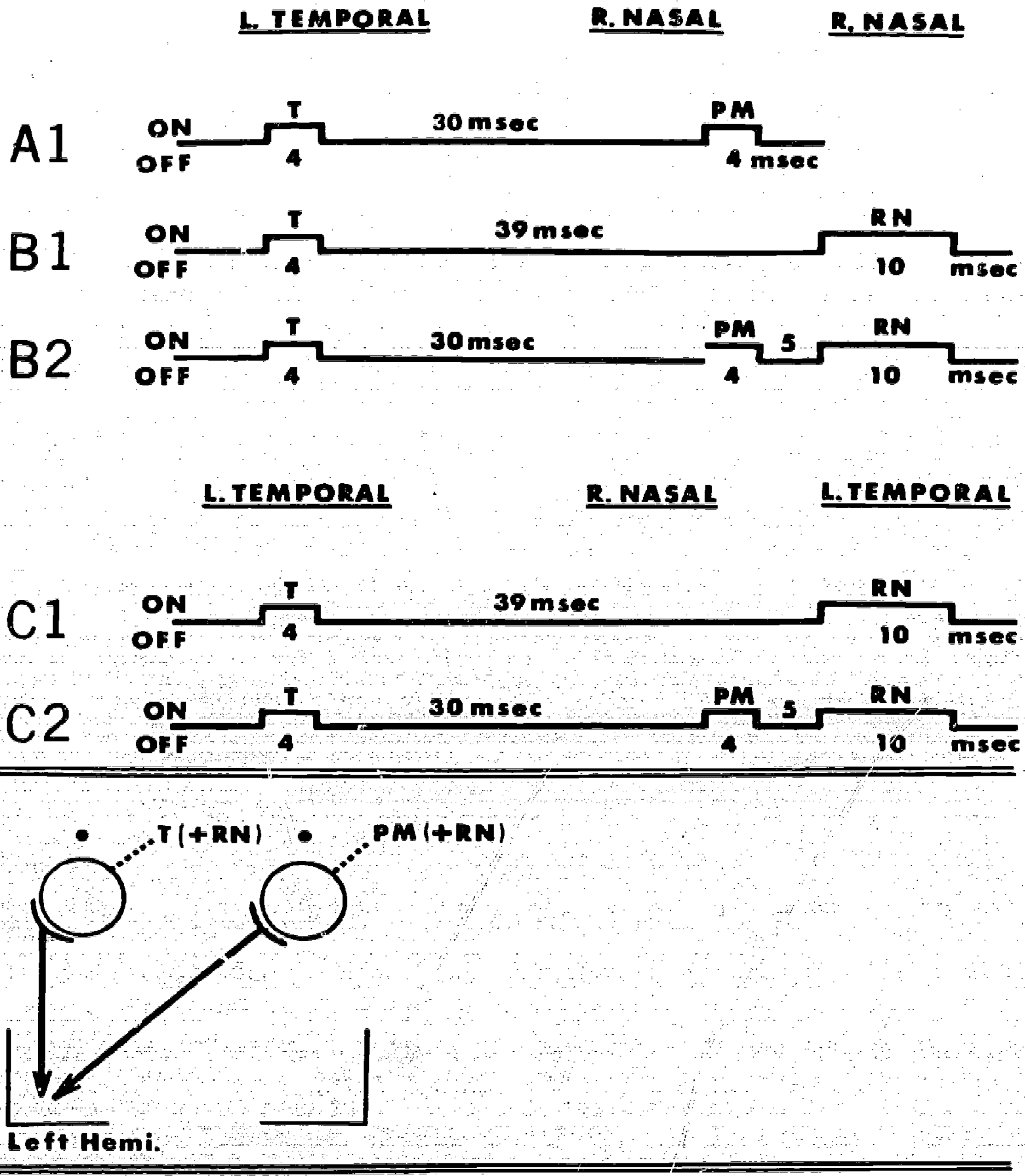


Figure 8: Order and mode of stimulus presentation in Exr. IV.

Each of six Ss were tested in all five conditions with twenty stimuli presented for identification in each condition. All six Ss went through the five conditions in sequence (i.e., A1, B1, B2, C1, C2) four times, with five observations made each time in each condition. Predictions based on the assumption that RN operated only peripherally were as follows: (a) In condition A1, T would be masked by PM. (b) In conditions B1 and C1, T would be seen and identified against RN as background. Note that in B1 the relation between T and RN was dichoptic; in C1 both stimuli were presented at the left temporal hemiretina. The particular ISI value was chosen to insure that RN would not monoptically mask T in condition C1. (c) In condition B2, T would be seen and identified against RN as background because RN would mask PM in the transmission channel, thereby preventing the central interference of T by PM. On the other hand, in condition C2, T would still be masked by PM. Note that the only difference between B2 and C2 was that in the former, RN was on the same transmission line as PM.

Results and Discussion

The data of all six Ss conformed to the predictions. All six Ss failed to identify any of the twenty letters in conditions A1 and C2; since Ss were not told to guess their typical response was "nothing." All six Ss identified every letter in condition B2. As a check on the phenomenon, to see if asynchrony of the three stimuli was essential to the effect (cf., Robinson, 1968), each S was tested in conditions A1, B1, and B2 with the three stimuli delivered simultaneously (i.e., SOA = 0 msec). The results of simultaneous presentation were identical to those of successive presentation: no masking in conditions B1 and B2, complete masking (i.e., no identification of the T letters) in condition A1. Temporal separation of stimuli was therefore not necessary for the effect. Further investigation also showed that this "recovery of T effect" could be obtained with T and masks presented to the entire retina rather than hemiretinas. The phenomenon was easily demonstrated informally many times subsequent to the experiment.[5]

"Disinhibition" effects are not uncommon in the literature on masking. Robinson (1966, 1968), Dember and Purcell (1967), Purcell and Dember (1968), and Schiller and Greenfield (1969) have all demonstrated that the masking effect of a stimulus can be inhibited by a subsequent stimulus. A mechanism proposed in the literature (e.g., Robinson, 1968) for the "recovery of T" phenomenon is lateral inhibition, which is often expressed as the likely mechanism underlying masking in general (see Weisstein, 1968). Although some investigators have criticized a lateral inhibition explanation of masking (e.g., Eriksen and Marshall, 1969; Kahneman, 1967a; Uttal, 1970), it remains for the most part a forceful explanation of disinhibition. In fact Weisstein (1968) views "recovery of T" or disinhibition experiments as direct tests of a lateral inhibition model of masking.

Lateral inhibition refers to the suppression of neural response by neighboring neural responses. Consider a stimulus delivered to the "on"

---

[5] A homogeneous light flash of energy greater than T produced the same effect when substituted for RN.

region of a receptive field of a cell--the cell fires above its normal spontaneous rate. If the stimulus had been delivered in the "off" region, the cell would fire below its normal rate. If both the "on" and "off" regions are stimulated simultaneously, the cell fires neither at the onset nor at the offset of stimulation. The lateral inhibition explanation of masking assumes that the neurons responding to the mask inhibit the neurons responding to the T stimulus. In its barest essentials the lateral inhibition explanation of the "recovery of T" is that the responses of the neurons stimulated by the second mask suppress the responses of the neurons stimulated by the first mask, thus freeing the T stimulus (or rather the neurons responding to T) from the inhibiting influence of the first mask. Presumably the cells responding to the second mask are not particularly close neighbors of those responding to T.

In the present experiment the masking of T by PM was of central origin; the masking of PM by RN was of peripheral origin. This cannot be explained by a model such as that proposed by Weisstein (1968), which must maintain that the masking of T by PM and of PM by RN occur at the same locus with the same underlying neural net. Comparing condition C2 with B2 clearly shows that the two masking effects did not have the same locus. Robinson (1968) reported a failure to elicit disinhibition when the disinhibiting second mask was delivered to the eye that did not receive the T stimulus and the first mask. This result was interpreted by Robinson to mean that disinhibition could be obtained only when all three stimuli were input to the same eye, and that disinhibition was, therefore, due to recurrent lateral inhibition influences in the retina. Obviously, that interpretation cannot apply to the present data.

In short, the present experiment brings into question the appropriateness of the term "disinhibition" and the concept of lateral inhibition as applied to the "recovery of T" phenomenon.

Kolers's Clerk-Customer Analogy

A preferred approach to the data of Exp. IV and to those of the experiments which follow is given in an analogy proposed by Kolers (1968). "A customer who enters a store is usually treated as fully as the attending clerk can treat him; a second customer then entering, the clerk tends to shorten the amount of time he spends with the first. In a store whose customers enter aperiodically, the amount of treatment given to anyone depends upon whether a second enters; if he does, treatment of the first is usually shortened. In this analogy, the visual inputs are the 'customers' and the central processor the 'clerk'" (p.38). The analogy is revealing. It would suggest that in the present experiment the loss in perceptibility of T when PM is presented cannot be because T is "erased." On the contrary, T may persist but what is known of T is limited. The clerk can find out a great deal from his customer: how he feels today, how the wife is, whether he wants brand X or brand Y, etc. With the appearance of another customer, however, much of this is left undone. If the second customer is particularly compelling and close on the heels of the preceding customer, the clerk may come to know very little, if anything, of his first customer's dispositions and wants. The analogy is further illuminating in that it implies that RN prevented PM (condition B2) from gaining access to the store housing the clerk, or central processor. For the analogy, the impairment in the perception of T by PM was not due to interference between the inputs, or customers, rather it was the result of their effect upon the central device, or clerk. On the other hand, the loss of perceptibility of PM, and consequently, the loss of its masking effect on T, might have been due to degradation by RN with this interference taking place in the transmission channel itself.

It is evident that RN must gain access to a central processor. In the present experiment in conditions B1 and B2, for example, T and RN were seen clearly by S; in condition C2, S saw PM and RN clearly. Introspective accounts were that T or PM appeared "through" RN or "on top of" RN. This would suggest that figural analysis or synthesis (depending on one's predilections) of T (or PM) and RN were accomplished in parallel by different processors or neural systems (cf., Liss, 1968) or concurrently by the same processor. Indeed, RN should function as a central mask for some stimuli. All this implies that masking with dichoptic presentation occurs whenever the analyses of both T and the mask require the use of the same central mechanism, or the same components of a single central mechanism, and not otherwise. On the other hand, binocular and monocular masking, where peripheral interaction can occur, may not be so dependent on formal similarity between T and the mask.

EXPERIMENT V

Kinsbourne and Warrington (1962b) reported that the relation, T duration x ISI_c = a constant, described masking functions for dichoptic, as well as monoptic, presentation. Experiments I and II of the present series taken together imply that the proper independent variable for masking by RN was not duration of T but rather the total quantity of light in the T exposure. The Kinsbourne and Warrington relation was therefore rewritten: T energy x ISI_c = a constant. As Exps. III and IV showed, however, the origin or locus of the interfering effect of RN on the perceptibility of T was in the transmission line. Perhaps, then, the relation, T energy x ISI_c = a constant, speaks only to peripheral interaction, contrary to the report of Kinsbourne and Warrington. The variation on Kolers's (1968) clerk-and-customer metaphor described above hints at a difference between masking originating outside the store (peripherally) and masking originating inside the store (centrally). There are also several sources of data which suggest that masking under conditions of dichoptic presentation differs in a fundamental and interesting way from monoptic masking. Boynton (1961) and Schiller (1969) report experiments showing that dichoptic masking is relatively independent of stimulus intensity.

Method

The design of Exp. V was comparable to that of Exp. I, but stimulus presentation was dichoptic and PM was the after-coming mask. The configuration of lines used for PM was centrally located in the mask field, the T material was the set of centrally located symmetrical letters and both their luminances were 4 ft L. (Thus, the present experiment contrasts with the preceding two in that presentation was to retinas rather than to hemiretinas.) Four naive Ss were presented with T to the left eye and PM to the right eye. For two durations of T, 4 msec and 10 msec, ISI_c was determined for the following exposure durations of PM: .5, 1, 2, 3, 4, 5, 6, 8, 10, 25, and 50 msec.

Results and Discussion

The functions relating ISI_c to PM exposure duration for the two durations of T are reproduced in Figure 9. Each data point represents the average of the four Ss.

The most important aspect of Figure 9 is the absence of any dramatic separation between ISI_c for the two exposure durations of T. Compare this figure with the data of Exp. I, in which stimulus presentation was binocular and the after-coming mask was RN; there the ISI_c separation between T = 2 msec and T = 6 msec in Figure 4 was about 35 msec.

The data reproduced in Figure 9 shows that the relation, T duration x ISI_c = a constant, does not describe dichoptic masking by PM. If that relation was in effect, then ISI_c for T = 10 msec should have been on the order of, at most, 14 to 16 msec, given that the mean ISI_c for T = 4 msec was about 36 msec. Again it should be noted that Kinsbourne and Warrington (1962b) did find that the relation, T duration x ISI_c = a constant, held for masking in the dichoptic mode. The reason for the disparity between the data of the present experiment and those of Kinsbourne and Warrington is unclear.

Two other aspects of Figure 9 deserve comment. First, ISI_c tends for both T exposure durations to decrease with an increase in PM duration beyond 10 msec. Subsequent informal experiments revealed that this was a fairly common occurrence. One hypothesis about this somewhat unexpected observation was that it perhaps reflected the dependence of masking in the dichoptic mode on stimulus-offset asynchrony. However, phenomenological description suggested an alternative possibility. At the longer durations of PM, Ss reported that the field surrounding the configuration of lines (see Figure 2) was very bright and that the pattern itself appeared degraded. As a check on the importance of duration per se, the luminance of PM at the longer durations was reduced. The result was the ISI_c remained relatively invariant across mask durations from 10 to 100 msec.

It seems, therefore, that as exposure duration increases and intensity is held constant, a display such as PM, consisting of a figure on a ground, may partially mask itself. The surrounding bright area may degrade the form of the central dark area. This effect probably takes place on the transmission channel itself rather than centrally. Purcell, Stewart, and Dember (1969) have made a similar observation: within certain limits, increasing luminance or duration increases the susceptibility of a stimulus to masking.

Second, and more important, comparison of Figure 9 with Figure 4 of Exp. I indicates that for a given T duration masking in the dichoptic mode was obtained with mask exposure durations of less than the minimal duration found to be effective in binocular (and monocular, e.g., Exp. II) conditions. Moreover, the minimal duration of the after-coming PM, which substantially masked T in the present experiment, was not contingent on the duration of T itself. A 3-msec duration of PM was as effective a mask for T = 10 msec as

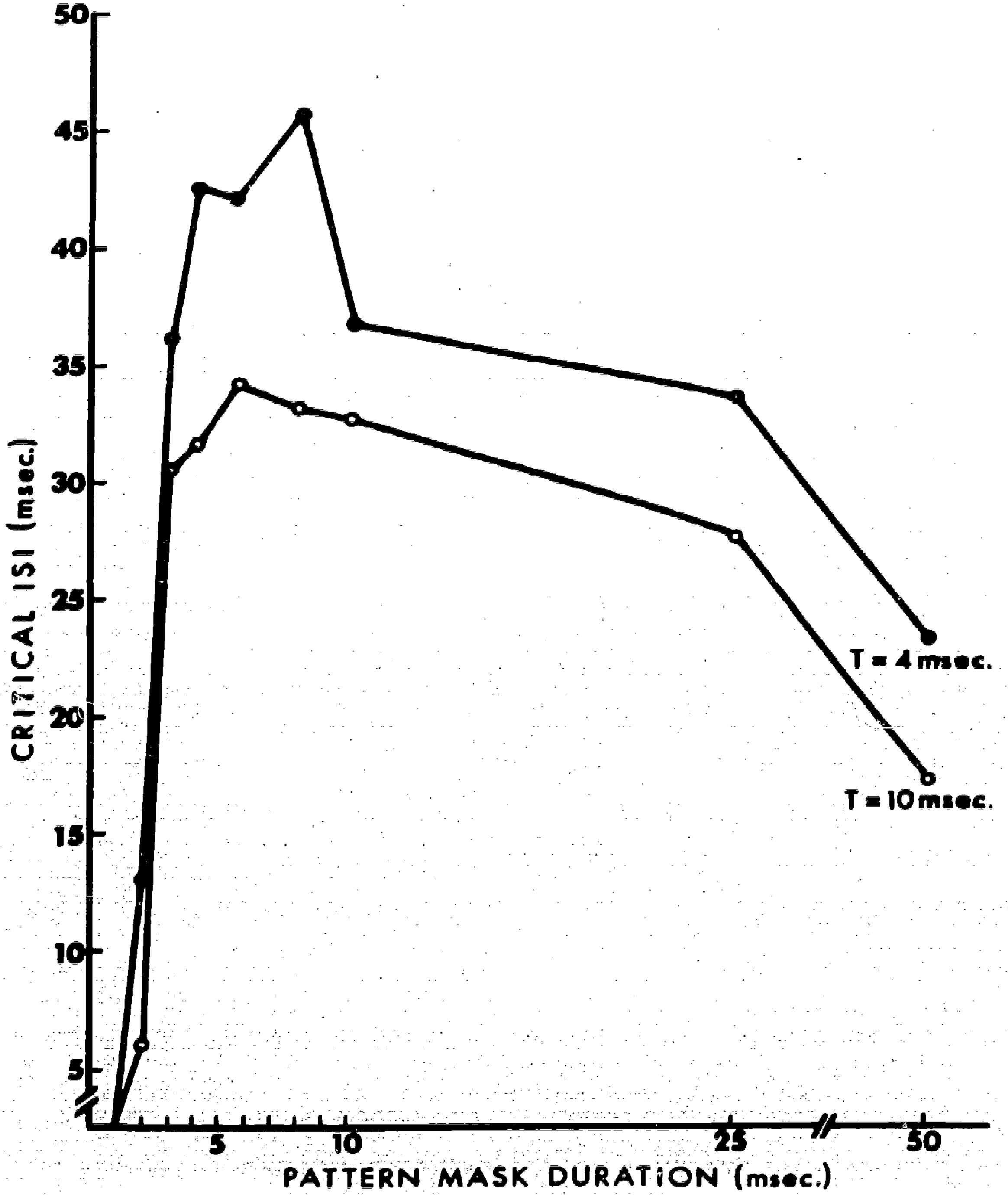


Figure 9: Relation between PM duration and mean ISI_c for dichoptic presentation of stimuli in Exp. V.

it was for T = 4 msec. This is in sharp contrast to the results obtained binocularly and monocularly with RN, where the minimal duration of the mask which impaired the perception of T was a direct function of T duration.

EXPERIMENT VI

Experiment VI was conducted to compare monoptic masking by RN and PM, with dichoptic masking by PM. A conclusion of Exp. V was that mask durations which fail in the monoptic and binocular situations to mask T of a given duration do function effectively under conditions of dichoptic presentation. That conclusion, however, had to be accepted with some reservations since the functions under comparison were obtained with different masks. The monocular and binocular data were obtained with RN as mask; the dichoptic data were obtained with PM.

Method

Four naive Ss participated in the experiment over two days. Two Ss were tested in the dichoptic mode on Day 1 and the monoptic mode on Day 2. The other two Ss received the reverse order. In the monoptic condition for two Ss, masking was examined first with PM and then with RN as the aftercoming stimulus; the other two Ss were tested in the reverse order. The T stimuli were presented to the left eye.

The exposure duration of T was 4 msec for both monoptic and dichoptic conditions. The luminances of T, PM, and RN were each 10 ft L. The set of centrally located symmetrical letters were the T stimuli. For each S in the monoptic conditions, ISI_c was determined by the usual procedure for the following mask durations in the order shown: 1, 2, 4, 6, 10, 50, and 100 msec. In the dichoptic conditions, ISI_c was determined at PM exposure durations of 1, 2, 4, 6, and 10 msec.

Results and Discussion

The results for both monoptic and dichoptic presentation are given in Figure 10. Each data point represents the average ISI_c of the four Ss.

First, inspection of Figure 10 shows that masking in the dichoptic mode can be produced by mask exposure durations which are ineffective in monoptic presentation. In the monoptic condition masking by PM at durations of 1, 2, or 4 msec was practically nonexistent. Figure 10 also shows that masking in the monoptic mode by PM was more severe than masking by RN.

For the present, what is important about Exp. VI is that it adds to the suspicions aroused in Exp. V that dichoptic masking is governed by somewhat different principles than monoptic masking. In dichoptic presentation central devices receive "clean" stimuli, i.e., imputs that are free from the possible confounding effects of the between-stimulus interference introduced when both stimuli have come to the central device by a common peripheral route. The between-stimulus interference which results when the stimuli travel on a common transmission line appears to be due, in part, to the comparative strengths of the stimuli. A 1-msec duration of PM failed to mask monoptically a 4-msec duration of T of the same luminance because T had more

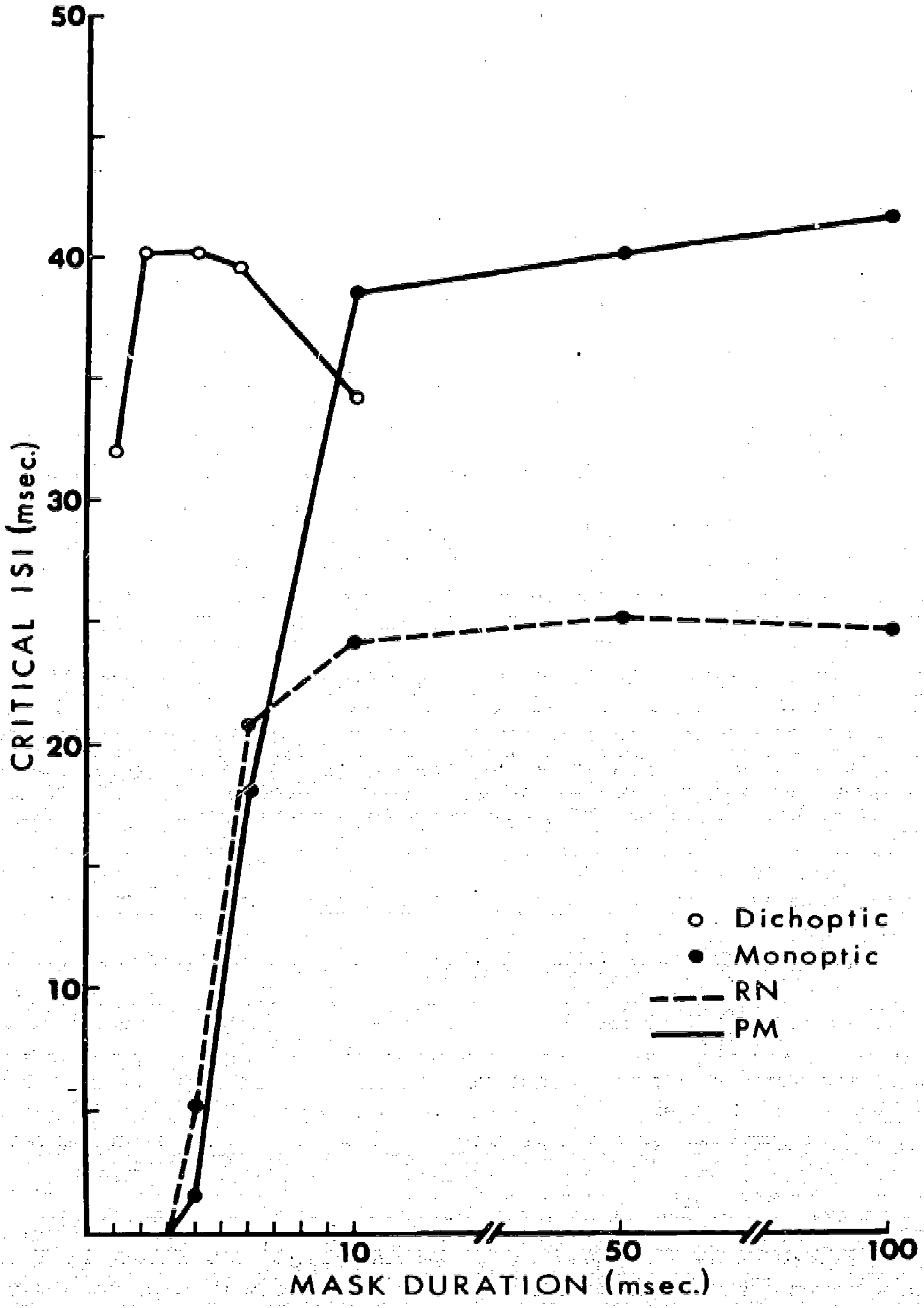


Figure 10: Relation between mask duration and mean ISI_c for monoptic masking by RN, and monoptic and dichoptic masking by PM, in Exp. VI.

energy than PM. Stimulus strength, however, was not a prominent factor in central processes. In dichoptic presentation, a 1-msec duration of PM masked a preceding 4-msec duration of T of the same luminance in spite of the fact that T was the stronger signal.

EXPERIMENT VII

A tentative conclusion drawn from Exp. V was that the relation, T duration x ISI_c = a constant, did not hold true for dichoptic viewing. That experiment seriously questioned the status of T duration as a determining temporal variable in dichoptic masking. Furthermore, in Exps. V and VI mask exposure duration was not as important a variable in dichoptic presentation as it was in monocular or binocular presentation. The relevant temporal parameter in dichoptic masking is suggested in the masking functions of Exp. V (Figure 9). The average separation between ISI_c for T = 4 msec and 10 msec was approximately 7 msec, which is roughly the difference between the two durations. That coincidence implicates SOA as the likely candidate for the role of determining temporal variable in dichoptic masking. Kahneman (1968) and Haber (1969b) advocate SOA as the important temporal variable in masking rather than stimulus duration or ISI. However, the data of Exp. I, and of Kinsbourne and Warrington (1962a), are compelling evidence against the theory that SOA is the only relevant variable. Clearly, the conditions under which SOA, rather than ISI and/or stimulus duration, determines the masking function have to be delineated.

Experiment VII was conducted to examine the hypothesis that SOA was the proper temporal variable for dichoptic masking. The logic of the experiment was simple. If SOA was the relevant variable, then the following relation should hold: critical T duration = T duration + ISI_c = a constant, where critical T duration is the minimal duration of T which permits evasion of masking when ISI = 0 msec.

Method

The procedure of the experiment was as follows. There were two general conditions. In Condition A, T and PM duration were held constant and ISI_c was determined in the usual manner. In Condition B, PM duration was held constant, ISI was 0 msec, and critical T duration was determined. In Condition A there were three T duration-PM duration combinations: T = 8 msec, PM = 2 msec; T = 20 msec, PM = 10 msec; T = 5 msec, PM = 5 msec. For each T - PM combination ISI_c was determined. The PM exposure durations for Conditions A and B were the same. For each PM duration in Condition B, both critical T duration and ISI_c were determined as that value at which S correctly identified four consecutive T letters.

The T stimuli were the set of centrally located symmetrical letters. Luminances of T and PM were equal at 10 ft L. Three Ss participated in the experiment. One S (S1) had had considerable experience with tachistoscopic presentations, the other two Ss were naive. The Ss were tested in a partially counterbalanced design. For a given PM duration each S received Condition A first and then Condition B. Across the three Ss each PM duration appeared once in each test-order position.

Results and Discussion

Table 1 gives the data for the three Ss. A Treatment x Ss analysis of the SOA's revealed that the six treatments did not differ, $\overline{F}(5,10) = 2.44$, $p > .05$, which suggests that under conditions of dichoptic presentation SOA, rather than T duration or PM duration, is the relevant variable.

Inspection of Table 1 shows no profound differences in the estimates of SOA for S whether SOA is computed from Condition A, in which T duration was held constant and ISI_c determined, or from Condition B, in which ISI was set at 0 msec and critical T duration was determined. In any event the picture is obviously quite different from that of Exp. I. The data of the present experiment show a complementarity between T duration and ISI_c implying that dichoptic masking by PM is best described as T duration + ISI_c = a constant.

EXPERIMENT VIII

The thrust of Exps. V, VI, and VII was that dichoptic masking by PM was fundamentally different from monoptic masking by RN. The earlier experiments in the present series showed that stimulus energy is important in determining the interference between stimuli traveling the same transmission channel. There was little, if anything, in Exps. V, VI, and VII which suggested that energy is similarly important in determining the perceptual impairment resulting from two stimuli arriving over separate channels, rather Exp. VII showed that the time elapsing between onsets was crucial for dichoptic presentation.

Experiment VIII compared monoptic masking by PM with dichoptic masking by PM using T and PM intensity as the independent variable.

Method

Four Ss participated in the experiment, two Ss in each of two parts.

Part A. The intensities of T and PM were manipulated in a 2 by 2 factorial design for both dichoptic and monoptic conditions. The two luminances were 5 and 10 ft L. For the two modes of presentation there were, therefore, four T-PM intensity combinations: 5-5, 5-10, 10-5, 10-10. The PM duration was 10 msec. For each T-PM intensity combination critical T duration was determined in the standard fashion of the present series of experiments. The set of centrally located symmetrical letters served as T stimuli and were presented to the left eye.

Part B. Intensities of T and mask were manipulated in the manner of Part A. The two intensities in this instance were 1.25 ft L and 2.5 ft L. The stimuli were the consonant trigrams, and the mask was the line-configuration of PM reproduced in triplicate, once at each of the locations of the consonant letters in the T displays. For future reference this mask will be referred to as PM3. The exposure duration of PM3 was 10 msec. Critical T duration was determined for both Ss in both presentation modes at each of the intensity combinations. The $\overline{T}$ stimuli were presented to the right eye.

TABLE 1

EXP. VII: CRITICAL ISI, CRITICAL T DURATION AND SOA
FOR DICHOPTIC PRESENTATION

Subjects	Critical ISI and Estimated SOA (= T + ISI_c)						Critical T Duration and SOA (= critical T duration)		
	T = 8 msec PM = 2 msec	SOA	T = 20 msec PM = 10 msec	SOA	T = 5 msec PM = 5 msec	SOA	PM = 2 msec	PM = 10 msec	PM = 5 msec
<u>S</u>1	31	39	18	38	35	40	40	42	43
<u>S</u>2	59	67	50	70	65	70	78	70	73
<u>S</u>3	50	58	48	68	54	59	59	64	64

Results and Discussion

The results of Part A are given in Table 2 and the results of Part B in Table 3. Inspection of the two tables reveals a monoptic-dichoptic difference which corroborates an earlier observation reported by Schiller (1969): stimulus luminance has a pronounced effect on masking in the monoptic mode but little, if any, systematic effect on masking in the dichoptic mode. Furthermore, it is evident on inspection that masking dichoptically was more severe than masking monoptically.

TABLE 2

EXP. VIII: MEAN CRITICAL T DURATIONS FOR PART A

		Monoptic		Dichoptic	
		PM intensity		PM intensity	
		5 ft L	10 ft L	5 ft L	10 ft L
T intensity	5 ft L	24.5	46.5	56.0	40.5
	10 ft L	9.5	33.5	67.0	67.0

TABLE 3

EXP. VIII: MEAN CRITICAL T DURATIONS FOR PART B

		Monoptic		Dichoptic	
		PM3 intensity		PM3 intensity	
		1.25 ft L	2.5 ft L	1.25 ft L	2.5 ft L
T intensity	1.25 ft L	19.0	40.5	73.5	71.0
	2.5 ft L	7.0	15.0	81.0	75.5

The dichoptic data of Table 2 are a little untidy, which may be attributed in part to the phenomenon reported in Exp. V. At longer durations and at greater intensities, there is a loss in clarity of stimuli which are of the dark figure--light background type. This phenomenal decrease in form clarity may induce disparities in critical T duration estimates across varying luminance conditions.

A further point needs to be added. In pilot work it was observed that with dichoptic presentation the severity of masking was not equivalent for the two eyes. Cursory examination indicated that the severity of dichoptic masking was very much a matter of which eye received the T stimulus and which eye received the mask. This suggests that factors related to ocular dominance and binocular rivalry are probably involved in the dichoptic paradigm.

EXPERIMENT IX

Stimulus intensity systematically affects critical T duration in monoptic presentation but not in dichoptic. That was the outcome of Exp. VIII. Experiment IX examined the effect of mask duration on critical T duration under conditions of monoptic and dichoptic presentation. On the basis of Exp. VIII, and for that matter of Exp. VI, it was predicted that in the monoptic case, critical T duration would be directly proportional to mask exposure duration, but dichoptically, critical T duration would be unaffected by mask exposure duration.

Method

Three Ss participated in the experiment. One S, S1, was not naive to tachistoscopic viewing. The T stimuli were the consonant trigrams and the mask was PM3. The luminances of T and PM3 were both 2.5 ft L. Three exposure durations for PM3 were used for both monoptic and dichoptic presentation: 4, 10, and 20 msec. The trigram stimuli were presented to the left eye. The mask stimulus followed on the left eye for monoptic and on the right eye for dichoptic. Each S was tested monoptically first.

Results and Discussion

The data for each of the three Ss are given in Table 4. The data are unequivocal. Increasing exposure duration of PM3 increased the minimal duration of T necessary to escape the masking of PM3 monoptically but not dichoptically. Common to Exp. VIII and the present experiment is the fact that critical T duration was considerably larger for dichoptic presentation than for monoptic presentation. This is contrary to a frequently quoted generalization that dichoptic presentation produces less interference than monoptic (e.g., Kolers, 1968, p. 39).

GENERAL DISCUSSION OF EXPERIMENTS I - IX

"Peripheral" and "Central" Defined

The terms "peripheral" (or transmission line) and "central" as used in the present communication have served as convenient ways of talking about

TABLE 4

EXP. IX: CRITICAL T DURATION AS A FUNCTION OF PM3 DURATION FOR MONOPTIC AND DICHOPTIC PRESENTATION

PM3 Duration (msec)	SUBJECTS					
	S1		S2		S3	
	Monoptic	Dichoptic	Monoptic	Dichoptic	Monoptic	Dichoptic
4	6	50	10	152	6	90
10	20	50	34	156	14	95
20	38	48	38	161	28	91

the loci of particular effects. They are, however, loaded terms because they imply two distinct and separable anatomic regions. In reality the interface between the sensory pathways and cortical structures is not at all a sharp boundary but rather a gradual merger. In addition, the term "transmission line" connotes a passive conduit via which exact images of physical stimuli are conveyed from the peripheral receptor to the brain. To the contrary, the electrophysiological evidence available thus far (see Chung, 1968) indicates that visual information is subject to drastic recoding as it proceeds along the pathways of the nervous system, with the degree of recoding and modification increasing as the input proceeds further centrally. In other words, en route to the cortex operations occur which give rise in output to something other than a mere relaying of the input array.

The definition of "peripheral" that has been implicit in the preceding discussions is one which includes retina, lateral geniculate nucleus, and striate cortex as its components. Preference is given to a view of the transmission line as a collection of devices signalling properties of the stimulus, and on this view the interface between peripheral and central is intentionally vague. Some cells of the striate cortex are seen as terminals of peripheral systems extracting basic stimulus parameters, while others are seen as enlisted in central processes that derive an identification of the stimulus from the data set so provided. A recent, relevant discussion of the functional organization of the striate cortex with respect to form perception is that of Pollen, Lee, and Taylor (1971).

The earliest point in the nervous system at which dichoptic masking may originate is probably in the region of the peripheral-central interface, although the question of whether the two eyes interact earlier, at the level of the lateral geniculate nucleus, has not gone unheeded. For the cat visual system, at least, there is some reason to believe that the two eyes might interact at the geniculate. Dichoptic interactions have been observed by Fillenz (1961) and by Lindsley, Chow, and Gollender (1967), and Bishop and his coworkers (Bishop, Burke, and Davis, 1959) have reported activation of geniculate cells by stimulation of either optic nerve. However, against this evidence is the work of Hubel and Wiesel (1961) and Sturr and Battersby (1966) which implies that interactions at the level of the geniculate are minimal at best. Furthermore, it has been reported (Jung, 1961) that at the level of the primary visual projection in the cortex, true binocular convergence is comparatively rare and most cells respond only to afferents from the ipsilateral or the contralateral retina. The implication of this is that dichoptic masking may arise at a relatively late stage in the cortical processing of visual data.

Two Loci for Backward Masking

In short, there are two possible loci for the perceptual impairment resulting when two visual stimuli follow in rapid succession. The impairment may have its locus in the transmission channel or in a central processor. Impairment localized in the transmission channel is best viewed as the effect one stimulus exerts on the other. Impairment localized in a central processor can be of two sorts: an interaction between the stimuli, similar in kind to that occurring in the transmission channel, or a distortion induced in the operation of a central processing mechanism (see Kolers, 1968). The proposition that backward masking reflects a disturbance in the proper functioning of a central device is to emphasize that the masking is not due to the effects exerted by stimuli on each other. With reference to the clerk-customer analogy, the second customer does not have a direct effect on the fate of the first; rather, he exerts an indirect effect by causing the clerk to be hurried and less thorough in his treatment of the first.

Masking by RN under conditions of monocular and binocular presentation was an instance of interference in the transmission channel. That the effect RN exerted upon T did not have a central locus was revealed by the absence of any masking by RN under conditions of dichoptic presentation. Between-stimulus interference arising in the transmission channel was defined by the following relation: T energy x ISI_c = a constant. Masking by PM, on the other hand, could have a central locus. The data suggested that the relevant independent variable for masking by PM under conditions of dichoptic presentation was SOA. Moreover, the nature of masking by PM in the dichoptic mode was not affected in any serious fashion by energy properties of the stimuli. Those two observations, the relevance of SOA and the comparative irrelevance of energy variables, favor the interpretation that dichoptic masking by PM represents an interruption in the normal functioning of a central mechanism rather than the effects of serial stimuli upon each other.

The Peripheral Operation

What does the relation, T energy x ISI_c = a constant, tell us about the peripheral visual system? The answer seems to be this: whatever the operations performed by the peripheral visual system on an incoming stimulus, the rate at which those operations are conducted is directly related to the energy of the stimulus. To reiterate some essential points. The RN mask exerts an influence on a preceding T stimulus only if RN is input on the same retinal area and, therefore, on the same transmission line as T. It must be assumed that masking by RN at some ISI means that the peripheral processing of the T stimulus has not been completed by the time RN occurs. The minimal time between T and RN at which T evades the masking action of the after-coming event, i.e., ISI_c, is inversely related to T stimulus energy. Therefore, suffice it to say that peripheral processing time is inversely proportional to the energy of the stimulus. We can infer from the foregoing that peripheral processing may be completed within the duration of a stimulus, given the right order of stimulus intensity. Support for that conclusion is found in the experiments of Kinsbourne and Warrington (1962a, 1962b).

The Role of Mask Energy in Peripheral Masking

The energy of the mask (RN) in monoptic or binocular presentation had to be equal to or greater than that of T in order to impair the identification of T. But once RN energy was just slightly greater than T energy, as inspection of Figure 4 clearly shows, further increases did not extend the ISI over which masking could be obtained. A useful general conclusion follows from this fact. When backward masking does occur in monoptic and binocular conditions where the energy of the mask is less than that of the T stimulus, it is unlikely that the masking originates peripherally. Rather, we ought to conclude that the masking is of central origin. This conclusion may only apply to masking of form where the S's task is to identify the form, i.e., the masking of interest to the present communication.

To clarify the potential importance of this conclusion, consider two instances of masking: masking by a contourless light flash and dichoptic masking by pattern. When the mask is a homogeneous flash of light of energy less than or the same as that of the T stimulus, masking is generally not obtained (Schiller, 1969). It is also known that masking of a form by a contourless light flash of greater energy does not have any appreciable central component (e.g., Schiller, 1965). Therefore, we may conclude that if the flash energy is not greater than T energy and if the two stimuli are not on the same transmission line, then masking of a form by a flash of light cannot occur. In contrast, when the mask is a pattern (say, PM) and the stimulus presentation is dichoptic, masking does occur, and the condition that mask energy be greater than T energy is not a necessary condition for such masking. Presumably, therefore, monoptic or binocular masking by pattern could occur centrally rather than peripherally, and that means, of course, that monoptic or binocular masking could occur in conditions where the energy of the pattern mask is less than that of the T stimulus.

The Central Operation

An important distinction between peripheral and central processes was demonstrated in Exps. VI, VIII, and IX. Whereas the parameters of duration and intensity significantly affected masking of peripheral origin, their effect centrally was negligible. This distinction is put into relief by electrophysiological data which show that the further centrally a neuron lies, the more complex and specific become the stimulus parameters to which the cell responds. Thus, the more centrally a cell is located, the more likely it is that the cell will be affected by informational rather than energy characteristics of stimuli.

The relevance of SOA to masking of central origin suggests that the constraint on central processes is simply time elapsed since stimulus onset. We will presume, and not without reason, that the central machinery assumes the major burden of pattern recognition and that it uses as its raw material the visual data provided by the peripheral mechanisms.

We may assume for the present that the relation between the peripheral and central processes is that they are successive and additive. That is to say, the peripheral operation must be complete before the central operation can begin, and therefore, the time needed to identify a tachistoscopically presented letter would be the total time of the two operations combined. It will be part of the task of the experiments that follow to assess the validity of this hypothesis.

Backward Masking by PM

There is now the question of the nature of masking by PM under conditions of monocular or binocular presentation. As noted in the introduction to the present paper, it is not inconceivable that the masking effect of a particular stimulus could be exerted prior to the establishment of the T representation or subsequent to the establishment of the T representation. Therefore, when T and PM are transmitted on the same channel, the resulting perceptual interference could reflect effects at either, or both, loci. However, the impression gained from Exps. I - IX was that interference in the transmission channel and interference with the operation of a central processor were two very distinct phenomena such that any masking that might be observed reflected either one or the other, but not both.

Consider T and PM presented monocularly. If T and PM fuse in the transmission channel, as suggested by the integration hypothesis, the task of the central processor would be rather like that of trying to make sense of a photograph produced by a double exposure. What is important here is the fact that if the stimuli superimpose in the peripheral channel, then what the central device receives effectively is but one stimulus for analysis, not two. If, however, the two stimuli do not interact in the transmission channel, for whatever reason, then the central processor receives two stimuli in succession and the task now is that of trying to make sense of the first before the second arrives.

In the experiments that follow, an attempt is made to separate the peripheral and central components in monoptic masking by PM (PM3). More generally, the experiments are directed at the question of how the two operations, peripheral and central, relate.

EXPERIMENT X

The essence of the concluding comments on Exps. I - IX was that at some point a masking function for monoptic presentation, of the sort generated by the procedure of Exp. IX, must assume the characteristics of dichoptic presentation. What Exp. IX (and for that matter Exps. VI and VII) had shown was that central processes were unaffected by the exposure duration of a lagging mask stimulus. Presumably, therefore, in monoptic presentation a measure of masking, such as critical T duration, should asymptote at some value of PM duration. Such an outcome, perhaps, would be expected regardless of any theory. Yet, pilot work prompted this experiment in that Ss reported an interesting shift in their phenomenological description of the stimuli as the duration of PM (actually PM3), and accordingly, the duration of T, increased. At brief durations pilot Ss reported a relatively unclear, degraded stimulus. The experience was that of T and PM "mixed." At longer durations Ss reported seeing a "clear" T followed by a "clear" PM, the experience being that of "not having sufficient time to read T." The latter description had been used occasionally by Ss in the previous dichoptic conditions.

Method

Six Ss participated in the experiment. Two of the Ss were highly experienced in tachistoscopic experiments; they were members of the staff of Haskins Laboratories and had served as pilot Ss for a number of the preceding experiments. The remaining four Ss were naive to the apparatus and to the experiment.

For each S, critical T durations were determined in the given order for the following values of PM3 duration: 2, 3, 4, 5, 6, 8, 10, 25, 35, 50, 100, and 500 msec. Presentation was monoptic at the right eye. The T stimuli were the set of trigrams and the luminance of T and PM3 was 2.5 ft L. Following each stimulus presentation and report, S was required to describe his experience of the stimuli. The Ss were not told what to expect.

Results and Discussion

The averaged data are represented graphically in Figure 11. The individual S data are given in Table 5. Inspection of Figure 11 suggests a linear relation between critical T duration and PM3 duration up to PM3 duration = 10 msec, followed by what appears to be a relatively abrupt transition to asymptote.

Individual S data shown in Table 5, S2 and S6 for example, demonstrate this transition most vividly. In the region of this transition, Ss shifted in their description of what they were seeing as the T duration approached the critical value. Up to the transition region Ss described the T stimulus as "messy," "mixed up," "hard to make out," and "unclear." Subsequent to the transition region Ss gave the following descriptions: "pattern replaced letters"; "image of letters shortened by pattern";

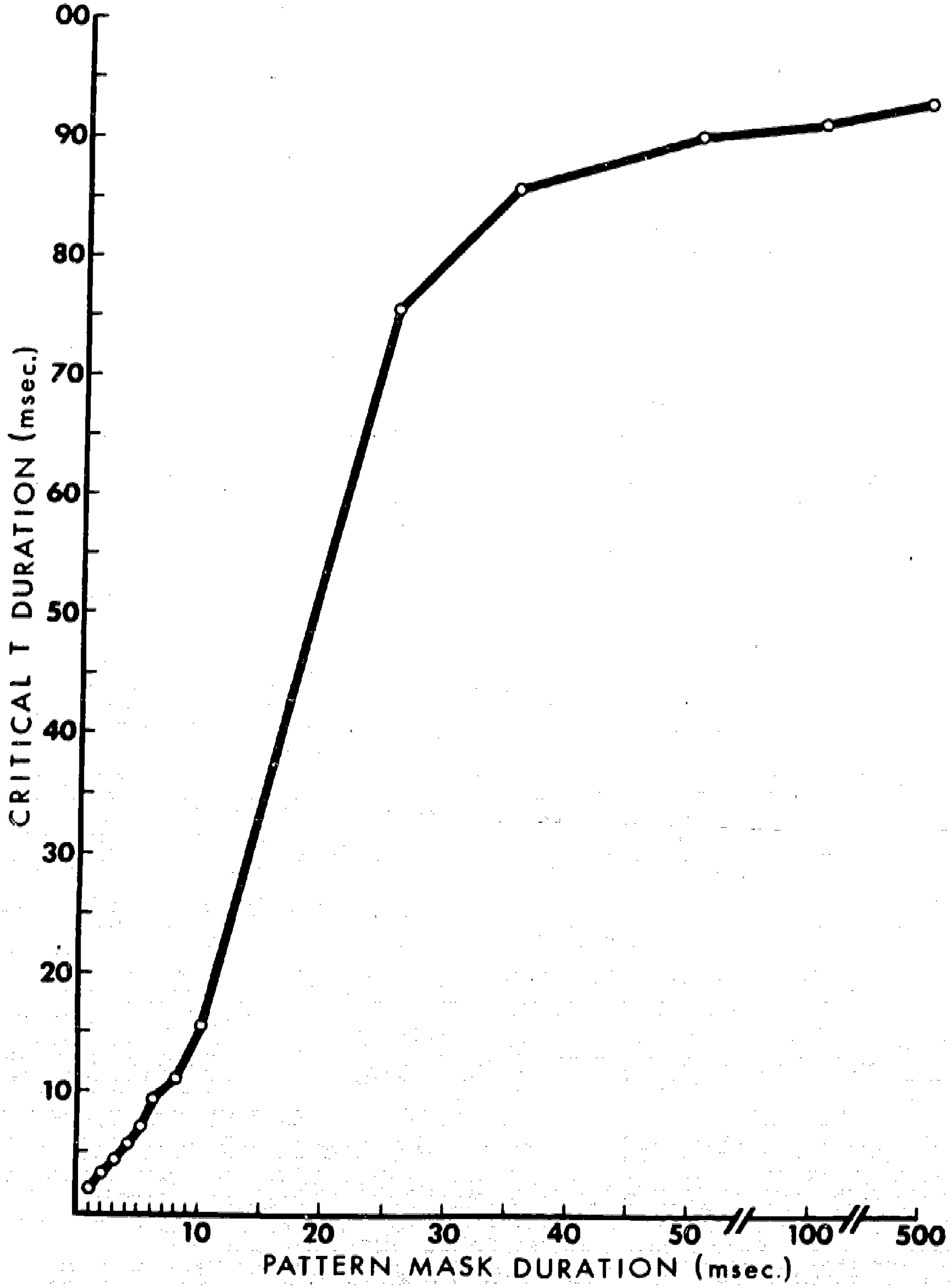


Figure 11: Relation between PM3 duration and mean critical T duration for monoptic stimulus presentation in Exp. X.

TABLE 5

EXP. X: CRITICAL T DURATION AS A FUNCTION OF PM3 DURATION FOR MONOPTIC PRESENTATION

PM3 Duration (msec)	Subjects					
	S1	S2	S3	S4	S5	S6
1	1.5	1.5	1.5	2.5	2.5	2.5
2	3.0	3.5	2.5	3.5	3.5	3.5
3	4.0	4.5	3.0	5.5	5.0	5.0
4	5.0	6.0	4.0	6.0	6.0	6.0
5	6.0	7.0	7.0	8.0	7.5	8.0
6	8.0	8.5	8.0	12.5	10.0	10.0
8	9.0	10.5	10.5	14.5	11.5	10.5
10	12.5	15.0	13.5	21.0	20.0	12.0
25	21.5	130.0	25.0	84.0	42.5	150.0
35	28.0	145.0	26.0	110.0	48.0	157.0
50	37.0	145.0	28.5	112.0	62.0	156.0
100	34.0	150.0	31.5	113.0	63.0	157.0
500	37.0	150.0	31.5	113.0	66.0	162.0

"pattern stopped me reading the letters." The principle phenomenological difference between the phase prior to and that subsequent to the transition region was that Ss described a shift from seeing one event to seeing two events in succession.

Further indication that masking in the PM3 duration range 1 msec to 10 msec was fundamentally different from that observed in the PM3 duration range 25 msec to 500 msec was provided by S errors and by between-S differences. A coarse examination revealed a fairly consistent pattern. For the mask range 1 msec to 10 msec, errors seemed to be evenly distributed across position; for example, Ss tended to commit as many errors in reporting the first letter of the trigram as they did in reporting the third letter. Moreover, all three letters became available at very much the same critical T duration (cf., Kinsbourne and Warrington, 1962a). In contrast, errors committed in the mask range 25 msec tended to relate to position in the trigram array. As T duration increased S was more likely to report the first letter correctly, less likely to report the second letter, and least likely to report the third. Omitting the third item of a trigram was common in PM3 range 25 to 500 msec, the Ss frequently responding that they did not have time to read it.

Between-S comparisons were also illuminating. Two Ss, S1 and S3, as noted above, were highly experienced in the task of reading material from a masked display. Inspection of Table 5 reveals a considerable difference between the performance of Ss 1 and 3 and the remaining Ss across the PM3 durations of 25 to 500 msec, yet little, if any, difference in the range 1 to 8 msec. Admittedly the possibility of large differences in critical T duration across PM3 exposure durations 1 to 8 msec was limited; this, however, does not detract from the fact that the increase in mask duration from 10 msec to 25 msec resulted in a clean separation of the sophisticated from the naive Ss. Moreover, errors committed by Ss 1 and 3 in the asymptotic part of the function were more evenly distributed across the trigram-letter positions.

EXPERIMENT XI

Experiment X reinforced the impression that two quite different processes could be isolated in monoptic masking by PM3. It was inferred that at briefer durations of T and PM3, the masking was similar to masking by RN, and at the longer durations, the perceptual interference was more like that seen dichoptically. The correctness of this inference could be tested on the basis of the data of Exp. VIII: manipulating luminance should affect the initial rising part, if that mirrored peripheral masking, but not the subsequent asymptotic part of the function relating critical T duration to PM3 duration. Experiment XI was designed to perform this test.

Method

The procedure of Exp. XI was similar to that of Exp. X. Critical T duration was estimated at the following durations of PM3: 1, 2, 3, 4, 5, 6, 8, 10, 15, 20, 25, 35, 50, and 100 msec. A single estimate was made for each of six naive Ss at each PM3 duration going in order from the shortest (1 msec) to the longest exposure duration (100 msec). The T stimuli were, as before, the set of trigrams. The principal feature of Exp. XI was that across the exposure durations of the after-coming stimulus, critical T duration was determined for three T stimulus-mask stimulus luminance ratios. The three ratios were: 1:1 (T = 2.5 ft L, PM3 = 2.5 ft L); 2:1 (T = 5.0 ft L, PM3 = 2.5 ft L); 1:2 (T = 2.5 ft L, PM3 = 5.0 ft L). In a partially counterbalanced design each ratio condition appeared twice across the six Ss at each of the three possible test-order positions. The stimuli were viewed with the right eye.

Results and Discussion

The average critical T durations for each of the three intensity ratios are shown in Figure 12.

The family of curves are reproduced in a log-log plot in order to give a clearer picture of the initial rising component of the functions. The hypothesis under test was that the ascending component of the function relating critical T duration to PM duration would be affected by the ratio of stimulus intensities but the asymptotic component would not, the idea being that the ascending and asymptotic phases reflected masking of two different origins. Inspection of Figure 12 shows, in accord with this hypothesis, that the ascending components of the three curves differed, while the asymptotic components did not.

EXPERIMENT XII

The data of Exps. X and XI invite the following hypothesis: the duration (energy) of the T stimulus determines whether the upper limit on monoptic masking by PM (or PM3) will reflect peripheral or central processes. Inspection of Figure 12 suggests that for all three ratios, the masking, for example, of a 3-msec exposure of T by a 5-msec exposure of PM3 was localized in the transmission line; on the other hand, the masking of a 60-msec T by a 50-msec PM3 was central in origin. The data of Exp. XI further imply that the origin of interference with a 3-msec T by a 50-msec and a 5-msec exposure of PM3 should be one and the same. That is, in both of these instances in which mask energy is greater than T energy and the T energy is comparatively weak, the locus of masking should be peripheral. The upshot of all this is that the locus of the interference induced by a PM3 of 50-msec exposure should shift from peripheral to central as T exposure duration increases.

The design of Exp. XII involved estimating ISI_c for eleven values of T duration, ranging from 2 to 64 msec, with mask duration held constant at 50 msec. Assuming the validity of the above reasoning, it was expected that at brief values of T, masking would display characteristics of between-stimuli interference in the transmission line; at longer durations of T, the masking would fit the central mold, i.e., complementarity would be observed between T duration and ISI_c. To provide a yardstick for interference in the transmission channel, ISI_c was determined across the eleven T durations, with RN as the after-coming mask. At the briefer durations of T, the function relating T duration to ISI_c with PM3 lagging should look similar to that with RN lagging. However, at the longer exposures of T, the two functions should assume very different characteristics.

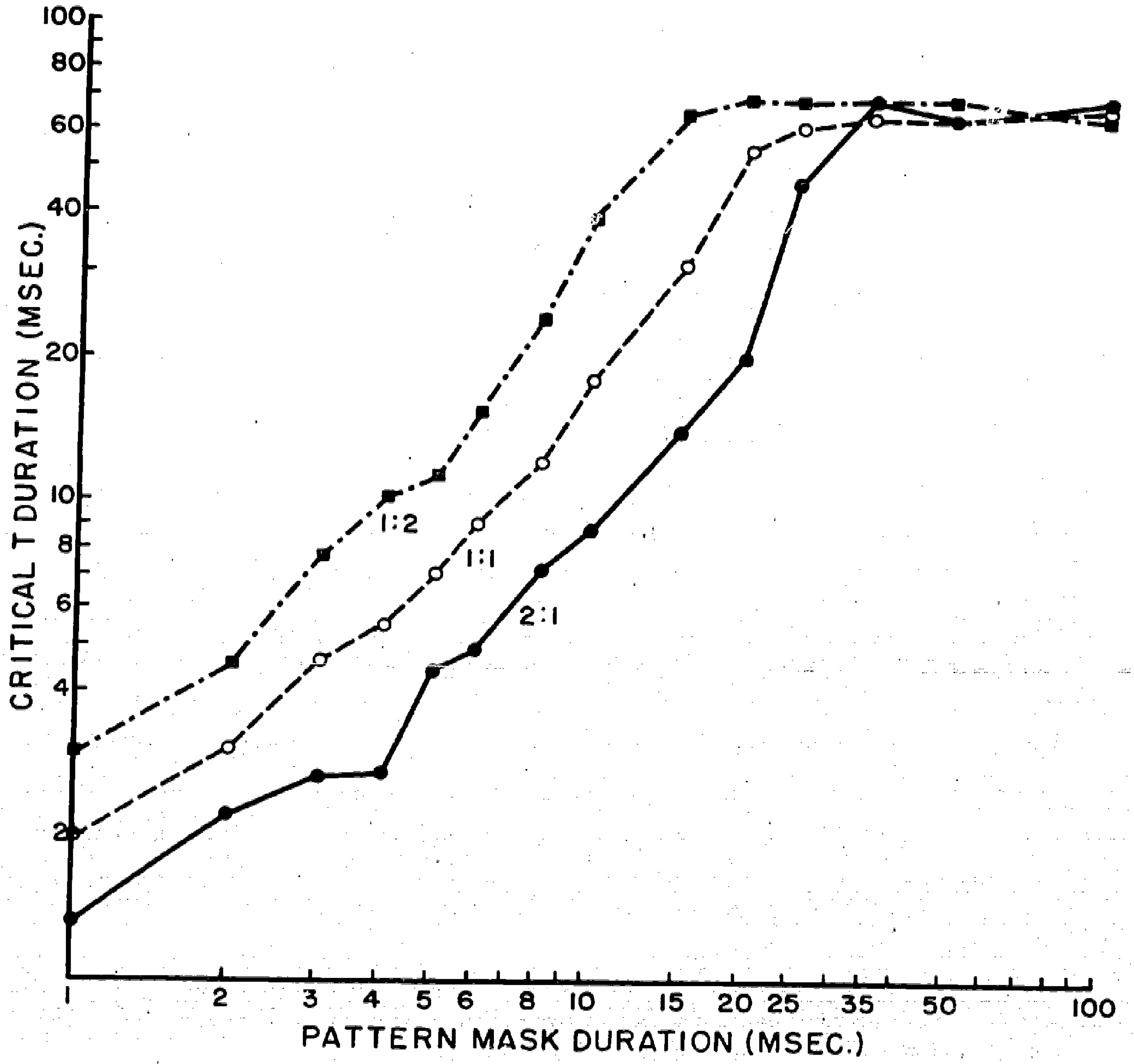


Figure 12: Log-log relation between PM3 duration and mean critical T duration for monoptic stimulus presentation at three intensity ratios in Exp. XI.

Method

The T stimuli were the trigrams. Six naive Ss received all conditions in a partially counterbalanced arrangement. Critical ISIs were estimated at each T duration going in succession from 2 to 64 msec. At each T duration, ISI_c was determined for both masks before testing at the next duration. Three of the six Ss were given RN first and the remaining three were given PM first. The ISI_c was determined in the usual manner. The luminances of T, PM3, and RN were equal at 5 ft L, and stimuli were viewed with the right eye.

Results and Discussion

The data averaged across the six Ss for PM3 and RN as the masking stimuli are plotted in Figure 13. Individual data are given in Table 6.

As before, masking by RN produced a simple relation between the exposure duration of the T event and the minimal time required to evade masking, i.e., T duration x ISI_c = a constant. The present RN function if plotted on log-log paper is virtually a straight line, and compares favorably with a log-log plotting of the RN function by Kinsbourne and Warrington (1962a). The only serious departure from the multiplicative relation occurs at T = 16 msec. At that exposure duration some Ss were still masked by the lagging RN, as inspection of Table 6 shows, but the multiplicative rule was obviously not in effect.

Masking by PM3, in sharp contrast to masking by RN, yielded a complex relation between T duration and ISI_c. At the very brief durations of 2 and 3 msec (and perhaps 4 msec), the PM3 function paralleled the RN function, i.e., the relation between exposure duration and ISI_c appeared to be multiplicative. For T = 2 msec and T = 3 msec the ISI_c's were 90 msec and 62 msec, respectively. Multiplying duration by ISI_c in these two cases yields very much the same values, 180 in the 2-msec case and 186 in the 3-msec case. In contrast, taking the next four values of T--4, 6, 8, and 16 msec--and multiplying them by their appropriate ISI_c's yields unequal products of approximately 208, 288, 384, and 1280, respectively. Thus, at the longer exposures of T, the PM3 function does not fit the multiplicative rule, and the relation between T duration and ISI_c is best described as T duration + ISI_c = a constant.

The conclusion of Exps. I - IX was that the multiplicative rule characterized peripheral and the additive rule characterized central processes. Indeed, the additive relation between T duration and ISI_c had been detected in dichoptic presentation. Thus, the present experiment may be viewed as a demonstration that peripheral and central masking are isolable and separable in conditions of monoptic (or binocular) presentation of stimuli. In Figure 13 the multiplicative and additive relations are referred to as Stage I and Stage II, respectively.

It is evident from Figure 13 and Table 6 that ISI_c at the very brief T durations was much greater in the PM3 function than in the RN function. The implication might be that for brief exposures of T both peripheral contamination and central distortion summate when PM3 (or PM) is the masking event. The position outlined in the general discussion of Exps. I - IX was

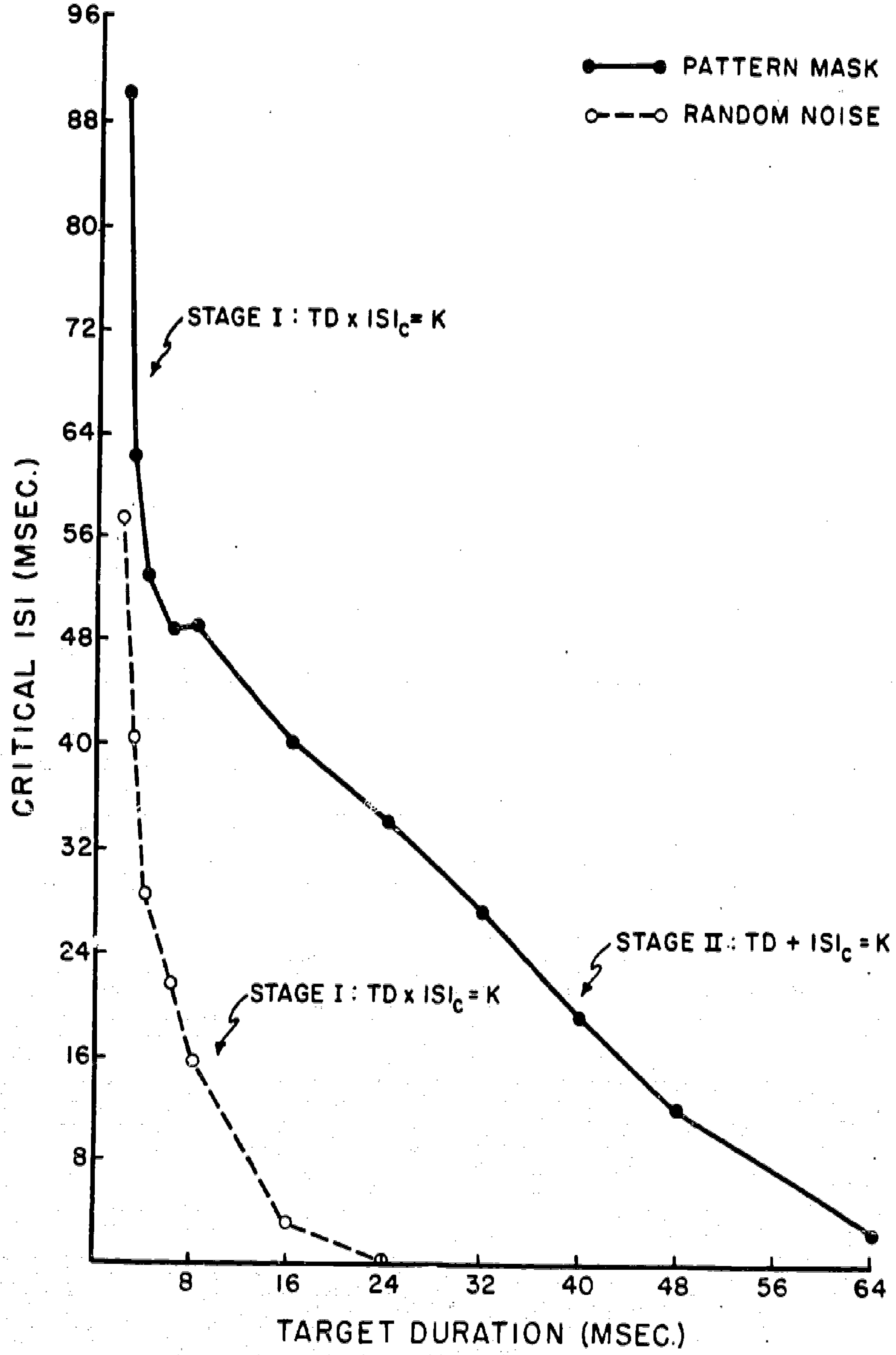


Figure 13: Relation between T duration and mean ISI_c for monoptic masking by RN and PM3 in Exp. XII.

TABLE 6

EXP. XII: MONOPTIC CRITICAL ISI AS A FUNCTION OF T DURATION WITH RN AND PM3 AS MASKS

T DURATION (MSEC)	RN							PM3						
	SUBJECTS						MEAN ISI_c	SUBJECTS						MEAN ISI_c
	S1	S2	S3	S4	S5	S6		S1	S2	S3	S4	S5	S6	
2	58	63	45	63	46	70	57.50	96	89	72	73	75	135	90.00
3	40	45	32	48	35	45	40.83	66	61	51	50	66	80	62.33
4	29	32	23	35	26	32	29.50	51	40	44	38	68	79	53.33
6	21	22	15	25	19	24	21.00	44	46	41	35	49	78	48.83
8	15	18	12	16	15	18	15.67	67	37	40	33	49	68	49.00
16	2	8	1	3	3	3	3.33	59	33	20	24	52	54	40.33
24	0	0	0	0	1	1	.33	53	25	23	18	40	45	34.00
32	0	0	0	0	0	0	0	54	13	12	6	38	39	27.00
40	0	0	0	0	0	0	0	44	7	5	2	27	29	19.00
48	0	0	0	0	0	0	0	37	1	0	0	14	25	12.83
64	0	0	0	0	0	0	0	12	0	0	0	0	3	2.50

that masking was due either to between-stimulus interference in the transmission channel or to distortion in the operation of a central device, but not to both. Preferably, therefore, the reason for the greater severity of masking by PM3 than by RN at brief T exposures should be sought elsewhere than in the notion that peripheral and central effects combine.

The more severe impairment in the perception of T by PM3 may be attributed to a greater confusion of contours owing to the greater similarity between T and PM3 than between T and RN. Admittedly the argument that PM3 is more like T than RN is like T, is based on the dichoptic effect of PM3 and the absence of such an effect with RN. However, there does seem to be some truth to the hypothesis.

Figure 14 shows a comparison between monoptic masking by RN and by a homogeneous flash of light of the same intensity. The data are from two Ss; the duration of T was 4 msec and its intensity was the same as that of the two masks. The experiment was conducted in the manner of Exp. I. Perception of T was more impaired by RN than by a contourless flash of light of the same intensity. It is known that masking of a form by a light flash does not take place dichoptically (e.g., Mowbray and Durr, 1964; Schiller and Wiener, 1963); its effect, like that of RN, is restricted to the transmission channel. Comparison of Figure 14 with Figure 10 of Exp. VI and Figure 13 of the present experiment suggests that masking in the transmission line owes allegiance to variables other than energy variables. Evidently the similarity between T and the mask is a determinant of the degree of between-stimulus interference in the transmission channel. However, for any given T and mask the degree and direction of interference in the transmission line varies as a function of their respective energies.

EXPERIMENT XIII

Experiment XII had isolated peripheral and central masking effects in monoptic viewing which prompted the question: Is the central effect in monoptic presentation the same as that in the "clean signals" case of dichoptic presentation? That is, would the minimal SOA for criterion performance be the same regardless of whether the two stimuli traveled to the central processor by the same route or by separate routes? Experiment XIII answers this question.

Method

At four values of T duration, 10, 20, 30, and 40 msec, four naive Ss were examined both monoptically and dichoptically. For each T duration in each mode, two estimates of ISI_c were made. Two Ss were given the following order of conditions: dichoptic, monoptic, rest (approximately 10 min), monoptic, dichoptic. The other two Ss were given: monoptic, dichoptic, rest, dichoptic, monoptic. The T stimuli were the consonant trigrams and PM3 was the mask held constant at 50-msec duration and at 2.5 ft L, the intensity of the T stimuli. Presentation was to the right eye.

Results and Discussion

The average of the two estimates of ISI_c made at each T duration for each S for both modes of presentation are given in Table 7. The data are

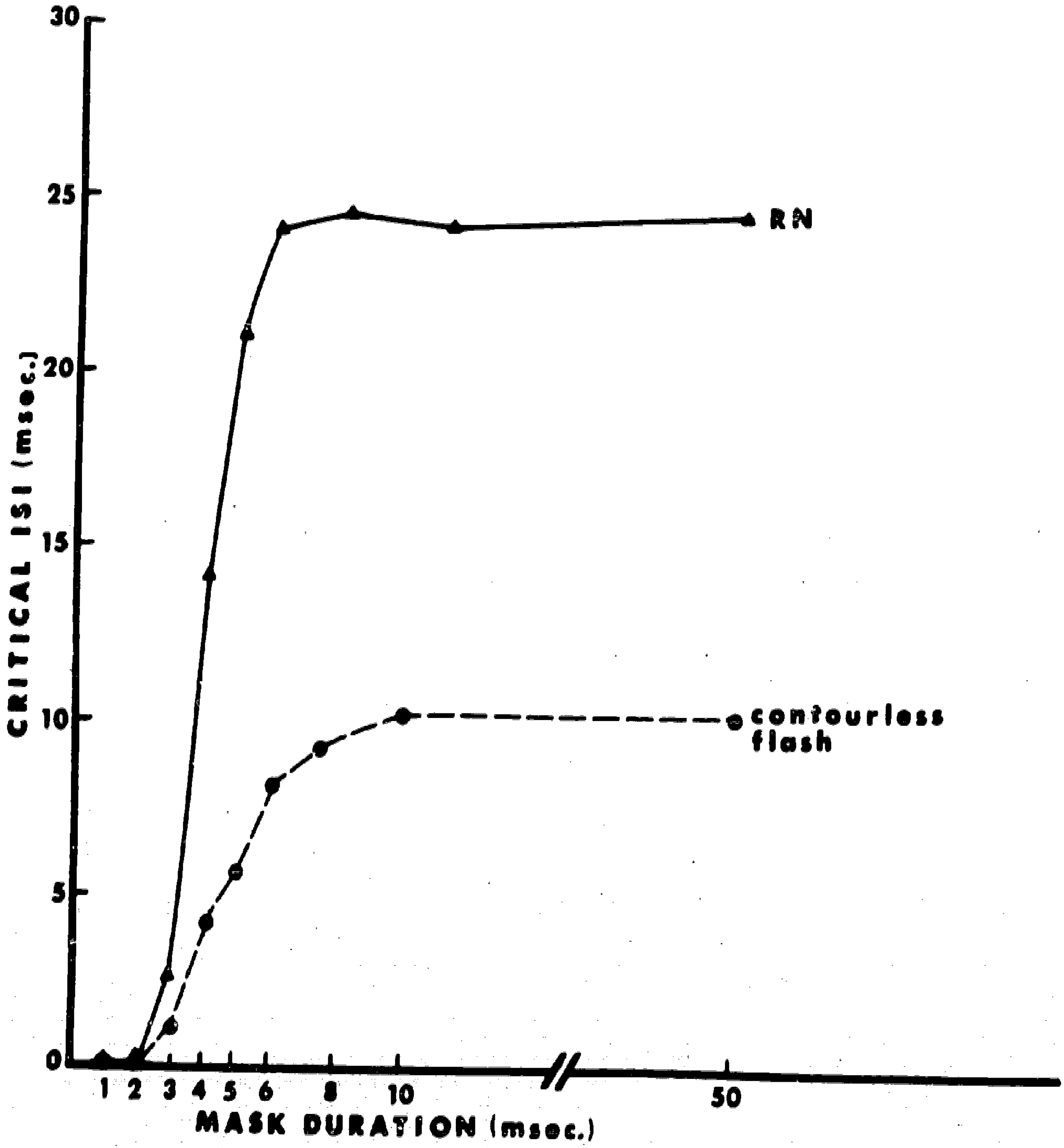


Figure 14: Relation between mask duration and mean ISI_c for monoptic masking by RN and a contourless light flash of the same intensity.

TABLE 7

EXP. XIII: COMPARISON OF MONOPTIC AND DICHOPTIC SOA: MEAN ISI_c AS A FUNCTION OF T DURATION

Subjects	Monoptic				Dichoptic			
	T Duration				T Duration			
	10	20	30	40	10	20	30	40
*S*1	99.0	86.0	68.5	63.0	87.0	74.5	69.5	66.5
*S*2	105.0	87.5	83.0	86.5	96.5	93.0	76.5	76.5
*S*3	67.0	69.5	59.5	44.0	75.5	65.0	64.0	51.0
*S*4	212.0	213.0	197.5	186.5	225.0	190.0	197.5	182.5
Mean ISI_c	120.75	114.00	102.13	94.88	121.00	105.63	101.88	94.13
Mean SOA	130.75	134.00	132.13	134.88	131.00	125.63	131.88	134.13

unequivocal. The minimal SOA for criterion performance was constant for monoptic and dichoptic masking. This suggests that the additive component isolated monoptically is the same as that isolated by the dichoptic procedure. In addition, the present data, taken together with those of Exp. VII, imply that central processing time is not influenced by peripheral processing time.

APPROXIMATIONS TO A MODEL FOR MASKING

Another Look at the Relation Between the Two Processes

In the general discussion of Exps. I - IX, the rudiments were spelled out for a theory of the recognition of visual stimuli. It was proposed that visual pattern recognition involves at least two distinct stages and that these two stages are successive and additive, the two stages in question corresponding to the processes represented by the multiplicative and additive rules. Several current theories of pattern recognition take the same form. Neisser (1967), for example, has proposed that an initial preattentive process which segregates objects in the optical array, and which may signal the presence of easily discriminable physical features, precedes a second stage of focal attentive processing which makes extensive contact with long-term storage and is essential for stimulus recognition. Sternberg (1967) has similarly argued for a successive-additive model of stimulus classification. While the present research may be viewed as substantiating one aspect of such theories, which is that there are several distinct processes underlying pattern recognition, the results of Exp. XII raise serious doubts about the postulation that the processes are successive and additive.

Figure 13 shows two functions relating T duration to ISI_C. One of these functions was generated by RN as mask and the other by PM3, and as noted above, these functions are fundamentally different. For present purposes what is important about the PM3 function is the invariance of T duration + ISI_C in the T duration range from 4 to 64 msec. What this means essentially is that the central process as identified by the additive rule was requiring a certain amount of time between stimuli onsets to identify the T stimulus and that this amount of time was constant and unaffected by the duration of the T stimulus. Inspection of the RN function, on the other hand, tells us that the time needed to complete peripheral processing varied with the duration of the T stimulus. This is how ISI_c in the multiplicative rule has been interpreted; it identifies the minimal time needed by peripheral processes to signal the features of the stimulus. The problem for the successive-additive postulation lies in this fact: while the peripheral processing time varied with T duration, the central processing time was constant. But if the processes or stages are sequential and the central processing time is measured as the elapsed time between onsets, then central processing time must include peripheral processing time as well. The implication is that the two processes are not conducted in sequence but instead overlap in time.

In recent tests of sequential two-stage theories such as Neisser's (1967) some evidence has appeared which, like that of Exp. XII, questions the sequential-additive assumption. For example, Ellis and Chase (1971) have shown in a variation of Sternberg's character-recognition paradigm (see Sternberg, 1969) that the time for item recognition or size discrimination alone is the same as the time for item recognition or size discrimination in a combined task. Item recognition is assumed to require focal attention, and size discrimination can be performed by preattentive processes; the conclusion, therefore, was that focal attentive and preattentive processes can occur in parallel. Beller (1968) reached a similar conclusion. He used a task in which Ss searched through displays of two- and three-digit numbers for targets whose size was specified at the outset. Nontargets of the same size and nontargets of different size comprised the noise items. Beller observed that varying the difficulty of the size discrimination affected the time needed to reject the different-sized nontargets (i.e., preattentive processing time) but not the time needed to reject the same-sized nontargets (i.e., focal attentive processing time). However, it must be borne in mind that these observations, while theoretically illuminating, need not necessarily be speaking to the stages reflected in the present data.

What is needed on the evidence of Exp. XII is a restatement of the relation between peripheral and central processes, or more precisely, between the processes symbolized by the multiplicative and additive rules. Two criteria must be met. First, any proposed relation must account for the invariance in the central processing time with varying peripheral processing time as evident in the range 4 to 64 msec of the PM3 function of Figure 13. Second, it must account for why the upper limit on the masking range of PM3 for the brief T durations of 2 and 3 msec is apparently set by peripheral processing time and not by central processing time.

Two possible hypotheses present themselves. One is that the processes symbolized by the multiplicative and additive rules are not allied at all; they are operationally parallel. The other is that the two processes overlap in time, but one is contingent on the other.

The first hypothesis requires discarding the notion that the multiplicative rule speaks to peripheral events and the additive rule to central. To say that the processes are operationally parallel is to say that they work independently of one another, and given the earlier anatomical localization of these rules, this is tantamount to saying that central processes are not contingent on the output of peripheral processes, which is nonsensical. On this view, the two principles, T duration x ISI_c = a constant, and T duration + ISI_c = a constant, are seen as representing simply two operations in vision rather than as indicants of peripheral and central processes.

To meet the criteria posed above, an operationally parallel view of the two processes must carry the rider that the rule describing the minimal time needed to evade masking by a pattern (such as PM3) must depend, for any circumscribed range of energy values of the T stimuli, on which of the two processes takes longer. In the PM3 function of Figure 13, masking at the T exposures of 2 and 3 msec is best described by the multiplicative relation, while at the longer exposures, the additive relation is more suited. From the data illustrated in Figure 13, it may be concluded that at the exposures of 2 and 3 msec the processing symbolized by the multiplicative rule took longer, and at the exposure durations of 4 to 64 msec it was the processing described by the additive rule which was more durable. Thus, for the exposures of 2 and 3 msec it may be inferred that the operation characterized by the additive rule was complete by a SOA of approximately 58 msec, while the operation characterized by the multiplicative rule was still in progress. Therefore, up to an SOA of 58 msec or so, PM3 could interfere with either or both processes; beyond that SOA, however, the aftercoming stimulus could only interfere with the process characterized by the multiplicative rule. Since the dependent variable was the minimal time needed to evade masking, the obtained estimate of that minimal time would, on this view, mirror the properties of the process underlying the multiplicative relation between T duration and ISI_c. This would hold only for those durations of T exposure at which the process underlying the multiplicative relation took more time than that underlying the additive relation. Where this criterion is no longer met, the estimate of minimal time needed to evade masking would mirror the properties of the process described by the additive rule.

In order for the operationally parallel view to account for the PM3 function, the assertion has had to be made that both operations must be concluded in order for the T stimulus to evade masking by PM3. This is equivalent to saying that for identification to occur, both operations must be complete, which implies perhaps that they cannot be orthogonal. On the other hand, it may imply only that some subsequent decision mechanism cannot output an identification until inputs from both processes are available.

Perhaps the strongest argument against the operationally parallel view is that the data of the present research point to a distinction between the two processes that is, in a nontrivial sense, anatomical. The multiplicative relation was most surely grounded in those circumstances which allowed for peripheral interaction, that is, in conditions of monoptic and binocular presentation. Indeed, the multiplicative relation was realized only in these conditions, Kinsbourne and Warrington (1962a, 1962b) to the contrary. In addition, only the multiplicative rule was engendered across T durations by RN, a mask which failed to impede letter perception in dichoptic presentation. Furthermore, the stimulus parameters of duration and intensity, immaterial to dichoptic masking by PM or PM3, were the determinants of monoptic and binocular masking by RN. In short, an anatomical distinction between the two processes along the lines peripheral-central is strongly demanded by the data.

A Concurrent and Contingent Model of the Peripheral-Central Relation

An alternative to the successive-additive and the operationally parallel interpretations is that the processes overlap temporally and that one process, the central, is contingent on the output of the other. This approach preserves the central/peripheral distinction nurtured in the earlier arguments of the present paper.

The essence of such a view is that the central process receives data intermittently from the periphery. This implies two things: there are a number of different peripheral systems or neural nets, and these peripheral systems may output data at different rates.

The form that such peripheral nets might take is suggested by a consideration of the selectivity manifested by individual cells in the visual systems of vertebrates such as cat and monkey. We know, for example, that certain neurons respond only if the input to the retina has a particular size, shape, or orientation or moves in a certain direction (Hubel and Wiesel, 1962, 1965, 1968). However, what is important to note here is that this selectivity is the result of an operation performed by a fairly large neural system, served in part by spatial summation and lateral inhibition and including many receptor units and intermediate neurons in addition to the cell in question. Neural systems of this sort exhibit certain features that are important to the present discussion (see Thomas, 1970). First, each is selectively responsive to a certain characteristic of stimulation. Second, although the different systems may have receptors and intermediate neurons in common, they are for the most part independent. Third, an input to the retina will affect several or all systems simultaneously, but only some will respond to it; i.e., only some systems can output a characteristic of the input. And fourth, each system has a "preferred feature condition," that is, it responds best when the feature to which it is selective is present in a particular way. With straight-line contour detectors, for example, the strongest response is given when the line is in a particular orientation. The strength of a system's output varies inversely with the degree of difference between the preferred condition of the feature and the actual condition. Thus, the output from these systems is graduated.

Evidence for parallel perceptual systems of this sort in the human has been accumulating. (See Weisstein, 1969, for a recent review.) The explanation of hue perception by reference to separate, parallel systems is, of course, not new. Recently several experiments have argued for the existence of systems which are both selectively sensitive to color and tuned to a narrow range of edge orientations (Held and Shattuck, 1971; McCollough, 1965). Sekuler and his colleagues (Pantle and Sekuler, 1968; Sekuler, Rubin, and Cushman, 1968) have proposed that mechanisms exist which are sensitive to the direction of movement and contour orientation. And several papers by Thomas (e.g., Thomas and Kerr, 1971) have argued that stimulus detection is mediated by mechanisms which are at least crudely size tuned.

There is also some evidence favouring the view that different properties of stimulation are ascertained at different rates. Kahneman (1967a), for example, has shown that brightness and contour data have different rates of formation. The experiments of Fehrer and Raab (1962) and Fehrer and Biederman (1962) reveal that information about stimulus onset is available well in advance of data on contour and that the former may be available in the phenomenal absence of the latter. And Cheatham (1952) has reported, albeit contrary to intuition, that the perception of contour precedes the perception of hue. In sum, there is reason to believe that different operations may be going on simultaneously at different rates (Kolers, 1967; Weisstein, 1971).

The following sketches the details of a concurrent-contingent model which relates the peripheral and central processes. The model is illustrated in Figure 15.

(i) I is an input to a particular retinal location from the set of all possible inputs to either or both eyes. For present purposes we will talk only about input to one eye.

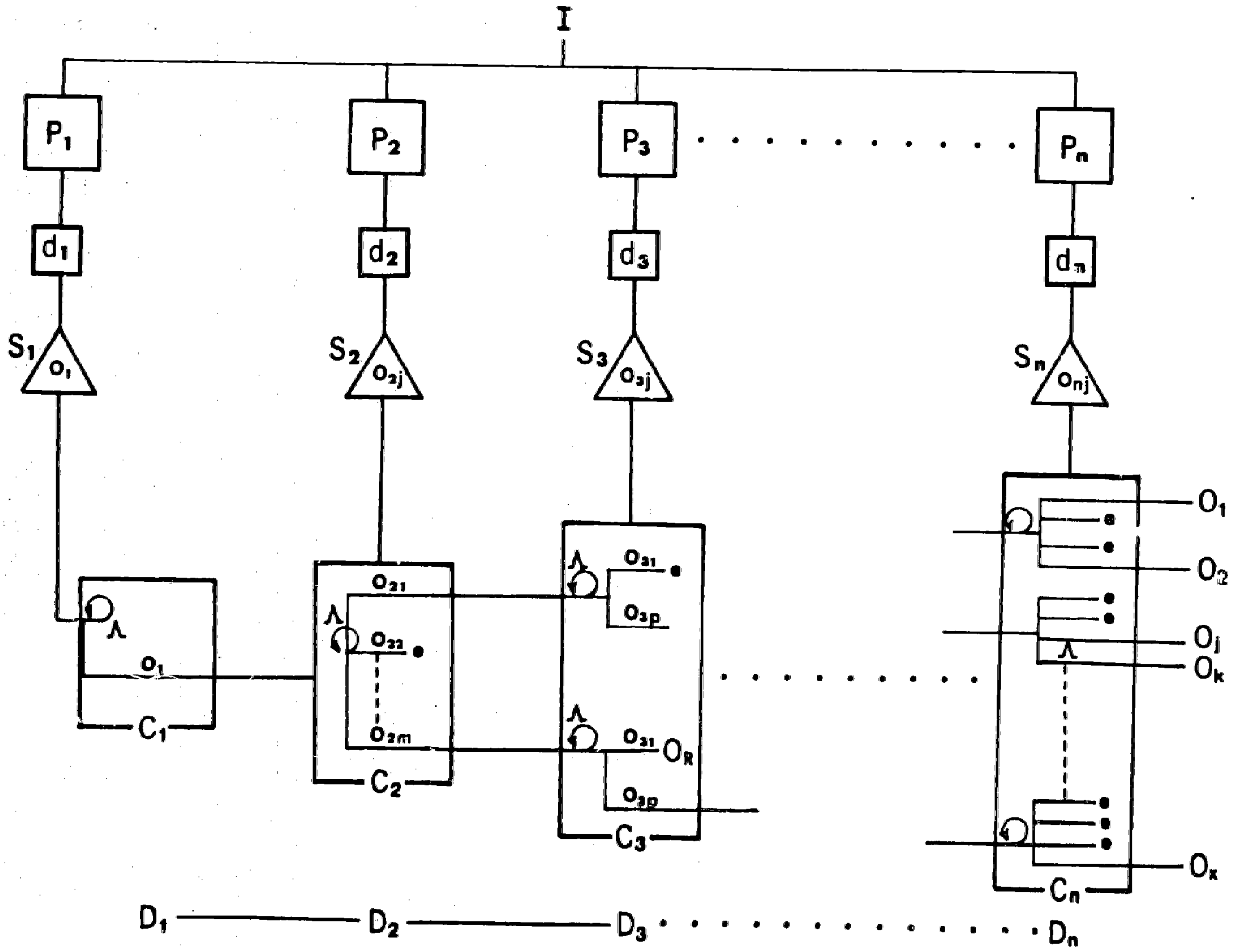


Figure 15: Schematic representation of the concurrent-contingent model.

(ii) The multiplicative rule characterizes the workings of the peripheral mechanisms, $P = [P_1, P_2, \ldots P_n]$, a set of "neural nets" or "logical units" which all have the same input I but which give rise to different outputs. The assumption is made that the peripheral nets are operationally parallel. We will presume that the two sets of peripheral systems, P_{right} and P_{left} are functionally equivalent. Only one P set will be discussed for simplicity.

(iii) O_{ij} is an output of a peripheral net P_i, where O_{ij} belongs to the set $[O_{i1}, O_{i2} \ldots . O_{in}]$ and n, the number of outputs for each P_i is finite and varies for different P_i.

(iv) Peripheral net outputs are realized at different times after I onset. Operating times for peripheral nets are symbolized $d_1, d_2, \ldots d_i, \ldots d_n$, such that, in general, $d_1 < d_2 < \ldots d_i < \ldots d_n$.

(v) For any peripheral net P_i, operating time varies as follows: (a) When T energy $<$ the minimal energy, E_{min}, required to elicit a terminal response in a peripheral net, there is no new output; (b) When $E_{min} <$ T energy $<$ maximum energy, E_{max}, operating time varies inversely with energy; (c) When T energy $> E_{max}$, operating time is at some fixed minimum.

(vi) Peripheral net outputs are stored in central storage units, S, for use by the central decision process. The base state of any storage unit, S_i, is Λ, the null state. This state can only be changed by the entry of real data, O_{ij}, from the peripheral net, P_i. The record of O_{ij} either decays with time, returning S_i to state Λ, or is replaced by the record of another O_{ij}.

(vii) We will presume that there is only one set of stores, $S = [S_1, S_2, \ldots S_i \ldots S_n]$, for the outputs of the two peripheral systems, P_{right} and P_{left}. In other words, for a corresponding region of the two retinas, the outputs from right-eye nets and left-eye nets are entered into the same storage units.

(viii) The central process, C, is also a set of nets, $[C_1, C_2, \ldots C_i, \ldots C_n]$, whose serial operations can be conveniently represented as a decision tree in which each C_i consists of a set of nodes on the tree. The additive rule characterizes the workings of the central process.

(ix) For any central net, C_i, two sources of data are necessary for a decision: an input from the appropriate P_i and a decision from the preceding central net, C_{i-1}.

(x) The final output of the central process is O_k belonging to the set $[e, O_1, O_2, \ldots . O_x]$, where e is the null element.

(xi) Figure 15 illustrates the decision process. For any given input branch, a decision by C_i will be made depending on the output O_{ij} (other than Λ) in S_i. Either a branch of the tree will continue to C_{i+1}, or it will terminate in a final output O_k, or it will terminate in the null output e, meaning no output possible. Thus we say that C is a pruned tree.

(xii) If a decision branch from C_{i-1} finds Λ as the record in S_i, then two possibilities arise since Λ means that no particular output, O_{ij}, has been made by the peripheral net, P_i. In general we would expect C_i to wait for some output, O_{ij}, by looping at the input (symbolized $\circlearrowleft$). This looping would have an upper time limit greater than that of the slowest operating time of P_i, at which point the decision branch would terminate in the final e state. Sometimes, however, Λ in S_i would be a permissible output from P_i, and C_i would compute its decision in the normal way.

(xiii) D_1, D_2,...D_i,...D_n are the operating times for the central nets. For any C_i, the operating time, D_i, is constant across input branches from C_{i-1} and outputs O_{ij} from P_i. Thus, as long as all central nets receive their inputs from their respective peripheral nets simultaneous with, or prior to, the input branches from the preceding nets, $D_1 + D_2 + \ldots D_n$ is a constant.

(xiv) P_1 and C_1 are, respectively, the peripheral and central nets detecting and identifying onset or change of stimulation

Peripheral processing time (PPT) refers to the time needed to complete a subset, or the complete set, of peripheral net operations, and central processing time (CPT) refers to the time needed to reach a particular decision. It is assumed that both PPT and CPT have varying upper limits determined by the characteristics of I and that the upper limit on CPT is also determined by the task, e.g., detection, identification. To be more precise, the subset of peripheral nets which will output data on I, i.e., the number and type of nets engaged, is constrained by the nature of I. A contourless flash of light will not occupy the same number or type of peripheral nets as would be occupied by a contoured flash; obviously networks determining intensity, duration, and size are involved in both, but networks determining inhomogeneities in the input array are needed only for the latter. Also, the full complement of central decision nets needed to identify an input as the letter A would not be needed to identify the occurrence of a stimulus (cf., Fehrer and Raab, 1962) or the presence of a vertical (vs. a horizontal) line.

Based on the data of Exp. XII, illustrated in Figure 13, the following statements can be made on the relation between PPT and CPT: (a) when PPT $<$ CPT, the upper limit on masking is CPT; (b) when PPT $<$ CPT, CPT is constant and does not vary with PPT; and (c) when PPT $>$ the constant CPT identified in (b), the upper limit on masking is PPT.

The model rationalizes (a), (b), and (c) as follows: when PPT $<$ CPT (a condition which is met when I energy is "substantial" as in the region 4 to 64 msec of Figure 13) outputs from peripheral nets are running ahead of decisions by the central decision nets to a degree depending on the energy of I. Thus for C_i, O_{ij} is stored in S_i awaiting the decision of the preceding stage, C_{i-1}. The decision process of C_i begins only when both O_{ij} and the decision of the preceding central net are available (see ix and xii above). Since the decision time for C_i is constant (xiii above), the decision of C_i is received by C_{i+1} after a constant delay; therefore, C_{i+1} cannot benefit from the earlier arrival of $O_{(i+1)j}$. In short, when PPT $<$ CPT, reducing PPT by increasing I energy will not decrease CPT; the constraint on CPT is the time-constants of the individual decision stages. Under these conditions, then, CPT sets the upper limit on masking by PM3. However, decreasing I energy retards PPT to a point where, for C_i (other than C_1), the decision from C_{i-1} is received prior to O_{ij}. In this case, the C_{i-1} decision is delayed (see xii above) and the constraint on CPT would no longer be solely the time-constants of the individual central nets but also the delay time for peripheral net outputs. In this instance, the upper limit on backward masking by PM3 would be determined by PPT.

We may now reexamine the issue of peripheral and central backward masking. The term "peripheral" has emerged in the present context as a rubric for systems which extend from receptor surface to cortex and which underlie the extraction of properties of visual stimulation.

Given the principles above, the second, and invariably stronger, stimulus in the present series of experiments would be processed by the peripheral systems more rapidly than the first. A situation, therefore, can exist in which a peripheral net is simultaneously occupied by two events presented in close succession. Under this condition of double occupancy the output of a peripheral net will depend on two things: whether both stimuli elicit terminal outputs from the net in question and the order of time elapsing between the two stimuli.

Consider the case where the net gives a terminal response only to the T stimulus. Since the mask covers the same receptor surface as T, the peripheral systems which will eventually output properties of T will at some early stage be affected by the mask. An early stage in a peripheral system may be so occupied by a response to the masking event that there is no room left for a response to the first stimulus. Or the response at an early stage may be to the combination of T and mask and thus the input to later stages of the peripheral net is distorted. The probability of perturbations of this sort occurring drops off sharply as the time elapsing between the two stimuli increases.

Both of the above means of affecting peripheral net function are included in the condition in which the peripheral net can give a terminal response to either T or the mask. However, since all stages of this peripheral net respond to the mask, the temporal range over which the mask may impair or occlude a terminal response to T is extended. In brief, the greater energy mask may in this case "overtake" the T stimulus at any stage in the peripheral net.

The implication is that when masking is peripheral in origin, the upper limit on ISI_c for a T of given energy is set by the slowest operating peripheral nets outputting data on T. The extent to which this upper limit is realized depends on the extent to which the second stimulus, the mask, elicits terminal responses from the same set of peripheral nets. Therefore, we should expect the severity of peripheral masking to vary as a function of the relation between T and mask. The earlier discussion on the differences between PM, RN, and a contourless light flash as masks is relevant to this point.

A purchase on masking of central origin may be gained by speculating on differences between masks that function only monoptically and masks that function either monoptically or dichoptically for a given set of T stimuli. In the context of the present series of experiments, this reduces to speculating on the differences between RN and PM.

It may be argued that the identification of RN for the most part is not based on outputs from the peripheral systems required for the identification of the T stimuli. Pursuing this further, it may be argued that if data on T have been laid down in a subset of the central stores, S, most of these data cannot be replaced by data on RN since data on RN are entered into a relatively nonoverlapping subset of S stores by virtue of the fact that different peripheral systems have extracted them. On this reasoning, RN can impede the identification of T only when it has the opportunity to affect the peripheral systems responding to properties of T. This impedance arises, as described above, by occluding outputs at early stages in a system, or by degrading outputs.

Consider RN and T presented dichoptically. In this situation RN obviously cannot occupy the peripheral systems abstracting properties of T. Thus at brief SOA's what is represented in the set, S, of central stores are all the properties of T and all the properties of RN, represented respectively in independent subsets of S. What is perceived is both T and RN.

In some circumstances it is conceivable that properties of two successively presented stimuli may be represented simultaneously in relatively independent subsets of S, and yet confusion, i.e., failure to identify the T stimulus, may occur. An example may be found in experiments using computer-generated dot stimuli, in which patterns are masked by nonoverlapping dynamic visual noise (Uttal, 1970, 1971b). The central decision process does not yield a distinction between the two dot stimuli; both are "perceived" and the masking results from failure to segregate the signal from the noise. A similar situation could also lead to fusion in which the two stimuli are integrated to yield a single identifiable form (e.g., Eriksen and Collins, 1967). In either of these cases, however, the degree to which masking or fusion occurs is dependent on the time elapsing between the two stimuli and the extent to which data on the first have decayed, i.e., the extent to which the central stores have returned to the null state.

In contrast to the argument on the identification of RN, it may be argued that the identification of PM does rely on outputs from some peripheral systems in common with those underlying the identification of T. Peripheral masking by PM should occur for the reasons cited above, and we should expect such masking to extend over greater intervals than the corresponding masking by RN. On the other hand, central masking by PM arises from the fact that data on PM can replace data on T in the set of central stores. Assume that a complete peripheral description of T is available in the central stores before the input of PM. The processing of the mask by the peripheral nets leads to a change in some of the stores, $S_1, S_2, ..S_n$. The number of stores that change depends on the number of peripheral nets common to the processing of T and PM. During the peripheral processing of PM, the central mechanisms have been making decisions on the nature of T. At some point in the decision series, however, data on the mask will enter into the ongoing decision on T. The point at which PM data enter into this decision process is determined, in part, by how soon data on PM replace data on T in the set of central stores. If replacement occurs before the decision process has progressed very far, the central mechanisms may fail to architect any perception of T whatsoever. In this circumstance the C_i decision may have specified a different branch in the subsequent stage, C_{i+1}, from that which would have been taken if mask data had not replaced T data in S_i, or C_i may have reached a null decision. The latter is unlikely if only low-level decisions had preceded C_i.

With increasing time before replacement, the central decision may have proceeded to the point where substantial data on T and PM are incorporated into the decision process. The result is a composite perception, but one which does not allow for a segregation of T from PM.

Quite obviously masking of central origin does not occur when entries in S_1, S_2,..S_n are changed after completion of the central decision process or when entries in earlier stores, e.g., S_3, S_4, are changed as the decision process is reaching the later stages, e.g., C_{n-1}, C_n.

In the model, the peripheral nets and central decision nets have been described for a single item input to one retinal location. Simultaneous presentation of several items to several locations would be represented by a simple replication of the basic model. With several objects or figures present at input, the peripheral-central net complexes serving the items, one complex to each, would yield a number of final outputs, one to each of the corresponding C_n decision nets. Thus peripheral-central net complexes operate in parallel over the visual field.

The concurrent-contingent model as described is not so much a formal theory as it is an example of a particular class of theory of visual masking. In its emphasis on stimulus-analyzing mechanisms, on selective interference with stimulus attributes, and on central decision mechanisms, it contrasts with theories of the integration and interruption type which view masking in terms of relatively global processes. Because the model is intended mainly to exemplify an approach, certain details have been left unspecified (for example, the identity of the peripheral nets' output, i.e., the kind of features represented; the relation of the concept of central stores to the concept of iconic memory; the form of the C_n net output; and the decision processes, if any, beyond C_n). The issues involved in making explicit these aspects of the model will be taken up in a subsequent discussion. For the present, attention is directed to an examination of forward masking of peripheral and central origin.

EXPERIMENT XIV

Kolers (1968) proposed the clerk-customer metaphor in response to the question: Why is greater interference exerted on the preceding rather than on the subsequent presentation? i.e., why are masking effects primarily backward? However, Kolers's metaphor as it stands does not rule out forward masking. When two customers enter a store the later-arriving customer usually has to queue while the clerk takes care of the earlier customer. An implication of queuing is that central forward masking should occur. Yet, since queuing is not the same as receiving insufficient service, we should not expect, on the analogy, forward and backward masking to give rise to the same type of perceptual interference.

To pursue Kolers's reasoning a little further: "The phenomenon of backward masking itself identifies a 'formation time' and a perceptual 'refractory period' in the nervous system governing the construction of a perceptual representation" (p. 38). The lagging mask stimulus, therefore, disturbs that process identifying (constructing) the earlier T input. With the mask leading, this disturbance in the identification of T is absent; at worst, T is denied immediate access to the central process.

The concurrent and contingent model does provide for the occurrence of a mild, central forward masking effect of a somewhat different nature than that implied by queuing. If data on the mask have already been entered into the central stores, S, and data on an after-coming T are now entered, the T data will replace some, but not all, of the mask data. This is so because the mask tends to cover a slightly larger retinal area than T, and, thus, while the two stimuli have some peripheral nets--and, therefore, central stores--in common, other nets, and their stores, are only responding to, and storing, data on the mask.

A situation may, therefore, exist in which the central stores contain data on both stimuli. The decision process in this circumstance may yield a composite perception in which T is inseparable from the mask. The probability of failing to identify T in this circumstance, however, should extend over a relatively small range of delays between the two stimuli. On the concurrent and contingent model forward masking of central origin must, by necessity, be a rare event; the later-arriving stimulus always overrides the earlier stimulus in the set of stores which they share. Thus, data sufficient for identifying T are always available to the central decision process in central forward masking, which is not true of central backward masking.

The model similarly predicts little, if any, peripheral forward masking. A higher-energy mask would pass through the peripheral nets well in advance of a following lower-energy T. In this case the data in the central stores would be mask data, replaced soon after by T data, thus bringing about the situation described in the preceding paragraph. In short, serious forward masking of peripheral origin should not occur unless it is assumed that processing a stimulus raises the threshold in the peripheral nets, thus suppressing subsequent lower-energy stimuli. Peripheral forward masking is to be expected even if central is not; a survey of the literature (Kahneman, 1968) shows substantial evidence for monoptic forward masking in contrast to the sparse evidence for dichoptic forward masking.

The present experiment compares forward and backward masking by PM3 under conditions of monoptic presentation. The procedure follows, *mutatis mutandis*, that of Exp. XII. The expected outcome was as follows: at brief durations of T, masking should be severe for PM3 leading and lagging; at the longer durations, the lagging function should match the additive rule, the leading function should not, and only in the lagging case should the masking be pronounced.

Method

Four *Ss*, three naive and one experienced (*S1*), participated in the experiment. The T material was the set of consonant trigrams. The luminances of T and PM3 were 2.5 ft L and the exposure duration of PM3 was 50 msec. For these durations of T--2, 3, 6, 8, 24, 40, and 56 msec--ISI_c was determined in the usual manner. Forward and backward ISI_c's were determined in succession at any particular T duration. Thus, Ss 1 and 3 at each T duration were given the forward arrangement first, and Ss 2 and 4 were given the backward arrangement first. The T durations were examined in the order shown. Stimuli were presented to the right eye.

Results and Discussion

Individual S data are given in Table 8. Graphic representation of the averaged data is given in Figure 16.

An important feature of Figure 16 is the resemblance that forward masking by PM3 bears to backward masking by RN shown in Figure 13 of Exp. XII. The forward masking function, up to and including T = 8 msec, fits the multiplicative relation reasonably well. This suggests that PM3 forward masking was very much a peripheral event, a suggestion further advanced by the dissimilarity between backward and forward masking at the longer T durations. Here, the backward masking function was clearly of the central type with the appropriate description being, as before, T duration + ISI_c = a constant. There is nothing in the forward masking function at these durations to suggest a similar central effect, although there does appear to be some central forward masking. The fact that forward masking by PM3 still occurred at the T duration of 40 msec, well beyond the duration at which RN became ineffective in Exp. XII, means, perhaps, that a leading PM3 can exert some influence centrally.[6]

Another important feature of the data is the considerably greater ISI_c needed in the forward case at the brief T durations. If only the brief durations of T had been investigated, the conclusion would have been that monoptic forward masking by pattern was more severe than monoptic backward masking by pattern. Such a conclusion was reached by Smith and Schiller (1966) who used a T duration of 2 msec. However, what is obvious from inspection of Figure 16 is that whether forward masking is more severe than backward depends on whether the phenomenon is of peripheral or central origin. Indeed, it was abundantly clear to Smith and Schiller that although monoptic masking of a 2-msec stimulus was more severe when the mask led, dichoptic masking was most severe when the mask lagged.

Smith and Schiller (1966) concluded that "forward masking [by pattern] seems to be mainly a monoptic phenomenon" (p. 196), a conclusion substantiated by Greenspoon and Eriksen (1968) who noted that dichoptic forward masking by pattern was quite weak. This conclusion, given the present data, can be stated more usefully: forward and backward masking can both occur peripherally, but only backward masking occurs to any appreciable degree centrally. Therefore, when two stimuli are in competition for the services of the central decision process, it is the later-arriving one which is most completely identified. On the other hand, when two stimuli compete for the same peripheral nets, order of arrival is less important than energy.

---

[6] Observation of dichoptic forward masking by PM3 confirmed this: forward masking occurred but over a much smaller range than backward masking.

TABLE 8

EXP. XIV: CRITICAL ISI AS A FUNCTION OF T DURATION FOR FORWARD AND BACKWARD MASKING BY PM3

T Duration (msec)	Forward					Backward				
	Subjects				Mean	Subjects				Mean
	S1	S2	S3	S4	ISI_c	S1	S2	S3	S4	ISI_c
2	140	221	220	180	190.25	46	140	166	140	123.00
3	85	156	140	130	127.75	43	94	125	92	88.50
6	47	79	74	69	67.25	30	78	110	63	70.25
8	35	58	57	50	50.00	25	74	110	64	68.25
24	1	17	16	5	9.75	13	64	88	46	52.75
40	0	12	16	2	7.50	0	35	72	31	34.50
56	0	1	1	0	.50	0	28	50	12	22.50

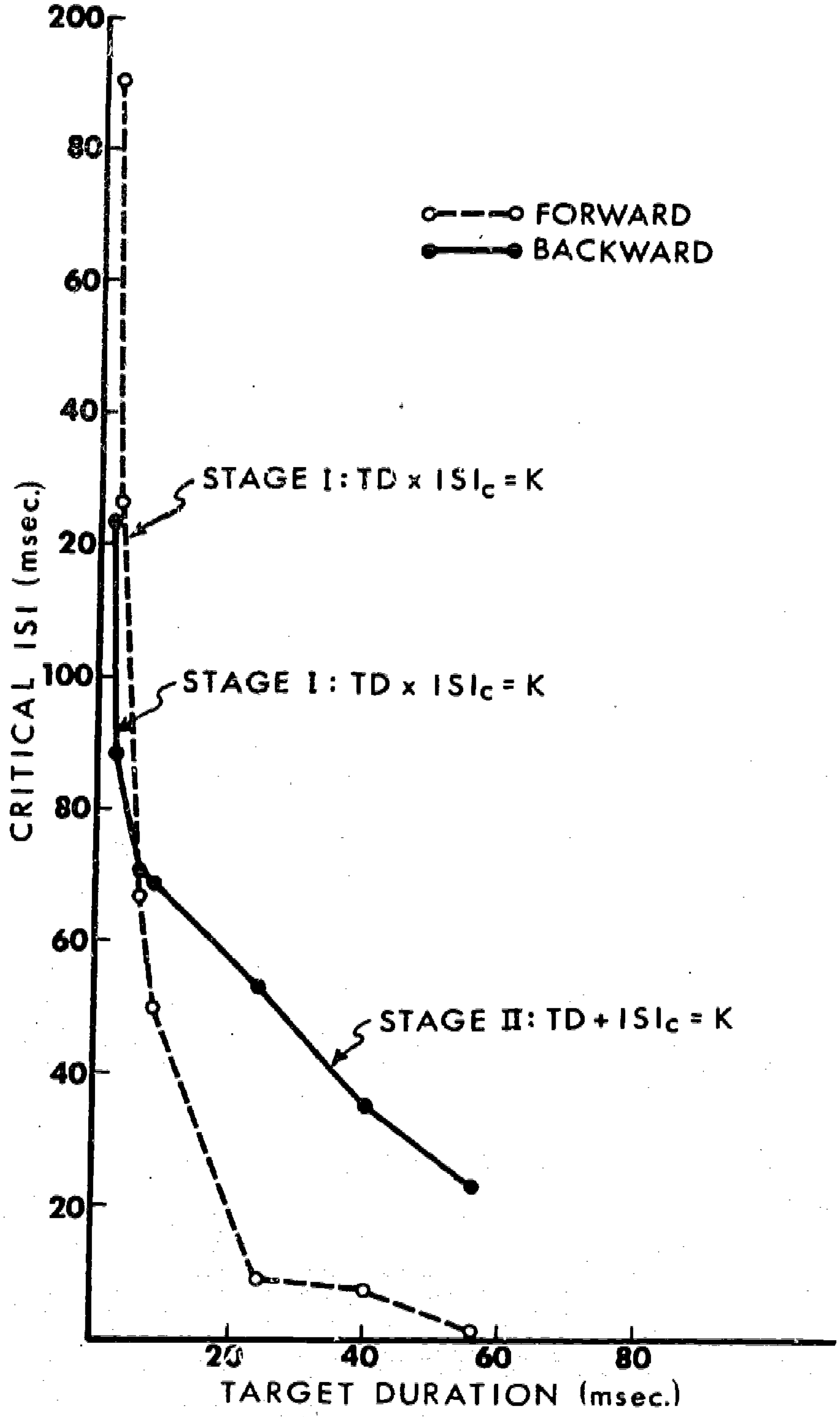


Figure 16: Relation between T duration and mean ISI$_c$ for monoptic forward and backward masking by PM3 in Exp. XIV.

The stimulus of greater energy, whether it leads or lags, will be the one whose properties are likely to be output by peripheral nets.

EXPERIMENT XV

The data of Exp. XIV suggested that a leading mask may be more disruptive peripherally than a lagging mask. The purpose of Exp. XV was to verify that this was the case, and for this reason the experiment used RN as the masking stimulus instead of PM3. Also, the experiment looks at the possibility implied by Exp. XIV and the experiments of Kinsbourne and Warrington (1962b) that the same rule, T duration x ISI_c = a constant, applies to both forward and backward masking by RN.

Method

The duration of RN was 50 msec and its luminance was 2.5 ft L, equal to the luminance of the T stimuli. Critical ISI was estimated for each of two Ss in both forward and backward masking conditions at each of these T durations: 2, 4, 6, 8, 10, 12, 16, and 20 msec. The two Ss, who were not naive to masking experiments, were tested as follows: at each T duration, going in order from 2 to 20 msec, ISI_c was determined for S1, first with RN lagging and then with RN leading; ISI_c estimates for S2 were collected in the reverse order. The T stimuli were the trigrams and the usual criterion was used to assess ISI_c. Stimulus presentation was to the right eye.

Results and Discussion

The relation between T duration and ISI_c for both forward and backward masking is illustrated in the log-log plot of Figure 17. The log-log plot facilitates comparison with Figure 3 of Kinsbourne and Warrington (1962b). Both figures demonstrate that the forward and backward masking curves relating T duration to ISI_c are of identical slope and that the relation, T duration x ISI_c = a constant, holds whether RN leads or lags. Furthermore, the absolute value of ISI_c is greater at any given duration of T when RN leads. In the present experiment the ratio of ISI_c in forward masking to that in backward masking at any T duration was approximately 2:1.

The inference to be made, therefore, is that forward masking of peripheral origin is more severe than backward masking of the same origin. A difference in this direction between forward and backward masking has been demonstrated several times by Schiller and his associates (Schiller, 1966; Schiller and Smith, 1965; Smith and Schiller, 1966) and others (e.g., Kietzman, Boyle, and Lindsley, 1971). What the data of Exps. XIV and XV do is to point to the transmission line as the locus of this difference.

EXPERIMENT XVI

Although some authors (e.g., Eriksen and Lappin, 1964) have argued that both forward and backward masking reflect a single underlying process, the data of the preceding experiments and others (e.g., Kinsbourne and Warrington, 1962b; Schiller and Smith 1965; Smith and Schiller, 1966) suggest the contrary. The present communication has argued, and demonstrated, that backward and forward masking of central origin are fundamentally different processes, and that central forward masking is rather modest at best. At the peripheral level there is some support for the notion of a common process.

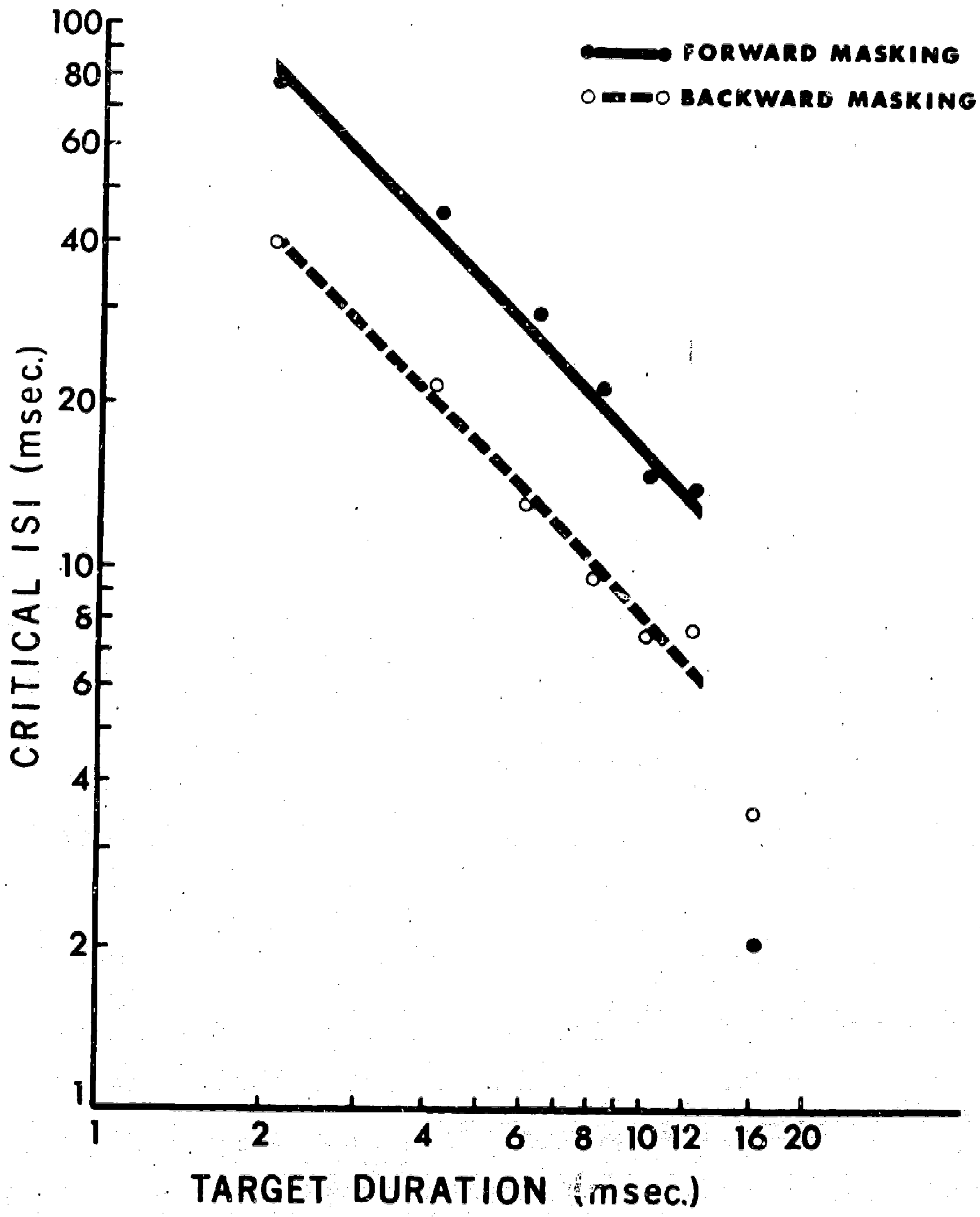


Figure 17: Log-log relation between T duration and mean ISI_c for monoptic forward and backward masking by RN in Exp. XV.

Both forward and backward masking occur substantially and they seemingly obey the same rule, but they do differ; forwarding masking, as we have seen, is more pronounced.

Experiment XVI looks for further evidence of dissimilarity between forward and backward masking of peripheral origin. The point has been made above that masking originating in the transmission line requires that the mask stimulus be of greater energy than the target stimulus. Yet, as was determined in Exps. I and III on backward masking by RN, once this criterion was met, further increase in mask duration with luminance held constant did not amplify the masking effect, i.e., ISI_c was unchanged. The present experiment compares forward and backward masking by RN as a function of mask intensity and mask duration to determine whether the two masking arrangements are differentially affected by these variables. There is some research by Schiller (1966) which has pointed to a greater sensitivity of forward masking to mask intensity.

Method

The experiment was relatively straightforward. The T stimuli, the trigram set, always appeared at the same level of illumination, 2 ft L, and the same exposure duration, 5 msec. In one condition the mask stimulus, RN, was always presented at the same exposure duration of 50 msec, but its level of illumination was 2, 8, or 20 ft L. In the other condition RN was always presented at the same level of illumination, 10 ft L, but its exposure duration was 10, 40, or 100 msec. Thus, in both conditions the T to RN energy ratios were the same. Four Ss, two of whom had served in a previous experiment, were given both conditions. Of the two possible orderings of the two conditions, two Ss were given one order and two Ss the other. Critical ISI_c's for forward and backward masking were determined in succession at each level of RN intensity or duration going in order from the lower to the higher value. Within a condition, one S of each pair of Ss was tested with the forward arrangement first. However, the S given forward masking first in one condition was given backward masking first in the other. The usual procedure was used for estimating ISI_c, and presentation of stimuli was monocular, to the right eye.

Results and Discussion

The ISI_c's averaged across the four Ss for forward and backward masking by RN in each condition are shown in Figure 18.

Since each duration had not been paired with each intensity, a single Treatment x Treatment x Ss analysis of variance could not be performed on the present data. Instead, a separate Treatment x Ss analysis was conducted, in turn, on the intensity-varying and duration-varying conditions. These analyses revealed that for both conditions, the difference between forward and backward masking was highly significant: for the intensity condition, $F(1,3) = 130.66, p < .001$, and for the duration condition, $F(1,3) = 44.23$, $p < .001$. The effect of RN luminance on ISI_c was significant, $F(2,2) = 19.19$, $p < .05$. However, the significant interaction between stimulus order (forward or backward masking) and intensity, $F(2,6) = 7.66$, $p < .05$, coupled with inspection of the left panel of Figure 17, suggests that only forward masking was affected by intensity. The duration-varying condition revealed no significant effect of duration ($F < 1$) on either forward or backward masking.

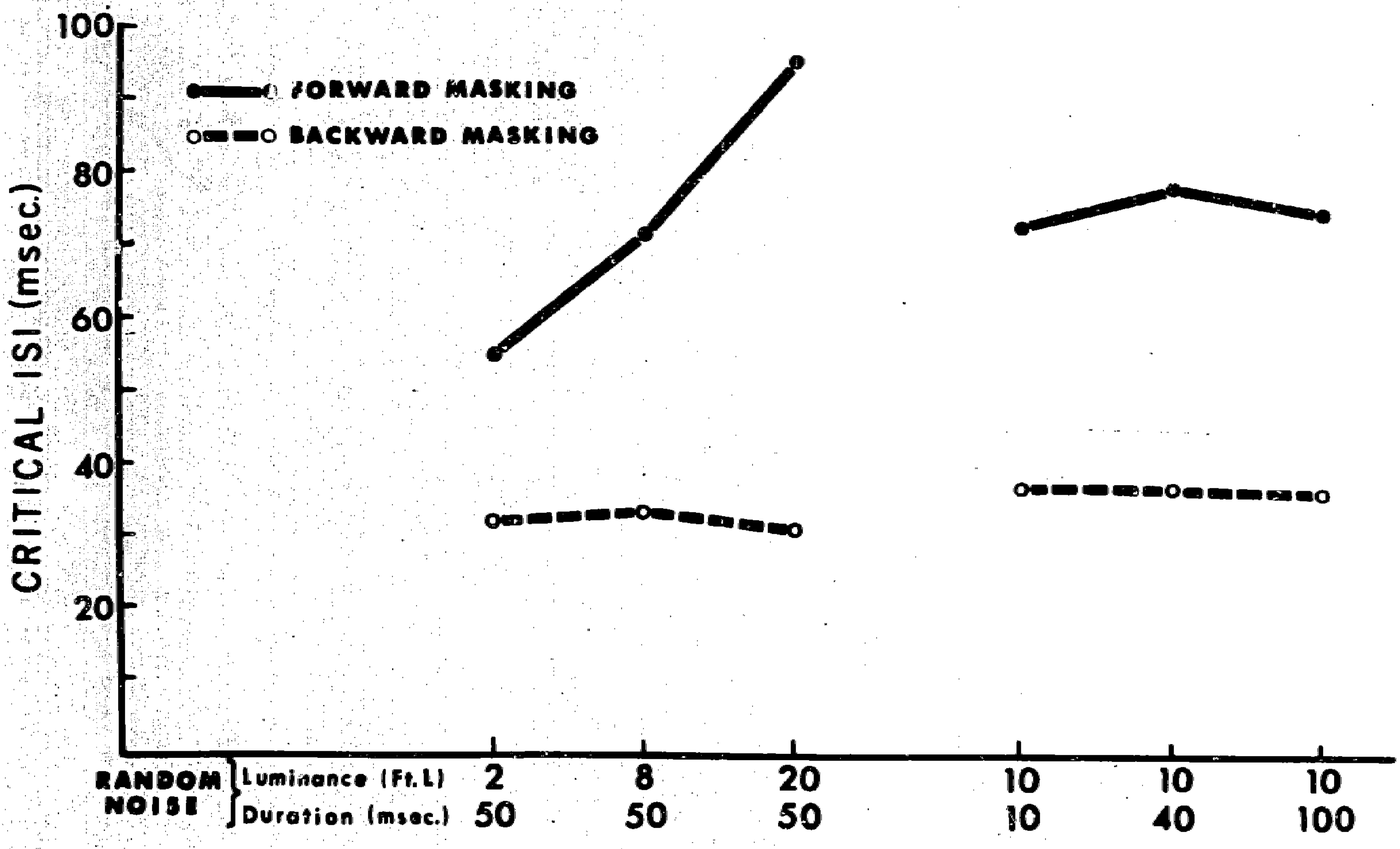


Figure 18: Monoptic forward and backward masking by RN as a function of RN intensity (left panel) and RN duration (right panel) in Exp. XVI.

In sum, the ISI_c in backward masking was invariant with respect to increase in RN intensity or RN duration. On the other hand, ISI_c in forward masking varied directly with increase in RN intensity, but like ISI_c in backward masking, it was unaffected by increases in RN duration.

The data are in complete agreement with those of Schiller (1966) and Kinsbourne and Warrington (1962b, Exp. I). Schiller showed that once mask energy was greater than T energy, increases in mask intensity were not accompanied by increases in interference in backward masking but were accompanied by increases in interference in forward masking. Kinsbourne and Warrington observed that given a mask of energy greater than T energy, increases in mask duration with luminance held constant did not extend the temporal range over which forward masking was obtained.

Forward Masking and the Concurrent-Contingent Model

The fact that the multiplicative relation between T duration and ISI_c is invariant with change in stimulus order (Exp. XV) favors the proposition that peripheral backward and forward masking reflect the same underlying process. On the other hand, the magnitude difference between the two orders and the selective effect of mask intensity on forward masking argues for two different processes rather than a single identical process.

Electrophysiological evidence suggests that forward masking arises from adaptation, i.e., reduced sensitivity, in the peripheral nets which previously responded to the mask (e.g., Nakayama, 1968; Schiller, 1968). Presumably the nets recover in sensitivity with time, and the efficacy with which previously occupied peripheral systems can process an after-coming T is determined by both the energy of T and the time since mask offset. A simple addition to the description of peripheral nets, therefore, can account for the similarities and dissimilarities between forward and backward masking functions of peripheral origin: a preceding event reduces the sensitivity of peripheral nets to subsequent events, with the sensitivity recovering exponentially as a function of time. The reduction in sensitivity of peripheral nets should not be taken to imply that all peripheral nets output a characteristic of the mask. While some nets do output a mask feature, others do not. Thus, in some nets, lowered sensitivity exists throughout the system, in others, it is limited to the early stages. The upper limit on peripheral forward masking is set by the recovery times of the slowest-recovering nets common to both mask and T stimuli. We can add one final comment on peripheral forward masking and that is, of course, that the reduction in sensitivity is directly related to mask intensity (Exp. XVI).

Forward masking of central origin is slight. The interpretation proposed above for the small effect generally found (e.g., Greenspoon and Eriksen, 1968; Smith and Schiller, 1966; and Exp. XIV) was that for a fairly limited range of delays between the two stimuli, mask data and T data are treated as a composite by the central decision process, resulting in a failure to detect and identify T. Another view of central forward interference is suggested by Kolers's clerk-customer analogy. The idea is that a later-arriving event may have to queue to gain access to a central decision process. On the perspective of the concurrent-contingent model, queuing would be interpreted as a delay in the replacement time in the set of central stores; for example, replacing mask data by T data takes longer than replacing the null state by T data. At all events, the forward interference implied by queuing would be manifest more as a delay in perception than as an impairment in perception, such as failure to identify.

EXPERIMENT XVII

On Kolers's analogy, the second of a pair of events has to queue for some finite period of time before it gains access to the central decision process. If this is so, it should be possible to detect evidence for queuing even though evidence for perceptual impairment, such as failure to see or identify correctly the second stimulus, is absent.

The paradigm developed to examine this possibility had the following form. To one eye is presented a pair of contoured stimuli, t_1 and t_2, the second lagging the first by x msec, where x is greater than the peripheral processing time of t_1. To the other eye is presented a patterned mask, m, which follows t_2 after a delay of y msec. The delay of y msec is just sufficient for t_2 in the absence of t_1 to evade the dichoptic masking action of m. Thus, when t_1 and t_2 are presented alone, t_2 is readily identified. When t_2 and m are presented alone, t_2 is again readily identified.

Now, if t_1 does in fact retard the entry of t_2 into the central decision process, then when t_1 precedes t_2, and t_2 is followed by m after y msec, failure of t_2 to gain immediate access to that process should make it susceptible to masking by m.

Experiment XVII was conducted as a demonstration of queuing rather than as a formal experiment. The stimuli chosen were as follows. The first stimulus, t_1, was the letter U located centrally; t_2 was two H's located on a slide such that if superimposed on the t_1 slide, they flanked the U. The separation between the arms of the U and the inner vertical components of the left and right H's was .18°. The m stimulus was PM3. The first and third line-configuration of PM3 overlapped the two H's and the middle line-configuration overlapped the U, if superimposed.

The three conditions described above, and depicted in Figure 19, were examined. Three naive and one experienced Ss participated. First, for each S an ISI was determined between t_1 and t_2 which yielded a fairly good and consistent metacontrast effect, i.e., S reported either that he failed to see U (only one S reported such a failure) or that the U was of "ghost-like" character and apparent movement was strongly present. For three Ss this ISI value was approximately 100 msec; for the remaining S it was 80 msec. Second, the minimal ISI between t_2 and m at which t_2 could always be seen and identified was determined for each S. This value, y, varied between 50 and 70 msec across the four Ss. Third, the three stimuli were presented in succession at the determined x and y values. Throughout, t_1 and t_2 and m were all exposed at 10 msec and 8 ft L. The right eye received t_1 and t_2, and the left eye received m.

Results and Discussion

The results of this demonstration were as follows. For each S in Condition 3, U was clearly seen with PM3 as background, and the pair of H's was not. Switching back and forth between Conditions 2 and 3, that is, simply

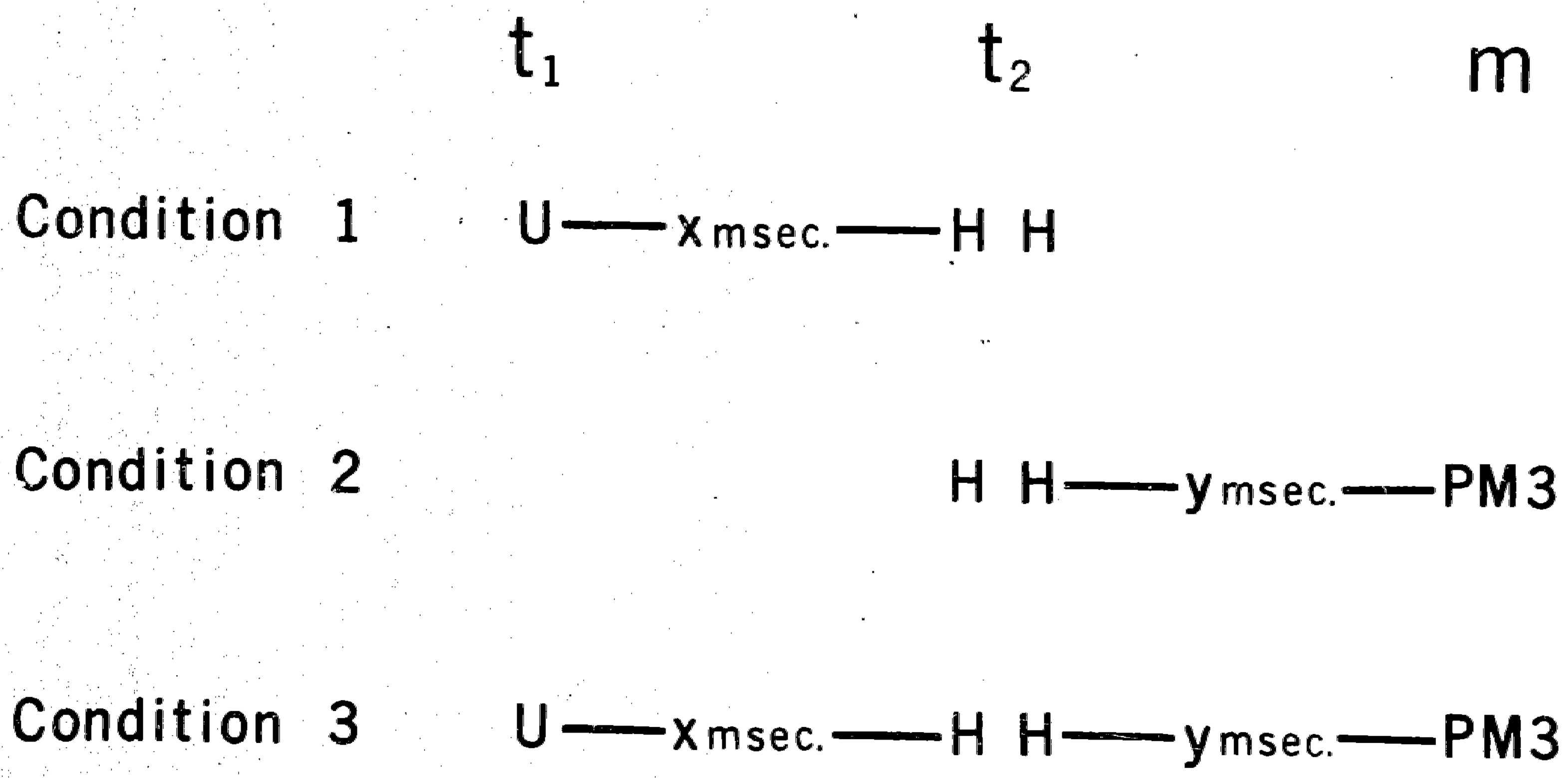


Figure 19: Schematic representation of conditions in Exp. XVII.

turning U on and off, revealed that whereas the pair of H's followed by PM3 were clearly seen in the absence of U, they were not seen in its presence. And, of course, the fact that U was more identifiable in Condition 3 than in Condition 1 is further evidence of the effect observed in Exp. IV, which is that an after-coming mask (PM3) may reduce or eliminate the interfering effect of a preceding mask (the pair of H's) on the perception of an earlier-presented T (in this case, U). The present experiment stands in contrast to Exp. IV in that here the "disinhibiting" effect (see discussion of Exp. IV) is purely of central origin.

Three further experiments/demonstrations were conducted with the same Ss. One showed that the queuing effect could be obtained with an overlapping U and H centrally located in their fields and with PM as the third stimulus. These x and y values were not identical to those of the main experiment. For expository purposes, these stimuli did not provide as good a demonstration as those described above. This experiment showed, however, that metacontrast conditions were not essential to the demonstration of queuing.

A second experiment showed that the queuing effect, i.e., the effect of t_1 upon the susceptibility of t_2 to m, could be obtained just as well when U was presented to one eye and the flanking H's and PM3 were presented, at the same x and y values as before, to the other. This rules out the notion that U retards the processing of the H's by lowering the sensitivity of peripheral nets, a possible interpretation of the queuing effect described in the main experiment. On the contrary, the queuing effect is indeed central.

A third and final experiment provided further corroboration of the queuing hypothesis. A prediction from this hypothesis is that in Condition 3 the likelihood of m masking t_2 at ISI = y msec should decrease with increases in x. This prediction was demonstrated for all four Ss by holding y constant and increasing the value of x. What was surprising, however, was that for the four Ss the value of x at which the flanking H's became visible was fairly substantial, of the order of 200+ msec. This implies that the locus of queuing was not in the central stores. If replacement time were of this order of magnitude, it would be difficult to account for any central backward masking.

In summary, Exp. XVII demonstrates that central queuing does occur with the effect probably localized at a relatively late stage of the central decision process. In addition, the several demonstrations of Exp. XVII suggest a methodology for investigating central processes in vision in some detail. Obviously estimates of processing time would have to take into account queuing time. Also, the present demonstrations would seem to raise serious questions about models of metacontrast which reduce the phenomenon to lateral inhibitory processes (e.g., Bridgeman, 1971; Weisstein, 1968). Off-hand it would seem that more complex processes are needed to handle the interplay between nonoverlapping stimuli.

EXPERIMENT XVIII

A shorthand account of the preceding research is that when two successive stimuli compete for the services of peripheral systems, the greater energy event wins; on the other hand, when two stimuli compete for the services of the central decision process, the victor is likely to be the one that arrives second.

Consider a T stimulus of relatively low energy, so that its peripheral processing time is longer than its duration. A mask event which follows immediately on the same eye will occupy the same peripheral nets as T. Whether the peripheral nets output data on T or not is very much dependent on whether the T stimulus has more or less energy than the mask. If the T stimulus has the greater energy, then the peripheral nets will, in the main, output data on T. On the other hand, if the mask has the greater energy, then what is represented in the set, S, of central stores is primarily data on the mask.

Let us now look at the case in which T energy is greater than mask energy and the mask is PM3. The temporal variable is SOA. At brief SOA's T and PM3 will occupy common peripheral nets but since T is of greater energy, T data, rather than mask data, will be output, i.e., PM3 will fail to mask T. At longer SOA's, however, peripheral processing of T is close to completion, or is in fact completed, prior to PM3 onset. As a consequence, the central decision process now receives in succession two sets of data. The central decision process is not affected by stimulus energy, and therefore, the energy superiority of T over PM3 is no longer relevant; what does matter is which data set arrives second. In this circumstance, PM3 can now successfully mask T. In short, for the condition in which T energy > PM3 energy, masking should vary nonmonotonically with SOA and a U-function should be obtained.

Quite to the contrary is the case in which T energy < PM3 energy. At brief SOA's PM3 masks T because of its energy superiority. At longer SOA's, PM3 masks T not because of the energy difference but because data on PM3 replace, or are interwoven with, data on T in the central stores, thus distorting the central decision process. Hence, when T energy < PM3 energy, masking should be a monotonic function of SOA. Experiment XVIII tests these predictions.

Method

There were two conditions with four naive Ss receiving both. In one condition, the luminance of the T stimuli, the set of trigrams, was twice that of PM3 (Condition 2:1); in the other, the luminance of T was half that of PM3 (Condition 1:2). The luminance values were 5 ft L:2.5 ft L and 2.5 ft L:5 ft L, respectively, and both T and PM3 were exposed for 10 msec.

At each of eighteen SOA's, ranging from 0 to 184 msec, all four Ss in both conditions viewed twenty trigrams followed by PM3, with a different set of twenty trigrams given at each SOA. The number of consonants correctly identified was recorded for each trigram presentation.

In both conditions, Ss were tested in ascending order from SOA = 0 msec to SOA = 184 msec. Two Ss were given Condition 2:1 first, and two were given Condition 1:2 first. All stimuli were presented monocularly, to the right eye.

Results and Discussion

The mean number of letters correctly identified (without respect to position in the trigram) at each SOA value are shown for individual Ss in Table 9, and Figure 20 shows these mean scores averaged across Ss. The expected nonmonotonic and monotonic functions were obtained.

TABLE 9

EXP. XVIII: MEAN NUMBER OF LETTERS IDENTIFIED AS A FUNCTION OF SOA FOR TWO RATIOS OF T AND PM3

Subjects	T:PM3	SOA(msec)								
		0	4	8	12	16	20	24	32	40
S1	2:1	3.0	2.8	2.9	3.0	3.0	2.1	1.3	.7	.5
	1:2	0	0	0	0	0	.1	.1	.1	.2
S2	2:1	3.0	2.8	2.8	2.8	3.0	2.7	2.2	2.2	1.2
	1:2	0	0	0	0	0	0	.1	.1	.8
S3	2:1	2.8	2.6	2.9	3.0	2.7	2.5	2.3	2.6	2.4
	1:2	0	0	0	0	0	0	0	0	.9
S4	2:1	2.8	2.7	2.4	2.6	2.6	1.9	2.1	1.6	1.6
	1:2	0	0	0	0	0	0	.1	1.0	1.3

Subjects	T:PM3	SOA(msec)								
		48	56	64	80	96	112	136	160	184
S1	2:1	.8	.5	.7	.7	1.0	1.1	1.7	2.6	3.0
	1:2	.8	.8	.5	.9	1.5	2.4	2.1	2.7	3.0
S2	2:1	1.0	1.3	1.0	2.0	2.7	3.0	3.0	3.0	3.0
	1:2	.9	1.5	2.0	2.9	3.0	3.0	3.0	3.0	3.0
S3	2:1	1.0	1.3	1.0	2.0	2.7	3.0	3.0	3.0	3.0
	1:2	.6	1.2	1.4	2.3	2.8	3.0	3.0	3.0	3.0
S4	2:1	1.7	1.6	1.7	2.4	2.5	2.6	2.4	2.6	3.0
	1:2	2.1	1.9	2.3	2.1	2.8	2.7	2.9	3.0	3.0

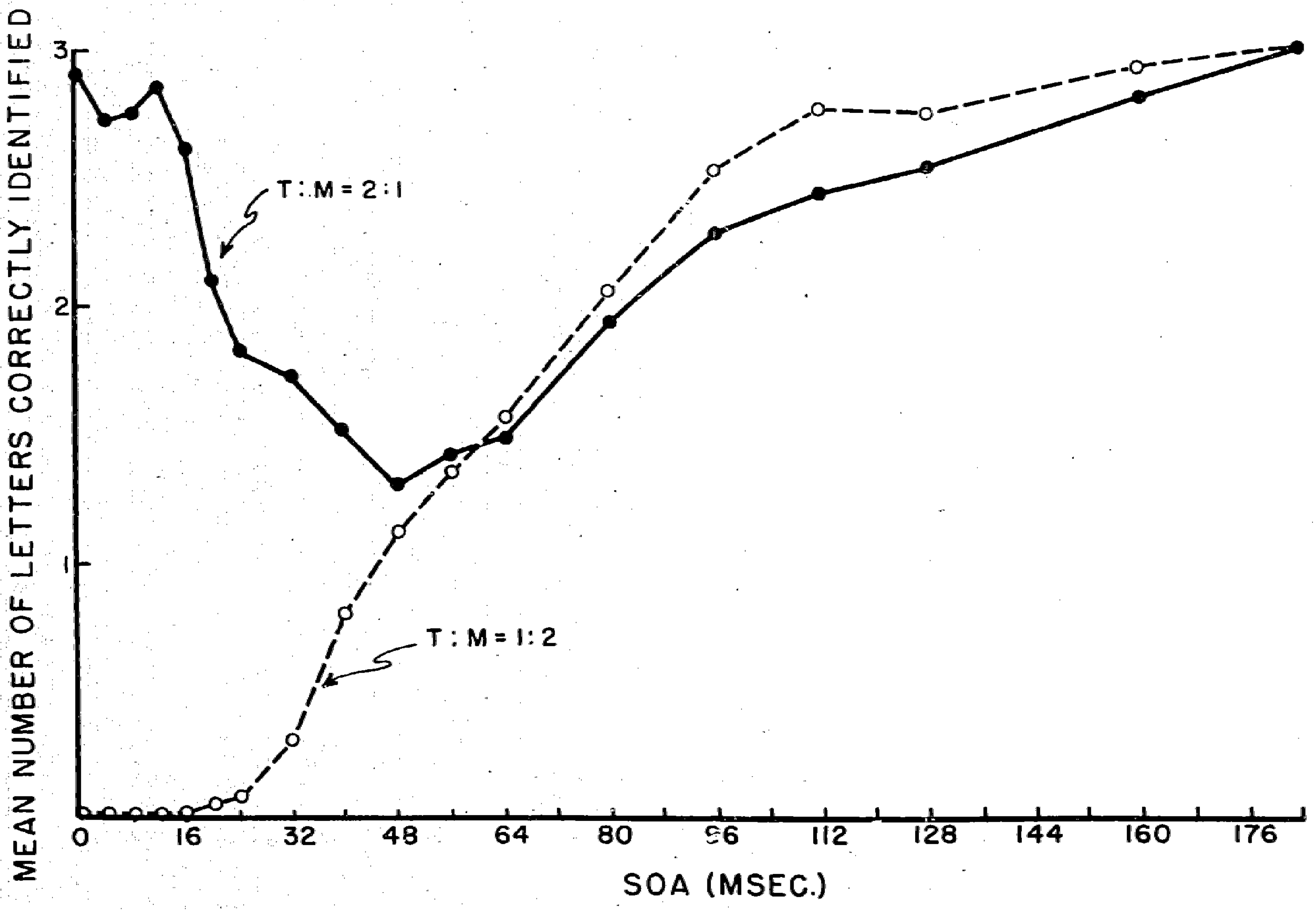


Figure 20: Relation between SOA and mean number of correct identifications for monoptic masking under two T-PM3 energy ratios in Exp. XVIII.

The minimum of the U-function in Condition 2:1 is at 48 msec. The two curves converge a little later than this in the region of 56 to 64 msec. In principle, if the T was of the same energy in both ratio conditions, the minimum of the U-function and the point of convergence with the monotonic function should be at the same SOA. The SOA value at which masking is most severe in nonmonotonic functions should vary inversely with T energy. Therefore, if the lower-energy T (10 msec x 2.5 ft L) had been pitted against a mask of 10 msec x 1.25 ft L, then the minimum of the U would have shifted to a longer SOA value and the two functions would have converged at this minimum.

U-shaped masking functions have generally been observed only under conditions of metacontrast, that is, conditions in which the contours of the mask do not overlap spatially with those of the T stimulus. The prevailing sentiment is that U-functions are unique to metacontrast paradigms (e.g., Bridgeman, 1971) and that metacontrast is therefore a very special type of visual masking. Although the present data do not necessarily refute the latter, they do show that U-functions can be obtained with T and mask overlapping, an observation buttressed by recent experiments of a very similar nature conducted by Purcell and Stewart (1970). These investigators, in accord with the present research, report U-functions with overlapping T and pattern mask when T is of greater energy than the after-coming stimulus.

The interpretation presented here for the U-function in backward masking is that it results from the differential effect of stimulus energy on masking of peripheral and central origin, coupled with the priviledged nature of a stimulus arriving centrally as the second of a pair. The question arises: is this interpretation applicable to U-functions generated by nonoverlapping T and mask?

A notable feature of metacontrast effects is that they tend to be associated with highly labile responses and are, for the most part, highly dependent on the criterion used by S. Generally the requirement for the metacontrast effect is that S uses a high criterion to determine his response (see Kahneman, 1967b, 1968; Schiller, 1969), which may be interpreted to mean that fairly complex, central processes underlie the effect (Schiller, 1969; Uttal, 1970). In addition to this idiosyncracy, there are several sources of evidence which strongly imply that the perceptual interference obtained with nonoverlapping stimuli is primarily of central origin.

First, the effect can be obtained dichoptically (Kolers, 1962; Kolers and Rosner, 1960; Weisstein and Gowney, 1969). Second, Schiller's (1969) microelectrode recordings from the lateral geniculate nucleus of the cat show that there is no physiological evidence of response depression in metacontrast-like stimulus conditions. Depression or suppression are found, either in neural response (Schiller, 1969; Fehmi, Adkins, and Lindsley, 1969) or in evoked potential (Donchin, Wicke, and Lindsley, 1963), in situations in which the two stimuli overlap in receptive fields.[7] Third, several lines of evidence suggest that metacontrast effects

---

[7] Recent experiments on visual evoked-potential correlates of sequental blanking by Andreassi et al. (Andreassi, Mayzner, Beyda, and Davidovics, 1971) are relevant both to this point and to the general thesis of the present

are not only central but, indeed, arise at a very late phase of the central decision or construction process. Werner (1935) showed that metacontrast effects rapidly diminish when the similarity between the contours of the two stimuli decrease. More recent evidence implies that the effect is most pronounced when the two stimuli, such as a form as target and two flanking forms as mask, are identical (Buchsbaum and Mayzner, 1968; Parlee, 1969; Uttal, 1970). The implication is that metacontrast masking may depend in many circumstances on the achievement of a central state approximating the identification of the form and not simply upon an interaction between contour-forming processes (see Uttal, 1970, 1971a).

On the present view, it would have to be argued that metacontrast in monoptic conditions, which follows a monotonic function with maximum masking at SOA or ISI = 0 msec, arises in part because of interference in the transmission line. The condition for monotonic masking functions is that mask energy be greater than T energy. In this situation, since no peripheral suppression can be found with nonoverlapping stimuli of equal energy, it would have to be assumed that the masking originating in the transmission line is caused by phenomena similar to those governing masking by contourless flashes. The part of the mask field not occupied by, but bounded by, the mask form (in the case of disc-ring stimuli) or forms (in the case of a flanking mask) overlaps optically the T form and is of greater energy. Therefore, we may assume that this peripheral masking is of no special type and that it may be attributed to summation or occlusion effects of the sort previously described. What is special is the central component of a monotonic metacontrast function.

When in a metacontrast paradigm T and mask are of approximately equal energy a nonmonotonic U-function is generally obtained. Since there is no peripheral perturbation possible under conditions of equal-energy and nonoverlapping stimuli, this metacontrast U-function cannot be explained by the coupling of differential peripheral and central masking effects. The entire function must be said to originate centrally. U-functions in the metacontrast paradigm have been reported for both monoptic and dichoptic conditions of presentation (see Kahneman, 1968, Weisstein, 1968).

The two nonoverlapping, equal-energy stimuli are handled, so it may be assumed, by different peripheral nets, and data on the two are cast into relatively independent subsets of the set, S, of central stores. At brief SOA's close to 0 msec, for both monoptic and dichoptic presentation, data on both stimuli are represented and a construction of both is made. What is perceived is a composite, a single-stimulus event; in the case of disc and ring as T and mask, what is perceived is a "bull's eye" (Bridgeman, 1971). With increasing SOA, the likelihood increases that all the data on the T stimulus are laid down before the data on the mask, which leads to the question: How do the later-arriving data on the nonoverlapping mask induce a distortion in the central decision on T?

---

paper. In conditions where all the stimuli are of equal intensity, Ss, while not perceiving and recognizing blanked stimuli, do give a visual evoked potential to blanked stimuli. On the other hand, when the blanking stimuli are of greater intensity than the blanked stimuli, both perceptual and evoked-potential suppression occur.

The answer that has been given to this question when the stimuli overlap is that the mask data replace data on T in the set of central stores. With nonoverlapping stimuli having the same contours, replacement may occur in the stores of peripheral systems rather like the complex cells of Hubel and Wiesel which respond to a particular stimulus property appearing in a relatively broad region of the visual field. However, in this case the new data would be identical to the old data that they replace, and a successful construction of T should still be possible. In short, the argument appropriate to overlapping stimuli having similar, but not identical, contour conditions is not especially appropriate to the metacontrast paradigm.

What is needed, perhaps, is the notion that data gathered from other, proximate locations in the visual field, or more precisely, decisions on data in other locations, enter into the decision process on the contents of a particular location. And perhaps other-location decisions converge upon the decision process in question at a relatively late stage, at least beyond the point at which the property data from the relevant subset of central stores has been used (see Figure 15).

The idea is that metacontrast effects arise at a later and perhaps functionally different stage of the central decision process than the previously discussed masking effects induced by overlapping random patterns. A functionally different stage might be one in which the decisions derived from the data in the set of central stores make contact with the structures of long-term storage to achieve identification of the form in question.[8] In any event, metacontrast is here viewed as a phenomenon arising from processes beyond the level of the concurrent-contingent model described above. And while this view pinpoints the locus of the effect, it does not, of course, suggest why the effect should be more pronounced at SOA $>$ 0 msec.

EXPERIMENT XIX

Basically, Exp. XVIII showed that a stimulus which may escape monoptic masking at some relatively brief delay of mask may suffer masking at a longer delay. This result, besides implicating the interplay between peripheral and central processes in determining the shape of masking functions, suggests caution in assigning an upper temporal limit to the masking effect of a particular pattern on a given T form. This point is put into relief by considering the data of Exp. IX given in Table 4. There it can be seen that

---

[8] A number of geometric illusions, e.g., the Ebbinghaus and Ponzo illusions, are not too dissimilar from metacontrast. In both metacontrast and the illusions, a distortion in the perception of one element is induced by surrounding, or flanking, elements. Schiller and Wiener (1962) have shown with brief, dichoptic presentation of the test and inducing elements that these illusions are of central origin. This, of course, questions the notion that these illusions arise from recurrent lateral-inhibitory processes in the retina; other lines of evidence question any form of lateral-inhibition explanation (see Coren, 1970). An alternative view is that these illusions occur because the visual system is misled into entertaining incorrect hypotheses about the test element (cf., Gregory, 1970). Presumably these hypotheses would be generated at a relatively late stage in the central decision process.

the monoptic critical T duration for various durations of PM3 was relatively brief in comparison to the dichoptic critical T duration for the same PM3 values. An interpretation of this result might say that the minimal duration of T at which a particular duration of PM3 (at ISI = 0 msec) failed to mask monoptically defines the upper boundary on the monoptic influence of that PM3 on the perception of T.

The hindsight afforded by subsequent experiments, and in particular Exp. XVIII, says that this minimal duration of T at which criterion performance has been attained does not define an upper limit for monoptic masking. Rather it simply reflects the basic peripheral rule, which is that when two stimuli occupy common peripheral nets, the more energetic one is favored. Presumably if the T stimuli were presented for this duration and an ISI were introduced between T and mask, as the ISI increased, identification accuracy should first decrease and then increase, gradually returning to the original criterion level. The reason for this see-saw effect would be the transition from masking originating peripherally to masking originating centrally as witnessed in Exp. XVIII.

Experiment XIX was conducted to verify these assertions. The procedure used was essentially that of Exp. IX grafted onto that of Exp. XVIII.

Method

Critical T duration was first determined for each of four naive Ss with the set of trigrams at 2.5 ft L followed at ISI = 0 msec by a PM3 of 10 msec x 2.5 ft L. The usual criterion of four correct in succession was used. When this critical T duration had been determined, ten ISI's were introduced between T and PM3 in 10-msec steps from 0 to 90 msec, with T duration held constant at this critical value. At each ISI S was presented with five different trigrams, and the number of letters correctly identified at each was recorded. Stimuli were presented to the right eye.

Results and Discussion

Each S's critical T duration and the mean number of letters correctly reported at each ISI are given in Table 10. Inspection of Table 10 shows that mean identification performance declined from a maximum of three letters to approximately two letters at an intermediate ISI and then recovered to the original level. The ISI means were cast into a repeated measures analysis which showed a significant difference in identification performance as a function of ISI, $F(9,27) = 4.75$, $p < .001$. In short, identification accuracy varied nonmonotonically with ISI and critical T duration was obviously not the upper limit on monoptic masking by PM3.

CONCLUDING DISCUSSION

The theoretical persuasion of this paper has been a view of perception as a temporal sequence of events involving stages of storage and transformation (Posner, 1969). Accordingly, the broad conclusions from the preceding research and speculations on these conclusions are drawn within the general context of the information-processing approach described at the outset. Figure 1 will serve as a useful reference.

TABLE 10

EXP. XIX: MEAN NUMBER OF LETTERS REPORTED AS A FUNCTION OF ISI

Subjects	Critical T Duration (msec)	ISI (msec)									
		0	10	20	30	40	50	60	70	80	90
S1	16	3.0	2.4	2.8	1.8	1.8	1.0	2.0	2.6	2.6	3.0
S2	12	3.0	2.6	2.6	2.4	3.0	2.8	2.6	3.0	3.0	3.0
S3	14	3.0	2.6	2.2	1.4	1.8	2.4	2.6	2.2	2.6	3.0
S4	16	3.0	2.6	1.6	2.0	2.2	3.0	2.8	3.0	2.8	3.0

representation of the visual input and long-term storage. On current theorizing, the contact between input and memory may be described as a feature-match or, alternatively, as an analysis-by-synthesis operation (MacKay, 1967; Neisser, 1967). In either case the question, What (kind of) object is this? is answered by first extracting certain properties or features, which in turn raises the question: What kinds of features are suitable?

It is quite unlikely that an inventory of straight lines, curved lines, verticals, horizontals, diagonals, edges, colors, etc., present in the retinal input could provide a sufficient data set for stimulus identification.[9] Rather, what would seem to be essential for recognizing a visual object or figure is the existence at some neural level of a description of the input which embodies, but does not necessarily list, all the potential relations between the parts of the object. We must suppose that knowledge of what kind of object or figure something is relies very heavily on "features," which exist only as relations among the parts. In short, the prerequisite for answering the question, What object is this? is a global representation of the input, since it is only in the context of the whole that certain "features" can be specified and that things such as lines and curves are useful to pattern-recognition devices. As Neisser (1967) observes: "In terms of information processing the whole is prior to its parts" (p.91).

Yet, paradoxically, it has to be argued on the basis of the single-cell recordings of Hubel and Wiesel (1959; 1962; 1965) that any wholistic representation must be derived originally from an inventory of features much like that described above. We may have to distinguish between two kinds of "features": those <u>detected</u> by feature-detecting systems and used to reconstitute the global character of the input and those <u>abstracted</u> from the global representation and used to recognize it. Let us call features of the first kind context-independent and features of the second kind context-dependent.

Especially relevant to the present view are the recent comments of Pollen, Lee, and Taylor (1971) on how the striate cortex reconstructs the visual world. Pollen and his colleagues intimate that the simple-cell level identified by Hubel and Wiesel (1959, 1962, 1968) cannot specify uniquely a description of the stimulus; further processing is required until "a 'reconstruction' (by which we mean the derivation of an invariant description of a visual object) has been achieved in some set of neurons" (p.74). The transformation that occurs from the level of the simple to that of the complex cell (Hubel and Wiesel, 1962) is, to their way of thinking, only a beginning in the reconstruction process; a complex cell is tuned to only one spatial frequency and only one particular angle for a restricted region of visual space. A more complete specification of the visual form, an "invariant" description, is achieved via a gathering of information from all

---

[9] Recently several authors have argued quite vigorously against the claim that feature-detectors can account for how things are recognized or why things should look as they do (Pribram, 1971; Rock, 1970; Uttal, Bunnell, and Corwin, 1970). Others have been less vigorous, but equally poignant (Gregory, 1970; Neisser, 1967).

complex cells over the involved region, and this description, they argue, may then serve as a unique determinant for further elaboration of the perception through contact with the long-term store.

In light of the foregoing, the outputs of peripheral nets, i.e., the content of the set of central stores, can be viewed in either one of two ways: either the content of the stores is a list of context-independent features or it is a wholistic representation from which relational properties are abstracted.

Let us suppose that the data in the set of central stores are context-independent features. The central decision process in the concurrent-contingent model must be, therefore, at least in part, a series of operations by which a wholistic representation of the visual input is assembled, or in the language of Pollen et al. (1971), "reconstructed." Thus, paraphrasing Neisser, preliminary decisions emphasize the global rather than the particular in the figure they construct. Assuming, therefore, two fairly broad stages in pattern recognition (Neisser, 1967), the central decision process must first establish figural unity--which is Neisser's term for wholistic representation and then make decisions on the nature of this segregated object. The first stage works with the data set provided by the peripheral nets; the second stage, with the figural unit afforded by the first. Thus in Figure 15, which illustrates the concurrent-contingent model, the figural unit is represented by the output of C_n; decision nets beyond C_n are needed for the further elaboration and classification of the stimulus.

If, on the other hand, we suppose that figural unity is represented at the level of the central stores, then the central decisions, C_1 to C_n, illustrated in Figure 15, are those which determine the proper classification of the stimulus by means of context-dependent features abstracted from the wholistic representation. On this view, however, it would be difficult to account for the relation between peripheral processing time (PPT) and central processing time (CPT) described earlier. As was noted, in conditions of relatively low T energy, the upper limit on masking (by PM, PM3) can be set by PPT, which implies that, to some extent, the peripheral and central processes overlap in time. If central decisions beyond figural unity are decisions which make use of properties abstracted from, and therefore, determined by, the context of the whole, then it is impossible on this view of peripheral net output for peripheral and central processes to occur concurrently. The latter must await the completion of the former. Thus, of the two views, the one which more easily accommodates the data is that which describes the total output of peripheral nets as a list of context-independent features from which the object form is reconstructed. We may now address the question of whether the postulated set of central stores in the concurrent-contingent model and the concept of iconic storage (Neisser, 1967) are identical.

In the main, the description of iconic, or brief, visual storage has been derived from applications of the delayed partial-sampling paradigm introduced by Sperling (1960) and Averbach and Coriell (1961). Essentially, this paradigm involves presenting simultaneously an overload of items, usually letters or digits, in a brief tachistoscopic exposure, followed after a similarly brief period of time by a probe or indicator designating which element or subset of elements S has to report. Despite the fact that the display load generally exceeds the memory span, if the indicator occurs soon enough after the display, S can give a highly accurate report of the specified element(s). As demonstrated by Sperling (1960), this delayed partial-sampling procedure shows that S has far more information available than can be reported by the memory-span, or whole-report, technique. Generally this is intepreted as meaning that the information tapped by the partial report exists in a storage medium of such brevity that the memory-span, or whole-report, technique is too slow to reveal it. The superiority of partial report over whole report declines rapidly with delay of indicator. Estimates of the decay time of iconic storage inferred from the decline in accuracy of partial reports vary from 250 msec to several seconds (Averbach and Coriell, 1961; Averbach and Sperling, 1961; Keele and Chase, 1967).

It is generally proposed that iconic memory is literal, or precategorical (Broadbent, 1971; Neisser, 1967), a proposition supported, in part, by the kinds of selection criteria which allow for efficient performance in the delayed partial-sampling task. In the original experiments of Sperling (1960), Ss were presented with an array of several rows of letters or digits. The delayed indicator specified report by row or column. Partial report at brief delays of the indicator was superior to whole report, demonstrating, perhaps, that the spatial properties of the input were available in the iconic representation. However, in one of Sperling's experiments Ss were asked to pick out letters or digits from a mixture of both. In this instance, partial report with preinstruction was not superior to whole report, suggesting that the distinction between letters and digits is not available at the level of iconic storage. Such a distinction is based on a derived property of the stimulus, and presumably the time required to categorize a particular set of physical characteristics as representing an item belonging to the class "letters" or "digits" is considerable in the medium of iconic storage. In contrast, superior partial report over whole report can be clearly demonstrated when the criterion for selection is brightness, size (Von Wright, 1968), color (Clark, 1969; Von Wright, 1968), shape (Turvey and Kravetz, 1970), or as already indicated, location (e.g., Sperling, 1960). These data demonstrate that we are able to select or ignore items in iconic storage on the basis of their general physical characteristics. We cannot, however, with the same efficiency select or ignore items on the basis of their derived properties. In terms of the distinctions recently made by Broadbent (1971), we can select efficiently on the basis of stimulus set but not on the basis of response set. All this speaks to the precategorical nature of iconic storage.

Several other lines of evidence point to a difference between iconic storage and the immediately subsequent store for categorized data, generally referred to as short-term storage or primary memory (Atkinson and Shiffrin, 1968; Broadbent, 1971). Wickelgren and Whitman (1970) have argued that unlike short-term storage, iconic memory is nonassociative. Memory for the position of the elements is by an ordered two-dimensional array of locations, not by associations between the representatives of the elements. This conclusion is buttressed by Rudov's (1966) close examination of error production in the iconic memory task. Several studies (Glucksburg and Balagura, 1965; Standing and DaPolito, 1968; Turvey, 1967) have indicated that iconic memory is not affected by repetition although repetition does significantly influence the memory of material at the level of short-term or primary storage (Hebb, 1961; Melton, 1963). In addition, experiments by Turvey (1966) and Doost and Turvey (1971) suggest that iconic storage does not require central processing capacity for its maintenance, in contrast to short-term storage which does rely on the availability of central attentive processes (see Broadbent, 1971; Posner, 1966).

There are two interpretations of the experiments which show efficient partial report under stimulus set instructions. One is that the properties of the stimulus on which stimulus set selection is based are present in the iconic store; the other is that they are not present but they can be rapidly ascertained, more rapidly, that is, than the properties which allow for a response-set selection, say, between letters and digits. Take as an example selection on the basis of size or shape. On the first view these global properties of the stimulus would be "known" at the level of iconic storage, on the second they would not. On the second view these global characteristics of the stimulus would have to be derived from a data set consisting, presumably, of context-independent features.

In theory, the content of iconic storage could be either a description of a visual object or objects, suitable for subsequent operations of pattern recognition, or a conglomerate of "crude," context-independent features which requires some further operations before it is rendered into a form suitable for classification. On the basis of perceptual reports of Ss in the delayed partial-sampling paradigm, the second of the two views of iconic content seems unlikely. Generally, Ss' descriptions imply that they see far more items than they can report (e.g., Sperling, 1960), and indeed, they may know how many items were presented although they may not know what the items were (see Eriksen and Rohrbaugh, 1970). In other words, at the level of visual information processing isolated by the delayed partial-sampling paradigm something is known about the gross form of the input, and it is the persistence of this knowledge which has been called iconic memory.

The description of what is known at the level of iconic storage, provided in the main by selection criteria which yield efficient partial report and by the perceptual reports of Ss, contrasts with the data set postulated for the central stores in the concurrent-contingent model. The argument made was that the outputs of peripheral nets are context-independent features and that it is via means of a central decision process that the visual object is "assembled" and identification of that object eventually achieved. Perceptual reports of Ss, in those situations of the present series of experiments in which masking was described by the additive rule, shifted with increasing ISI from reporting no evidence of the presence of the T letter to an intermediary state of noting its presence and finally to reporting not only that it was present but that its form was clear and that the problem was to identify it before it was replaced by PM (cf., Haber and Standing, 1968; Liss, 1968). In other words, the perceptual report which defines the iconic memory experience, that of an image in which the global characteristics are clearly defined, emerges at a relatively late stage in the process embraced by the additive rule. The conclusion we would like to draw from this is that iconic storage and the central set of stores on peripheral nets output are not identical. Iconic storage for a single item is perhaps better viewed as a storage of a decision on peripheral data as opposed to a storage of peripheral data. The decision represented at the leval of iconic storage is an intermediary decision relating to the global properties of the stimulus object; the final category state (Broadbent, 1971) has not yet been achieved at this point in the flow of visual information. For example, what is stored for an input to a certain region of the visual field is the decision that the input in this region has this size, this brightness, this color, this general shape, etc., but whether the input was the letter "F" or one's loved one is not yet known. Thus, in the central decision process, iconic storage represents an interface between decisions based on context-independent features and decisions based on context-dependent features.

Given the foregoing, the "read in" to, and "read out" from, iconic storage may be described briefly as follows. First, a set of operationally parallel, peripheral visual systems which have the retina as starting point and the cortex as end point, signal fundamental, but context-independent, properties of the stimulus at a rate which serves inversely with the energy of the stimulus up to some limiting energy value. These properties are entered asynchronously into a set of central stores by virtue of the different processing rates of the different systems. In parallel with the peripheral signalling of properties, central decisions about the stimulus based on these properties are being made. At some point, and here we can talk only vaguely, a decision is reached which corresponds to a convenient description of the stimulus from the vantage point of the subsequent categorization process (Broadbent, 1971; Neisser, 1967). This decision state can persist for a relatively prolonged period, probably because the decisions which now occur (in read-out) are based on relational features which have to be abstracted (and abstracting the "right" features may on occasion require several attempts), and probably because these subsequent decisions tax the limited capacity of the information-processing mechanisms and thus, in the face of concurrent demands, cannot always be conducted as efficiently, and as swiftly, as is ideally possible. This decision state is iconic storage, and we may conjecture in the earlier notation that when PPT $<$ CPT, the read-in to iconic storage is relatively constant for varying energy values of the stimulus.

It will be recalled that this paper began with the adoption of a particular view on two theories of masking--the integration and interruption theories. This view proposed that integration localized masking by pattern in the read-in to iconic storage and interruption placed the effect of a patterned mask on read-out; moreover, in the visual information-processing framework both theories could, indeed, be true. An extension of this view, implicit in the general discussion of Exps. I-IX, was that "integration" described masking originating peripherally while "interruption" was a more appropriate description of central masking. These notions, especially the peripheral-central one, served to guide the design and interpretation of many of the experiments reported. However, we must now emphasize what is already manifestly apparent in the reported data and the description of the concurrent-contingent model and that is that neither integration nor interruption nor both theories combined can substantially accommodate the phenomena of masking. The point to be made is, perhaps, an obvious one: there are many ways in which one stimulus may impair the perception of another.

But let us pursue for a moment, in the context of the concurrent-contingent model, the general approach of pinpointing the masking effects of a pattern prior to, and subsequent to, the attainment of the iconic representation. The read-in to iconic storage consists of a number of operations which may be affected in several ways by an after-coming event. Outputs from early stages of peripheral nets may be occluded and/or terminal outputs may be distorted. Data on both stimuli may exist in the set of central stores leading to an iconic state which would, in effect, be a mixture of both. Or data on the leading stimulus in the central stores may be replaced almost entirely, and immediately, by data on the following stimulus so that no iconic representation of the first is achieved.

A preceding mask can similarly influence the read-in by temporarily prohibiting peripheral net outputs or by mixing with the target stimulus data in the central stores, thus giving rise to a blemished iconic representation of the target. The temporal range over which impairment of the latter kind can occur is limited by virtue of the fact that data on a later-arriving stimulus event will always replace data on an earlier event in any central stores that the two have in common. Quite obviously, in this perspective the effect of a patterned mask on read-in to iconic storage could not be classified exclusively as either integration or interruption or as any simple combination of the two.

The position taken in this final section is that the decision nets illustrated in Figure 15 represent the process by which the iconic, or wholistic, representation is established. Thus, as we have noted, the output of C_n represents the iconic form. The hypothesis with which we began proposed, in part, that interruption theory spoke specifically to the effects of a mask on read-out, which here is viewed as a series of decisions. Usually the interruption theory has been interpreted as saying that an after-coming stimulus erases or replaces the icon of an earlier stimulus thus curtailing the time available for processing (Haber, 1969b; Scharf and Lefton, 1970; Spencer and Shuntich, 1970; Sperling, 1963). In the concurrent-contingent model, therefore, the notion of replacement can be translated into "a change in the decision state of C_n." Presuming that subsequent central decision nets use the C_n output as their data base, changing C_n output on a first stimulus would cut short the time available for these decisions on that stimulus. This interpretation, of course, is similar to that suggested in the clerk-customer analogy. Moreover, it implies that the minimal SOA needed to evade a central mask defines the minimal time needed for read-in and for read-out from iconic store.

But a fundamental assumption of the information-processing approach is that the flow of information on the nervous system is characterized by successive changes over time in the content of the information (Haber, 1969a). The idea that the output from each central decision net is, in essence, a new form of the stimulus information and a further step on the way to answering the question of what kind of object the stimulus is, suggests a very different view of how iconic read-out might be disturbed.

One prediction of the theory that backward masking by pattern interrupts the processing of the icon by replacing the T icon is that all pattern masks that can mask a given type of T stimulus centrally should mask over the same temporal range. Thus, for example, the minimal SOA needed by a set of T letters, say the trigrams, to evade masking by PM3 should be the same as that needed to evade, say, a mask which is either another configuration of lines different from PM3 or three letters or a word. In all these cases we must suppose that the time needed to process the T letters from the icon is constant and, therefore, that the upper limit on central masking is set by this processing time. Replacing the T icon by any one of the masking forms cited before processing is complete should yield masking, but the interval at which no masking occurs should be identical for all.

Quite on the contrary are some informal observations we have made which imply that the upper limit on central masking for a given T stimulus set depends on the form of the mask. For example following the trigrams by a trigram mask, i.e., three letters overlapping the three letters of the T stimulus, requires a longer minimal SOA for evasion of masking than following the trigram by PM3. In addition, in a variation of the situation described in Exp. XVII, a letter U followed by a pair of flanking H's gives a maximum metacontrast effect at an interval far in excess of that at which PM fails to mask the U. What these observations imply is that the minimal SOA needed by a stimulus to evade masking does not necessarily define the minimal time needed to process that stimulus; rather it defines the maximal time in which this particular mask can interfere with the processing of this particular stimulus. In other words, the time to process an item may well extend beyond the temporal interval in which a given mask can impede perception. Essentially, this theme is expressed in the contrast between RN and PM.

We may therefore entertain an alternative to the icon-replacement notion. To begin with perception is, as we have noted earlier, a sequence of operations in time in which the iconic stage, we might now add, is a convenient point to introduce a delay if such a need arises (see Posner, 1963). But in most circumstances perception proceeds uninterrupted with the output of each decision net representing a further gain in knowledge about the stimulus. Masking arises subsequent to the C_n output, i.e., post-iconically, not because of icon-replacement, although that may occasionally be true, but because discovering what kind of object the mask is may require the services of decision nets beyond C_n which are presently engaged in discovering what kind of object the target is. The implication of this view is that the more similarities between the target and mask (and this similarity is not restricted to the physical dimensions), the greater the opportunity for masking and the greater the temporal range over which masking may occur.

In this respect a most instructive observation on sequential blanking or masking has been made by Mayzner and Tresselt (1970): if a non-word mask of five letters follows a non-word, five-letter target, masking occurs; on the other hand, masking does not occur if the non-word mask follows a five-letter word. This means, perhaps, that semantic similarity as well as geometric similarity may be grounds for central masking (cf., Uttal, 1971b).[10]

---

[10] Of even greater relevance to this point are unpublished experiments by Jacobsen cited recently by Coltheart (1972). These experiments show that if the mask is a word which is an associate of the target word (e.g., mouse-cheese), the interval over which masking is obtained is shorter than if the mask is not an associate (e.g., mouse-green). This finding suggests that central decisions on an earlier event may be facilitated, rather than hindered, by a subsequent semantically similar event.

SUMMARY

A series of experiments was conducted which explored visual masking of peripheral and central origin through the use of mask stimuli which masked either both monoptically and dichoptically or only monoptically. The major observations are summarized below.

(1) Backward masking of peripheral origin was characterized by a multiplicative rule relating the energy of the target stimulus to the minimal interstimulus interval needed to evade masking; thus, target energy x minimal interstimulus interval = a constant.

(2) Backward masking of central origin was characterized by an additive rule relating the duration of the target stimulus to the minimal interstimulus interval needed to evade masking: target duration + minimal interstimulus interval = a constant. This complimentarity between target duration and interstimulus interval implicates onset-onset time as the relevant temporal variable in central masking.

(3) While energy variables significantly affected the degree and direction of peripheral masking, they were relatively immaterial to masking arising centrally.

(4) Forward masking of peripheral origin was more pronounced than backward masking of peripheral origin; moreover, the severity of peripheral forward masking increased with increases in mask intensity, the severity of peripheral backward masking did not. Peripheral forward masking, like peripheral backward masking, was characterized by the multiplicative rule.

(5) In comparison to central backward masking, central forward masking was relatively weak and did not appear to obey the additive rule. In addition, a central forward masking effect was observed which delayed, rather than impaired, target stimulus perception.

(6) When two stimuli, target and mask, were presented monoptically in a backward masking arrangement, the upper limit on masking was set by either peripheral or central processes depending on the energy of the target and the relation between the target and mask patterns.

(7) A nonmonotonic U-function was obtained monoptically with overlapping target and mask, where target energy was greater than mask energy. The function reflects the transition from peripheral to central masking with increasing delay between the two stimuli.

(8) Individual differences were manifestly greater in central than in peripheral masking.[11]

---

[11] Two observations by Schiller (Schiller, 1965; Schiller and Wiener, 1963) speak to this point: monoptic and dichoptic masking by pattern declines with practice, more so for dichoptic than monoptic presentation, but practice does not significantly influence masking by a homogeneous flash. Both of these results would be expected on the principle that dichoptic masking by a light flash reflects disturbances in peripheral nets. The central process should be more susceptible to practice

(9) "Disinhibition" or "recovery of target" effects were observed which could not be easily accommodated by lateral-inhibition explanations.

(10) Peripheral and central processes, symbolized respectively by the multiplicative and additive rules, do not function in a sequential and additive fashion. Rather, the relation between the two is that they overlap in time, with the central processes contingent on the outputs of the peripheral processes. A model was developed which expressed this concurrent-contingent relation and rationalized the data of the present series of experiments.

REFERENCES

Allport, D.A. (1968) The rate of assimilation of visual information. Psychonom. Sci. 12, 231-232.

Andreassi, J.L., M.S. Mayzner, D.R. Beyda, and S. Davidovics. (1971) Visual cortical evoked potentials under conditions of sequential blanking. Percept. Psychophys. 10, 164-168.

Atkinson, R.C. and R.M. Shiffrin. (1968) Human memory: A proposed system and its control processes. In Psychology of Learning and Motivation: Advances in Research and Theory, K.W. Spence and J.T. Spence, eds. Vol. 2 (New York: Academic Press).

Averbach, E. and A.S. Coriell. (1961) Short-term memory in vision. Bell Syst. tech. J. 40, 309-328.

Averbach, E. and G. Sperling. (1961) Short-term storage of information in vision. In Symposium on Information Theory, C. Cherry, ed. (London: Butterworth).

Battersby, W.S. and I.H. Wagman. (1962) Neural limitations of visual excitability. IV: Spatial determinants of retrochiasmal interaction. Amer. J. Physiol. 203, 359-365.

Baxt, N. (1871) Ueber d. Zeit Welche nötig ist, damit ein Gesichtseindruck zum Bewusstsein kommt. Pflügers Arch. ges. Psychol. 4, 325-336.

Beller, H.K. (1968) Stages of processing in visual search. Unpublished doctoral dissertation, Brandeis University.

Bishop, P.O., W. Burke, and R. Davis. (1959) Activation of single lateral geniculate cells by stimulation of either optic nerve. Science 130, 506-507.

Boynton, R.M. (1961) Some temporal factors in vision. In Sensory Communication, W.A. Rosenblith, ed. (New York: Wiley).

Bridgeman, B. (1971) Metacontrast and lateral inhibition. Psychol. Rev. 78, 528-539.

Broadbent, D.E. (1971) Decision and Stress. (London: Academic Press).

Buchsbaum, W.H. and M.S. Mayzner. (1968) The effects of line length on sequential blanking. Psychonom. Sci. 15, 111-112.

Cheatham, P.G. (1952) Visual perceptual latency as a function of stimulus brightness and contour shape. J. exp. Psychol. 43, 369-380.

Chung, S-H. (1968) Neurophysiology of the visual system. In Recognizing Patterns, P.A. Kolers and M. Eden, eds. (Boston: M.I.T.).

Clark, S.E. (1969) Retrieval of color information from the preperceptual storage system. J. exp. Psychol. 82, 263-266.

Coren, S. (1970) Lateral inhibition and the Wundt-Herring illusion. Psychonom. Sci. 18, 341.

(experience) and a more significant source of individual differences. There is also some evidence to suggest that individual differences in masking may be quite stable (Dember and Neiberg, 1966).

Coltheart, M. (1972) Visual information processing. In Psychology 1972, P.C. Dodwell, ed. (Baltimore: Penguin).

Crawford, B.H. (1947) Visual adaptation in relation to brief conditioning stimuli. Proceedings of the Royal Society of London, Series B, 134, 283-302.

Davy, E. (1952) The intensity-time relation for multiple flashes of light in the peripheral retina. J. opt. Soc. Amer. 42, 937-941.

Dember, W.N. and A. Neiberg. (1966) Individual differences in susceptibility to visual backward masking. Psychonom. Sci. 6, 49-50.

Dember, W.N. and D.G. Purcell. (1967) Recovery of masked visual targets by inhibition of the masking stimulus. Science 157, 1335-1336.

Donchin, E., J.D. Wicke, and D.B. Lindsley. (1963) Cortical evoked potentials and perception of paired flashes. Science 41, 1285-1286.

Doost, R. and M.T. Turvey (1971) Iconic memory and central processing capacity. Percept. Psychophys. 9, 269-274.

Ellis, S.H. and W.G. Chase. (1971) Parallel processing in item recognition. Percept. Psychophys. 10, 379-384.

Eriksen, C.W. (1966) Temporal luminance summation effects in backward and forward masking. Percept. Psychophys. 1, 87-92.

Eriksen, C.W. and J.F. Collins. (1965) A reinterpretation of one form of backward and forward masking in visual perception. J. exp. Psychol. 70, 343-351.

Eriksen, C.W. and J.F. Collins. (1967) Some temporal characteristics of visual pattern reception. J. exp. Psychol. 74, 476-484.

Eriksen, C.W. and M. Hoffman. (1963) Form recognition at brief duration as a function of adapting field and interval between stimulations. J. exp. Psychol. 66, 485-499.

Eriksen, C.W. and J.S. Lappin. (1964) Luminance summation-contrast reduction as a basis for certain forward and backward masking effects. Psychonom. Sci. 1, 313-314.

Eriksen, C.W. and P.H. Marshall. (1969) Failure to replicate a reported U-shaped visual masking function. Psychonom. Sci. 15, 195-196.

Eriksen, C.W. and J.W. Rohrbaugh. (1970) Some factors determining efficiency of selective attention. Amer. J. Psychol. 83, 330-342.

Fehmi, L.G., J.W. Adkins, and D.B. Lindsley. (1969) Electrophysiological correlates of visual perceptual masking in monkeys. Exp. Brain Res. 7 229-316.

Fehrer, E. and I. Biederman. (1962) A comparison of reaction time and verbal report in detection of masked stimuli. J. exp. Psychol. 64, 126-130.

Fehrer, E. and D. Raab. (1962) Reaction time to stimuli masked by metacontrast. J. exp. Psychol. 63, 143-147.

Fillenz, M. (1961) Binocular interaction in the lateral geniculate body of the cat. In The Visual System: Neurophysiology and Psychophysics, R. Jung and H. Kornhuber, eds. (Berlin: Springer).

Fitzgerald, R.E. and R. Kirkham. (1966) Backward visual masking as a function of average uncertainty of the masking pattern. Psychonom. Sci. 6, 469-470.

Fry, G.A. (1934) Depression of the activity aroused by a flash of light by applying a second flash immediately afterwards to adjacent areas of the retina. Amer. J. Physiol. 108, 701-707.

Glucksberg, S. and S. Balagura. (1965) Effects of repetition and intra-array similarity upon very short-term visual memory. Paper presented at the meeting of the Psychonomic Society, Chicago.

Greenspoon, T.S. and C.S. Eriksen. (1968) Interocular non-independence. Percept. Psychophys. 3, 93-96.

Gregory, R.L. (1970) *The Intelligent Eye*. (New York: McGraw-Hill).

Haber, R.N. (1969a) Information processing analyses of visual perception: An introduction. In *Information Processing Approaches to Visual Perception*, R.N. Haber, ed. (New York: Holt, Rinehart and Winston).

Haber, R.N. (1969b) Repetition, visual persistence, visual noise and information processing. In *Information Processing in the Nervous System*, K.N. Leibovic, ed. (New York: Springer-Verlag).

Haber, R.N. and L.S. Nathanson. (1969) Processing of sequentially presented letters. Percept. Psychophys. 5, 359-361.

Haber, R.N. and L.G. Standing. (1968) Clarity and recognition of masked and degraded stimuli. Psychonom. Sci. 13, 83-84.

Hebb, D.O. (1961) Distinctive features of learning in the higher animal. In *Brain Mechanisms and Learning*, J.F. Delafresnaye, ed. (New York: Oxford University Press).

Held, R. and S.R. Shattuck. (1971) Color-and-edge-sensitive channels in the human visual system: Tuning for orientation. Science 174, 314-316.

Houlihan, K. and R.W. Sekuler. (1968) Contour interactions in visual masking. J. exp. Psychol. 77, 281-285.

Hubel, D.H. and T.N. Wiesel. (1959) Receptive fields of single neurons in the cat's striate cortex. J. Physiol. 148, 574-591.

Hubel, D.H. and T.N. Wiesel. (1961) Integrative action in the cat's lateral geniculate body. J. Physiol. 155, 385-398.

Hubel, D.H. and T.N. Wiesel. (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. J. Physiol. 160, 106-154.

Hubel, D.H. and T.N. Wiesel. (1965) Receptive fields and functional architecture in two nonstriate visual areas (18 & 19) of the cat. J. Neurophysiol. 28, 229-289.

Hubel, D.H. and T.N. Wiesel. (1968) Receptive and functional architecture of monkey striate cortex. J. Physiol. 195, 215-243.

Jung, R. (1961) Neuronal integration in the visual cortex and its significance for visual information. In *Sensory Communication*, W.A. Rosenblith, ed. (New York: Wiley).

Kahneman, D. (1967a) Temporal effects in the perception of light and form. In *Symposium on Models for the Perception of Speech and Visual Form*, W. Wathen-Dunn, ed. (Cambridge: M.I.T. Press).

Kahneman, D. (1967b) An onset-onset law for one case of apparent motion and metacontrast. Percept. Psychophys. 2, 577-584.

Kahneman, D. (1968) Method, findings and theory in studies of visual masking. Psychol. Bull. 70, 404-426.

Keele, S.W. and W.G. Chase. (1967) Short-term visual storage. Percept. Psychophys. 2, 383-386.

Kietzman, M.L., R.C. Boyle, and D.B. Lindsley. (1971) Perceptual masking: Peripheral vs. central factors. Percept. Psychophys. 9, 350-351.

Kinsbourne, M. and E.K. Warrington. (1962a) The effect of an aftercoming random pattern on the perception of brief visual stimuli. Quart. J. exp. Psychol. 14, 223-224.

Kinsbourne, M. and E.K. Warrington. (1962b) Further studies on the masking of brief visual stimuli by a random pattern. Quart. J. exp. Psychol. 14, 235-245.

Kolers, P.A. (1962) Intensity and contour effects in visual masking. Vision Res. 2, 277-294.

Kolers, P.A. (1967) Comments on the session on visual recognition. In Models for the Perception of Speech and Visual Form, W. Wathen-Dunn, ed. (Cambridge: M.I.T. Press).

Kolers, P.A. (1968) Some psychological aspects of pattern recognition. In Recognizing Patterns, P.A. Kolers and M. Eden, eds. (Cambridge: M.I.T. Press).

Kolers, P.A. and B.S. Rosner. (1960) On visual masking (metacontrast): Dichoptic observation. Amer. J. Psychol. 73, 2-21.

Laner, S., P. Morris, and R.C. Oldfield. (1957) A random pattern screen. Quart. J. exp. Psychol. 9, 105-108.

Leibovic, K.N. (1968) Observations on some experiments on visual masking. J. gen. Psychol. 78, 19-26.

Lindsley, D.F., K.L. Chow, and M. Gollender. (1967) Dichoptic interactions of lateral geniculate neurons of cats to contralateral and ipsilateral eye stimulation. J. Neurophysiol. 30, 628-644.

Liss, P. (1968) Does backward masking by visual noise stop stimulus processing? Percept. Psychophys. 4, 328-330.

MacKay, D.M. (1967) Ways of looking at perception. In Models for the Perception of Speech and Visual Form, W. Wathen-Dunn, ed. (Cambridge: M.I.T. Press).

Mayzner, M.S. and M.E. Tresselt. (1970) Visual information processing with sequential inputs: A general model for sequential blanking, displacement and over-printing phenomena. Ann. New York Acad. Scis. 169, 599-618.

McCollough, C. (1965) Color adaptation of edge-detectors in the human visual system. Science 149, 1115-1116.

Melton, A.W. (1963) Implications of short-term memory for a general theory of memory. J. verb. Learn. verb. Behav. 2, 1-21.

Mewhort, D.J.K., P.M. Merikle, and M.P. Bryden. (1969) On the transfer from iconic to short-term memory. J. exp. Psychol. 81, 89-94.

Monnier, M. (1952) Retinal, cortical and motor responses to photic stimulation in man. J. Neurophysiol. 15, 469-486.

Mowbray, G.H., and L.B. Durr. (1964) Visual masking. Nature 201, 277-278.

Nakayama, K. (1968) Analysis of forward and backward masking in terms of single unit responses in the cat visual system. Proc. 76th Annual Convention Amer. Psychol. Assoc. 3, 317-318.

Neisser, U. (1967) Cognitive Psychology. (New York: Appleton-Century-Crofts).

Pantle, A. and R.W. Sekuler. (1968) Size detecting mechanisms in human vision. Science 162, 1146-1148.

Parlee, M.B. (1969) Visual backward masking of a single line by a single line. Vision Res. 9, 199-205.

Pollen, D.A., J.R. Lee, and H. Taylor. (1971) How does the striate cortex begin the reconstruction of the visual world? Sicence 173, 74-77.

Posner, M.I. (1963) Immediate memory in sequential tasks. Psychol. Bull. 60, 333-349.

Posner, M.I. (1966) Components of skilled performance. Science 152, 1712-1718.

Posner, M.I. (1969) Perception and cognition as processing models. In Approaches to Thought, J.F. Voss, ed. (Columbus: Merrill).

Pribram, K.H. (1971) Languages of the Brain: Experimental Paradoxes and Principles in Neuropsychology. (Englewood Cliffs: Prentice Hall).

Purcell, D.G. and W.N. Dember. (1968) The relation of phenomenal brightness reversal and re-reversal to backward masking and recovery. Percept. Psychophys. 3, 290-292.

Purcell, D.G. and A.L. Stewart. (1970) U-shaped backward masking functions with nonmetacontrast paradigms. Psychonom. Sci. 21, 361-363.

Purcell, D.G., A.L. Stewart, and W.N. Dember. (1968) Spatial effectiveness of the mask: Lateral inhibition in visual backward masking. Percept. Psychophys. 4, 344-346.

Purcell, D.G., A.L. Stewart, and W.N. Dember. (1969) Backward masking: Facilitation through increased target-field luminance and duration. Psychonom. Sci. 15, 87-88.

Raab, D.H. (1963) Backward masking. Psychol. Bull. 60, 118-129.

Riggs, L.A., J.C. Armington, and F. Ratliff. (1954) Motions of the retinal image during fixation. J. opt. Soc. Amer. 44, 315-321.

Robinson, D.N. (1966) Disinhibition of visually masked stimuli. Science 154, 157-158.

Robinson, D.N. (1968) Visual disinhibition with binocular and interocular presentations. J. opt. Soc. Amer. 58, 1146-1150.

Rock, I. (1970) Perception from the standpoint of psychology. Perception and its disorders. (Research Publication of the Association for Research on Nervous and Mental Disease) Vol. XLVII.

Rudov, M. (1966) Dimensionality of human information storage. J. exp. Psychol. 71, 273-281.

Schiller, P.H. (1965) Monoptic and dichoptic visual masking by patterns and flashes. J. exp. Psychol. 69, 193-199.

Schiller, P.H. (1966) Forward and backward masking as a function of relative overlap and intensity of test and masking stimuli. Percept. Psychophys. 1, 161-164.

Schiller, P.H. (1968) Single unit analysis of backward visual masking and metacontrast in the cat lateral geniculate nucleus. Vision Res. 8 855-866.

Schiller, P.H. (1969) Behavioral and electrophysiological studies of visual masking. In Information Processing in the Nervous System, K.N. Leibovic, ed. (New York: Springer-Verlag).

Schiller, P.H. and A. Greenfield. (1969) Visual masking and the recovery phenomenon. Percept. Psychophys. 6, 182-184.

Schiller, P.H. and M.C. Smith. (1965) A comparison of forward and backward masking. Psychonom. Sci. 3, 77-78.

Schiller, P.H. and M. Wiener. (1962) Binocular and stereoscopic viewing of geometric illusions. Percept. mot. Skills 15, 739-747.

Schiller, P.H. and M. Wiener. (1963) Monoptic and dichoptic visual masking. J. exp. Psychol. 66, 386-393.

Sekuler, R.W. (1965) Spatial and temporal determinants of visual backward masking. J. exp. Psychol. 70, 401-406.

Sekuler, R.W., E.L. Rubin, and W.H. Cushman. (1968) Selectivities of human visual mechanisms for direction of movement and contour orientation. J. opt. Soc. Amer. 58, 1146-1150.

Sharf, B. and L.A. Lefton. (1970) Backward and forward masking as a function of stimulus and task parameters. J. exp. Psychol. 84, 331-338.

Smith, M.C. and P.H. Schiller. (1966) Forward and backward masking: A comparison. Canad. J. Psychol. 20, 337-342.

Spencer, T.J. and R. Shuntich. (1970) Evidence for an interruption theory of backward masking. J. exp. Psychol. 85, 198-203.

Sperling, G. (1960) The information available in brief visual presentations. Psychol. Monogr. 74, (whole No. 498).

Sperling, G. (1963) A model for visual memory tasks. Human Factors 5, 19-31.

Standing, L. and F. DaPolito. (1968) Limitations of the repetition effect revealed by partial report. Psychonom. Sci. 13, 297-299.

Sternberg, S. (1967) Two operations in character recognition: Some evidence from reaction time measures. Percept. Psychophys. 2, 45-53.

Sternberg, S. (1969) Memory-scanning: Mental processes revealed by reaction-time experiments. Amer. Scientist 57, 421-457.

Stigler, R. (1910) Chronophotische Studien über den Umgebungskontrast. Pflügers Arch. ges. Physiol. 134, 365-435.

Sturr, J.R. and W.S. Battersby. (1966) Neural limitations of visual excitability. VII: Binocular convergence in cat geniculate and cortex. Vision Res. 6, 401-419.

Thomas, J.P. (1970) Model of the function of receptive fields in human vision. Psychol. Rev. 77, 121-134.

Thomas, J.P. and L.G. Kerr. (1971) Evidence of role of size-tuned mechanisms in increment threshold task. Vision Res. 11, 647-655.

Thompson, J.H. (1966) What happens to the stimulus in backward masking? J. exp. Psychol. 71, 580-586.

Turvey, M.T. (1966) The effects of rehearsing analyzed information upon the retrieval of unanalyzed information. Psychonom. Sci. 6, 365-366.

Turvey, M.T. (1967) Repetition and the perceptual information store. J. exp. Psychol. 74, 289-293.

Turvey, M.T. and S. Kravetz. (1970) Retrieval from iconic memory with shape as the selection criterion. Percept. Psychophys. 8, 171-172.

Uttal, W.R. (1970) On the physiological basis of masking with dotted visual noise. Percept. Psychophys. 7, 321-327.

Uttal, W.R. (1971a) The effect of interval and number on masking with dot bursts. Percept. Psychophys. 9, 469-473.

Uttal, W.R. (1971b) The psychobiological silly season-or-what happens when neurophysiological data become psychological theories. J. gen. Psychol. 84, 151-166.

Uttal, W.R., L.M. Bunnell, and S. Corwin. (1970) On the detectability of straight lines in visual noise: An extension of French's paradigm into the millisecond domain. Percept. Psychophys. 8, 385-388.

von Békésy, G. (1971) Auditory backward inhibition in concert halls. Science 171, 529-536.

Von Wright, J.M. (1968) Selection in immediate memory. Quart. J. exp. Psychol. 20, 62-68.

Weisstein, N.A. (1968) A Rashevsky-Landahl neural net: Simulation of metacontrast. Psychol. Rev. 75, 494-522.

Weisstein, N.A. (1969) What the frog's eye tells the human brain: Single cell analyzers in the human visual system. Psychol. Bull. 72, 157-176.

Weisstein, N.A. (1971) W-shaped and U-shaped functions obtained for monoptic and dichoptic disk-disk masking. Percept. Psychophys. 9, 275-278.

Weisstein, N.A. and R.L. Gowney. (1969) Apparent movement and metacontrast: A note on Kahneman's formulation. Percept. Psychophys. 5, 321-328.

Werner, H. (1935) Studies on contour. I: Qualitative analyses. Amer. J. Psychol. 47, 40-64.

White, M.J. (1969) Laterality differences in perception: A review. Psychol. Bull. 72, 387-405.

Wickelgren, W.A. and P. Whitman. (1970) Visual very short-term memory is nonassociative. J. exp. Psychol. 84, 277-281.

Separate Speech and Nonspeech Processing in Dichotic Listening?*

Ruth S. Day+ and James C. Bartlett+
Haskins Laboratories, New Haven

ABSTRACT

Temporal order judgment (TOJ) in dichotic listening can be a difficult task. Previous experiments that used two speech stimuli on each trial (S/S) obtained sizeable error rates when subjects were required to report which ear led (TOJ-by-ear). When subjects were required to _identify_ the leading stimulus (TOJ-by-stimulus), the error rate increased substantially. Apparently, the two speech stimuli were competing for analysis by the same processor and so were overloading it. The present experiment used the same TOJ tasks but presented a speech and a nonspeech stimulus on each trial (S/NS). The error rate was comparable to that of S/S for TOJ-by-ear but did not increase for TOJ-by-stimulus. This would be expected if the speech and nonspeech stimuli are being sent to different processors, each of which performs its analysis without interference from the other. The interpretation of the data given here is consistent with the results of standard identification experiments reported elsewhere: when asked to identify both stimuli on each dichotic trial, subjects made many errors on S/S, while performance was virtually error free on S/NS.

Dichotic listening is presumably a task that creates a situation of information overload. Let us examine some of the data that support this notion.

When both stimuli are speech (S/S). Consider cases where a different speech stimulus is presented to each ear. When subjects are asked to report both stimuli, do errors occur? The answer is yes. Although there is a wide range of performance levels reported in the literature, significant error rates are obtained. One explanation for these results is that both speech stimuli are sent to a single processor. This processor cannot fully analyze two stimuli at the same time, hence errors occur.

When both stimuli are nonspeech (NS/NS). Consider cases where a different nonspeech stimulus is presented to each ear. Again, although overall performance levels vary in the literature, significant error rates are

---

*Paper presented at the 82nd meeting of the Acoustical Society of America, Denver, Colorado, October 1971.

+Also Yale University, New Haven.

obtained. An explanation complementary to the one given above for the S/S cases is that both nonspeech messages are sent to a single processor. This processor cannot fully handle two stimuli at the same time, and hence errors occur.[1]

When one stimulus is speech and the other is nonspeech (S/NS). Recently we asked a very simple question: what happens when we present speech to one ear and nonspeech to the other ear (Day and Cutting, 1971a)? Will errors occur in this "mixed" S/NS situation? The answer, somewhat surprisingly, is no. Thus, given a consonant-vowel syllable to one ear and a tone to the other, subjects are readily able to identify both stimuli. It makes no difference which ear receives the speech stimulus and which the nonspeech stimulus. It appears, then, that each stimulus is sent to a different processor. Each processor can do its work without competition from the other. Figure 1 summarizes the results and explanations for S/S, NS/NS, and S/NS identification tasks.

Elsewhere (Day and Cutting, 1971b) we have re-examined what it means to say that dichotic listening yields a situation of information overload resulting in "perceptual competition." We have argued that perceptual competition in these standard identification tasks occurs only when both stimuli are from the same broad class of events, that is, when both are speech (S/S) or both are nonspeech (NS/NS). Perceptual competition does not occur when one stimulus is speech and the other is nonspeech (S/NS).

The present paper examines a different kind of dichotic listening task: temporal order judgment (TOJ) tasks. Figure 2 illustrates the general manner in which we conduct these studies. The relative onset time of the members of a dichotic pair are varied over trials, as shown in the top part of the display. On some trials, stimulus A and stimulus B begin at the same point in time; thus, there is zero relative onset time between the two stimuli. On other trials, stimulus A precedes stimulus B by a short interval, for example, 25, 50, or 150 msec. There are also trials where stimulus B precedes stimulus A by these same intervals.

Two types of TOJ tasks are shown in the bottom portion of Figure 2. In the TOJ-by-stimulus task, the subject is asked to report which stimulus led. In terms of the schematic diagram, he would report either stimulus A or stimulus B. In the TOJ-by-ear task, we present the same stimuli but ask a different question. The subject is asked to report which ear led. He need not perform linguistic analysis and identify the leading stimulus; all he needs to do is determine which ear was the first to receive stimulation.

We are interested in comparing overall performance levels for the two TOJ tasks. First, consider what happens when both stimuli are speech. The data shown in Figure 3 have been pooled over several experiments reported elsewhere (Day, 1970, forthcoming; Day and Cutting, 1970, Day and Copeland, forthcoming-a). Stimuli were of the general form BANKET/LANKET. When the

---

[1] Note that we are discussing overall performance levels here and have put aside the whole question of ear-hemisphere advantages.

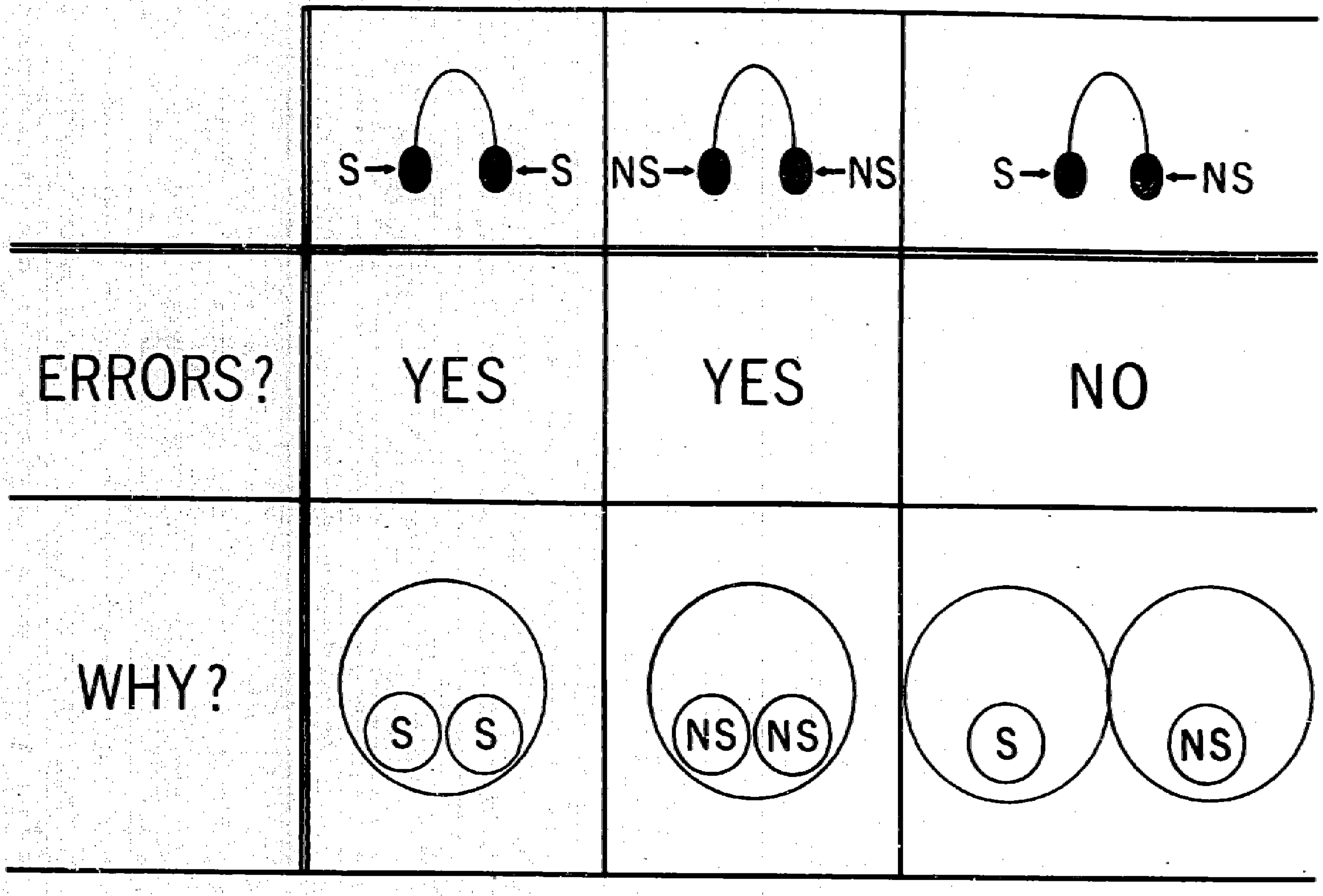
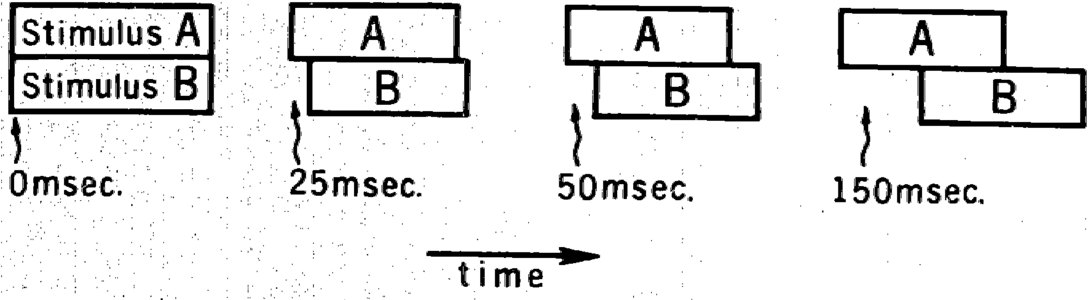


Figure 1: Results and explanations for dichotic identification tasks where both stimuli are speech (S/S), both are nonspeech (NS/NS), or one is speech and the other nonspeech (S/NS).

LEAD TIME CONDITION



TEMPORAL ORDER JUDGMENT TASKS

		Responses
1. TOJ-by-stimulus	"Which stimulus led?"	A, B
2. TOJ-by-ear	"Which ear led?"	Left, Right

Figure 2: General paradigm for temporal order judgment (TOJ) tasks.

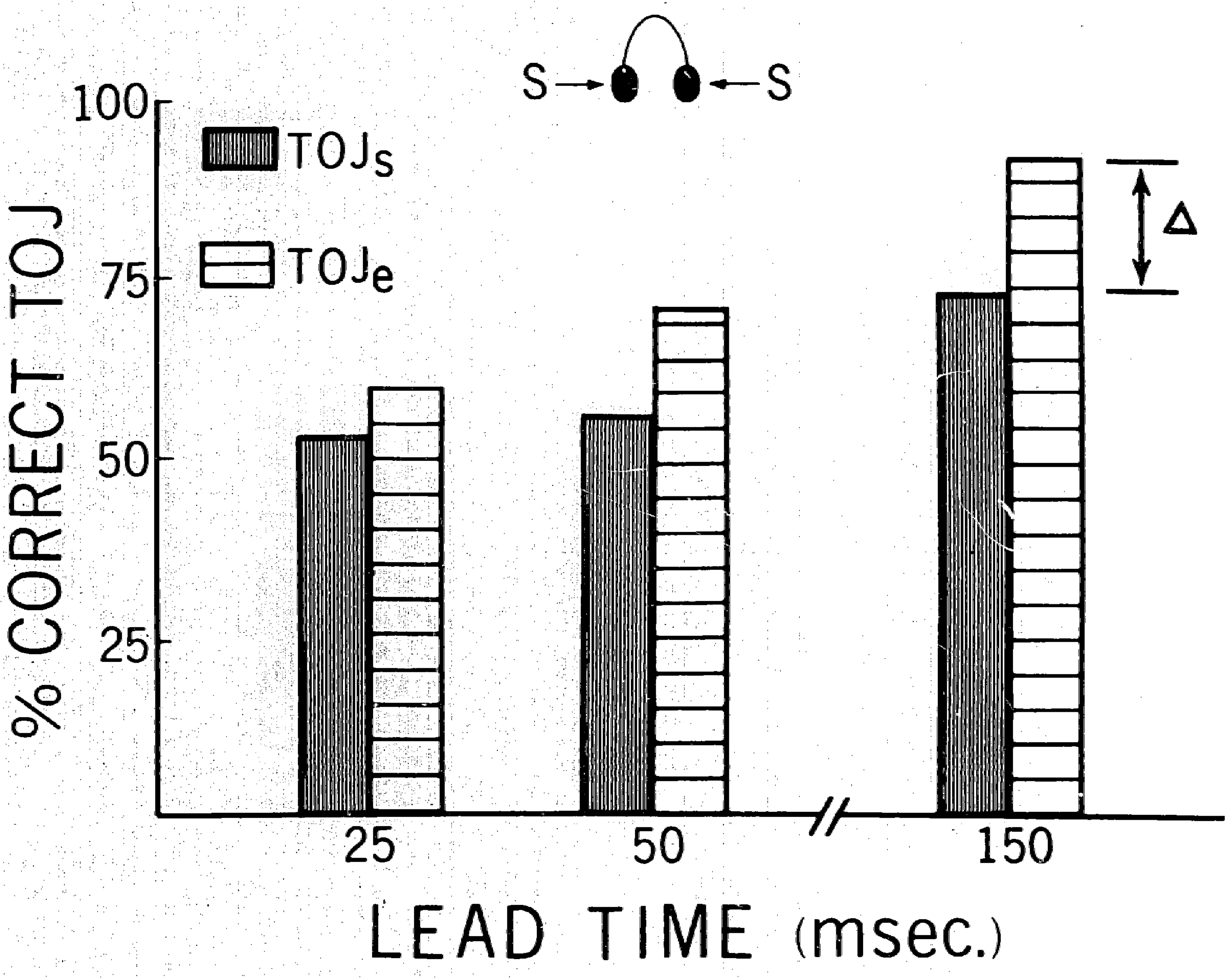


Figure 3: Percent correct for the two TOJ tasks when both stimuli are speech.

relative onset time of the two inputs was 25 msec, overall performance was 53% correct when subjects had to report which stimulus led (TOJ_s). For this same lead condition and identical stimulus pairs, performance was 60% correct when these subjects had to report which ear led (TOJ_e). Thus there was an improvement of 7% when subjects reported ear-of-lead rather than stimulus-of-lead. This improvement occurred at all lead conditions, with a 15% difference between the two tasks in the 50-msec condition and a 19% difference in the 150-msec condition. In addition to this task effect, there was a lead-time effect: subjects were better able to judge temporal order as the relative onset time between the two inputs increased. However, we are primarily interested in the difference between the two tasks as shown by the Δ on the right side of Figure 3. Recently the TOJ_e vs. TOJ_s comparison was made in a highly simplified situation (Day and Copeland, forthcoming-b). The same pair of consonant-vowel syllables (/bæ/ + /dæ/) was presented on every trial, with lead times of ±50 msec. Again, TOJ_e yielded superior performance.

What happens when we perform the same experiment but use "mixed" S/NS trials? The present experiment used the syllables /ba, da, ga/ as speech stimuli and 500-, 700-, and 1000-Hz tones as nonspeech stimuli. The stimuli were those used in the first S/NS study (Day and Cutting, 1971a). That paper describes the stimuli and the tape preparation. Briefly, all stimuli were 300 msec long, and the amplitude envelopes of the tone stimuli were matched to resemble those of the speech stimuli. Tapes were prepared on the pulse code modulation system at the Haskins Laboratories (Cooper and Mattingly, 1969) which insures an accuracy of ±1/2 msec in specifying relative onset time.

There were twelve subjects. All were right handed, had no history of hearing trouble, and were native American English speakers. A representative trial consisted of /ba/ to one ear and a low tone to the other. On the TOJ_s task, the subject would report either "ba" or "low," while on the TOJ_e task, he would report either "left" or "right."

Figure 4 summarizes the results of the experiment. Again, performance did improve as the relative onset time increased. However, there was no task difference. Thus the Δ-value was 3%, 4%, and 2% for the 25-, 50-, and 150-msec conditions, respectively. None of these differences is statistically significant.

In order to visualize the contrast between the previous S/S experiments and the present S/NS experiments, let us compare Δ-scores across experiment types. Figure 5 plots Δ-scores along the ordinate, representing a subtraction of the TOJ_s scores from the TOJ_e scores. For the previous S/S type of experiment, there were large task differences for all lead conditions. Also, the magnitude of this difference increased across the lead continuum. In the present S/NS type of experiment, there were small, nonsignificant task differences, and Δ did not change across the lead continuum.

These data can be viewed in still another manner. So far we have been looking at overall correct performance. Figure 6 replots the same data in terms of error scores, collapsed over the lead continuum. In the S/S experiments, when subjects only had to report which ear led, the error rate was 25%. When more complex information processing was required, namely when they had to report which stimulus led, the error rate increased to 39%. This increase,

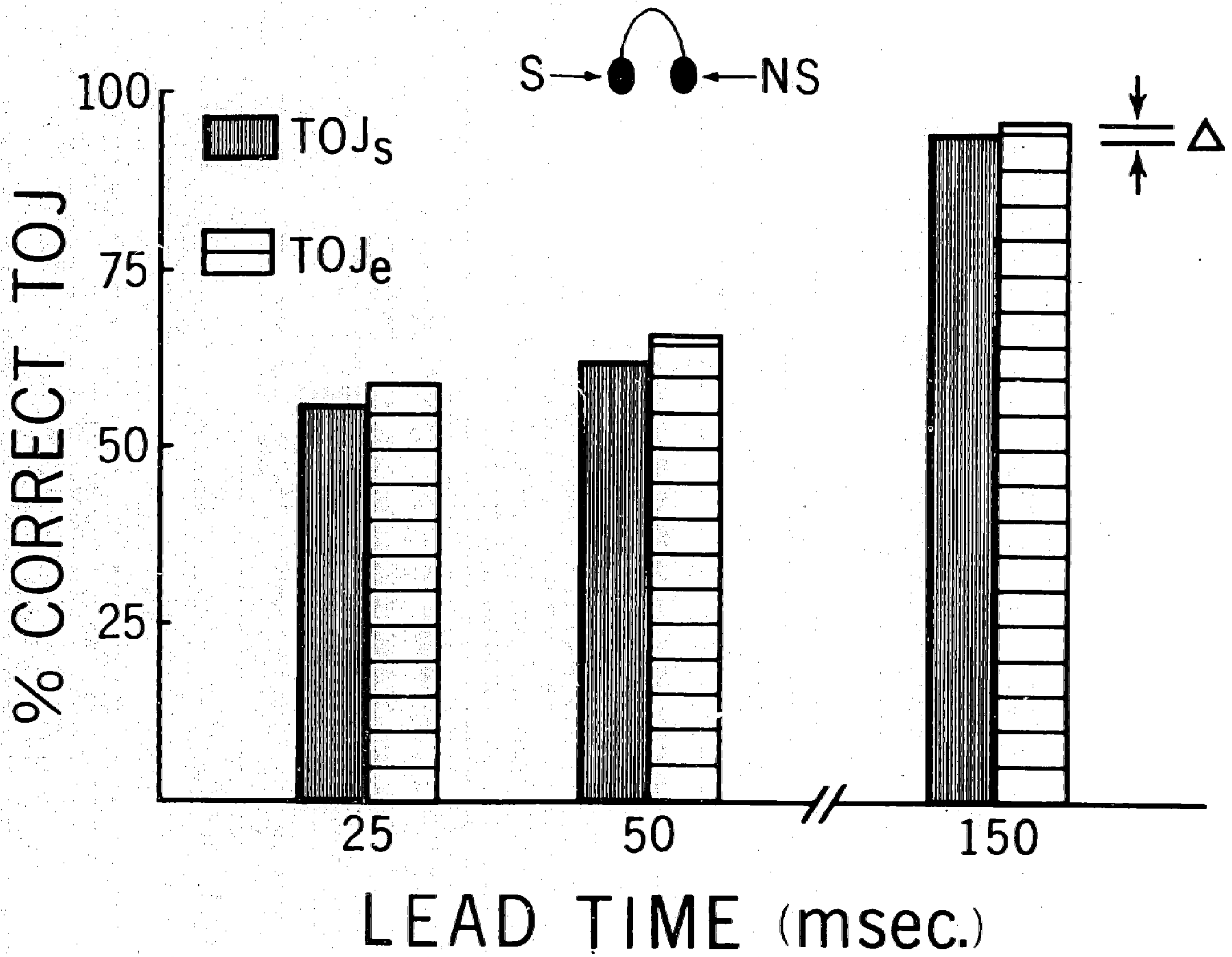


Figure 4: Percent correct for the two TOJ tasks when one stimulus is speech and the other nonspeech.

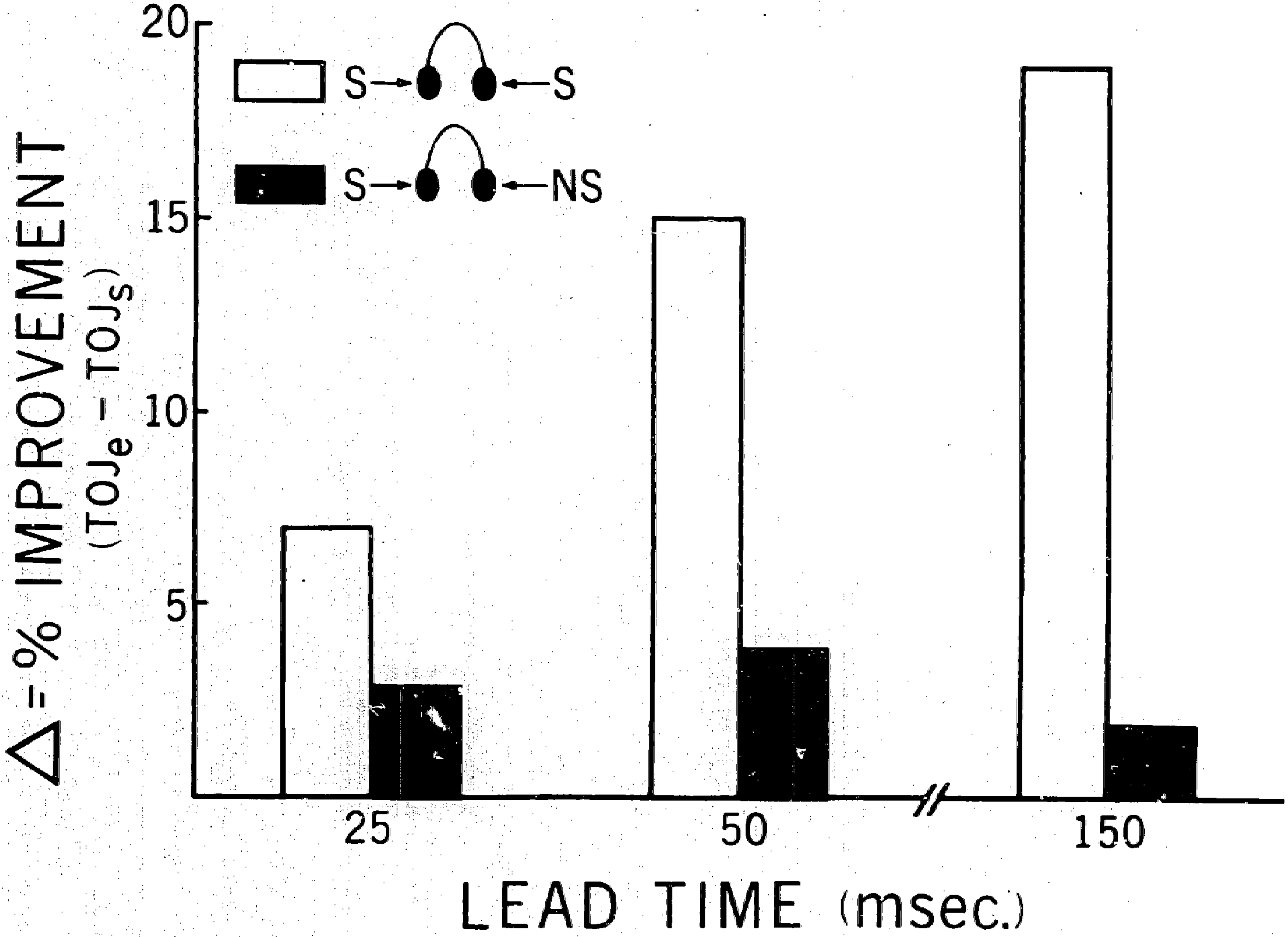


Figure 5: Performance differences between the two TOJ tasks for S/S and S/NS experiments.

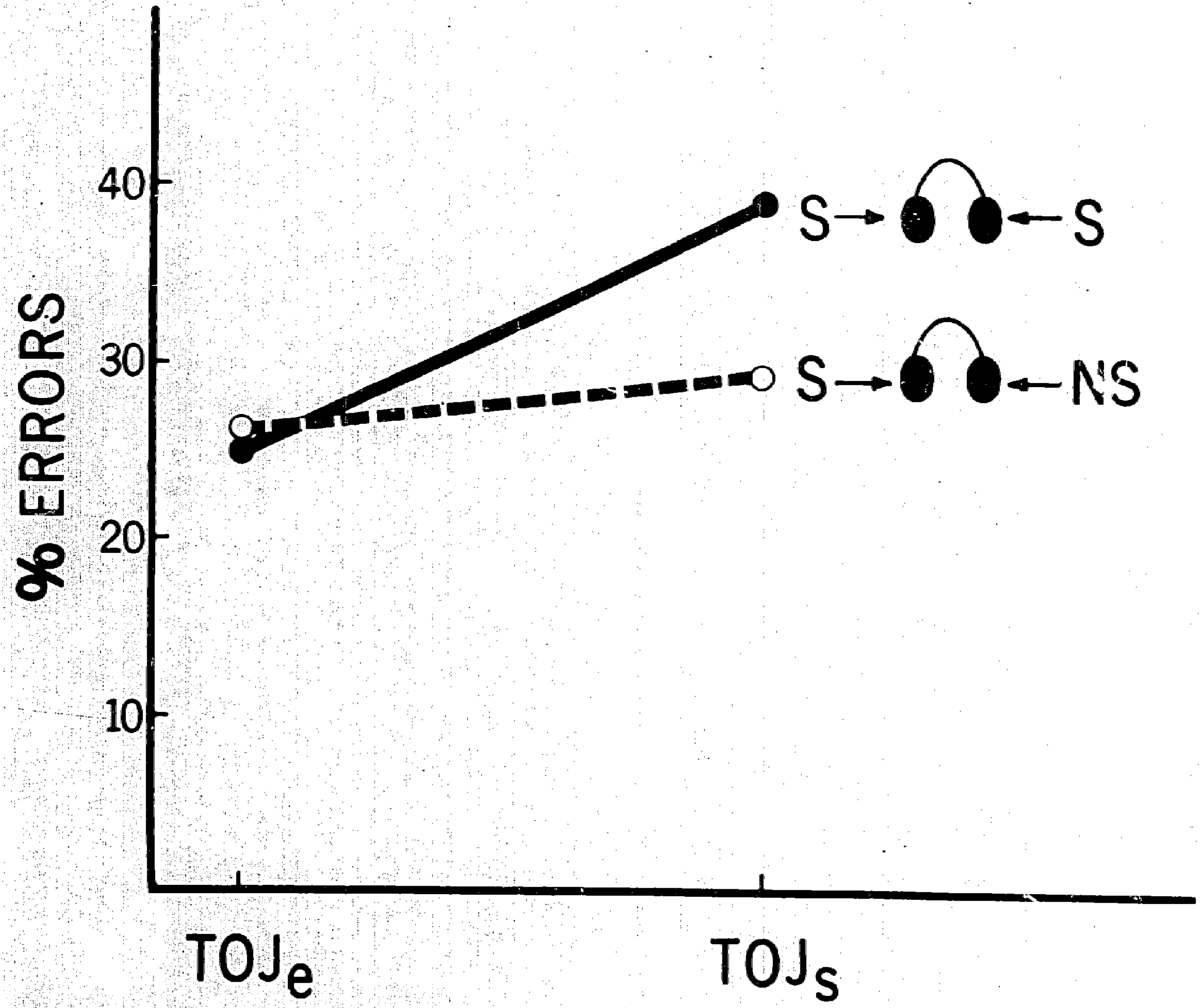


Figure 6: Error scores for the two TOJ tasks in S/S and S/NS experiments.

as we have seen, is highly significant. This is the type of result we would expect if both stimuli were requiring the services of a single processor.

The S/NS experiment yielded comparable error rates for both TOJ tasks: 26% errors for TOJ_e and 29% for TOJ_s. Remember that in the TOJ_s task, the subject must perform all the analysis functions necessary to identify the leading stimulus. Nevertheless, in the S/NS situation, this "extra work" did not yield increased error rates. Such results are consistent with the view that speech and nonspeech stimuli are sent to separate processors for analysis. The data and explanations presented here for the two TOJ tasks are compatible with those discussed earlier in terms of standard dichotic identification tasks (summarized in Figure 1).

We are also comparing S/S cases with S/NS cases in other situations. Recently we completed some studies using dichotic, binaural, and monotic modes of presentation. For S/S cases, a given level of correct performance was obtained under dichotic presentation; performance increased when these stimuli were presented binaurally and monotically (Day and Copeland, forthcoming-a). However, for S/NS cases, performance levels were identical for all modes of presentation (Day and Bartlett, forthcoming). Thus perception of S/NS items is independent of mode of presentation as well as type of TOJ task.

In view of the various findings discussed above, we cannot rule out the possibility that there can be separate processing mechanisms for speech and nonspeech.

REFERENCES

Cooper, F.S. and I.G. Mattingly. (1969) Computer-controlled PCM system for investigation of dichotic speech perception. J. acous. Soc. Amer. 46, 115(A).

Day, R.S. (1970) Temporal order judgments in speech: Are individuals language-bound or stimulus-bound? Haskins Laboratories Status Report on Speech Research SR-21/22, 71-87.

Day, R.S. (forthcoming) Release from language-bound perception.

Day, R.S. and J.C. Bartlett. (forthcoming) Perception of paired speech and nonspeech messages under dichotic, binaural, and monotic modes of presentation.

Day, R.S. and P.M. Copeland. (forthcoming-a) Phonemic fusion under dichotic, binaural, and monotic modes of presentation.

Day, R.S. and P.M. Copeland. (forthcoming-b) Perception of fusable and nonfusable dichotic stimuli.

Day, R.S. and J.E. Cutting. (1970) Levels of processing in speech perception. Paper presented at the 10th annual meeting of the Psychonomic Society, San Antonio, November.

Day, R.S. and J.E. Cutting. (1971a) Perceptual competition between speech and nonspeech. J. acous. Soc. Amer. 49, 85(A). (Also in Haskins Laboratories Status Report on Speech Research SR-24, 35-46.)

Day, R.S. and J.E. Cutting (1971b) What constitutes perceptual competition in dichotic listening? Paper presented at the Eastern Psychological Association meeting, New York, 17 April.

Dichotic Fusion Along an Acoustic Continuum[*]

James E. Cutting[+] and Ruth S. Day[+]
Haskins Laboratories, New Haven

ABSTRACT

When stimuli such as BANKET and LANKET are presented dichotically, phonemic fusions often occur: subjects report hearing BLANKET. Previous studies have shown that stop + /r/ and stop + /l/ items have different fusion properties. For example, /l/ was sometimes substituted for /r/ (but rarely vice versa): GOCERY/ROCERY→ (yielded) GLOCERY. The present experiment varied the liquid stimuli along an acoustic continuum involving the third-formant transition. For example, one set varied from RAY to LAY. Each was paired dichotically with an initial stop stimulus, in this case, PAY. All inputs (PAY, RAY, LAY) and possible fusions (PRAY, PLAY) were acceptable English words. When asked to report "what they heard," subjects gave many fusion responses. Of these, there was a preponderance of stop + /l/ fusions (88% vs. 12%). They occurred even for pairs where the liquid item was reported as an /r/ during separate binaural identification trials. Thus, given that an item was identified as RAY, the same subjects reported hearing PLAY when it was paired with PAY: PAY/RAY→PLAY. Despite the fact that the third-formant transition is crucial for perception of /r/ vs. /l/, this parameter was not responsible for the observed phoneme substitutions.

Most of the dichotic listening literature to date has dealt with the phenomenon of perceptual rivalry. Given a different stimulus to each ear, the subject typically reports hearing one or both of them. Different information contained in each stimulus is not combined into a single percept. Thus, given the dichotic digits ONE/FIVE, the subject does not report hearing FUN or WIVE. Perceptual fusion does occur, however, when certain psycholinguistic variables have been taken into account (Day, 1968). For example, given the dichotic pair BANKET/LANKET, subjects often report hearing BLANKET (Day, 1970a, forthcoming; Day and Cutting, 1970). This phenomenon of phonemic fusion has been obtained for various types of consonant clusters, including initial stop + liquid clusters such as BANKET/LANKET (Day, 1970b) and final stop + fricative clusters such as TASS/TACK (Day, 1970b).

One of the intriguing findings in the phonemic fusion studies is that some clusters fuse more readily than others (Day, 1968). For example, BACK/

---

[*] Paper to be presented at the 83rd meeting of the Acoustical Society of America, Buffalo, New York, April 1972.

[+] Also Yale University, New Haven.

LACK→BLACK[1] more readily than TACK/RACK→TRACK (Day, forthcoming-b). Studies using natural speech stimuli have obtained fusion rates for initial stop + /l/ clusters that are more frequent than those for initial stop + /r/ clusters. Day (1968) noted that such differential fusion rates cannot be explained by the relative frequency of these clusters in English. In fact, frequency data show the reverse trend: stop + /r/ clusters outnumber stop + /l/ clusters. She suggested that the differential fusion rates might be explained on an acoustic level.

Another curious finding further distinguished /r/ from /l/. Given GOCERY/ROCERY, subjects sometimes reported hearing GLOCERY. Thus /l/ was substituted for /r/. This /l/-for-/r/ substitution occurred quite often, while the reverse substitution rarely occurred.

The present set of studies was designed to explore the different fusion properties of /r/ and /l/ by varying a relevant acoustic cue. O'Connor et al. (1957) and Lisker (1957) have shown that the third-formant transition is crucial for the perception of initial /r/ vs. /l/. Hence this acoustic cue was systematically varied along an F_3 continuum from /r/ to /l/; the resulting stimuli were then paired with an appropriate stimulus beginning with a stop consonant, and the resulting fusion rates were observed.

General Method

Stimuli. Four fusion sets of the same general pattern were selected: the PAY set (PAY/RAY→PRAY, PAY/LAY→PLAY); the BED set (BED/RED→BREAD, BED/LED→BLED); the CAM set (CAM/RAM→CRAM, CAM/LAMB→CLAM); the GO set (GO/ROW→GROW, GO/LOW→GLOW). Thus, for each set, all phoneme strings were identical, except for the initial element. Dichotic trials consisted of a stop-initial string such as PAY presented to one ear and a liquid-initial string such as RAY or LAY to the other ear. All inputs and possible fusions were acceptable English words. Furthermore, all have a relatively high frequency of occurrence in the language [most have Thorndike-Lorge (1944) frequencies of 100 per million].

The stimuli were prepared on the parallel resonant synthesizer at the Haskins Laboratories. The acoustic form of each stop was identical on all presentations. However the liquids varied along an acoustic continuum as shown in Figure 1. For each liquid array such as the /rɛɪ/-/lɛɪ/ array, the stimuli were identical in all respects except for the first 150 msec of the third formant (F_3). F_3 was varied in such a manner as to yield perception of an initial /r/ at one end of the continuum and an initial /l/ at the other end. Each stimulus in the array began with a different initial F_3 value: stimulus A = 1524 Hz, B = 1849 Hz, C = 2180 Hz, D = 2525 Hz, and E = 2862 Hz. After an initial steady-state portion of 50 msec at these respective values, F_3 underwent the appropriate changes to reach the target value of 2525 Hz for the following vowel: stimuli A, B, and C rose for 100 msec, D held steady at the target value, and E fell for 100 msec. For the remaining duration of the stimuli, F_3 values were identical for all members of a given array. Meanwhile other formant information was held constant across a given liquid

---

[1] The arrow should be read as "yields."

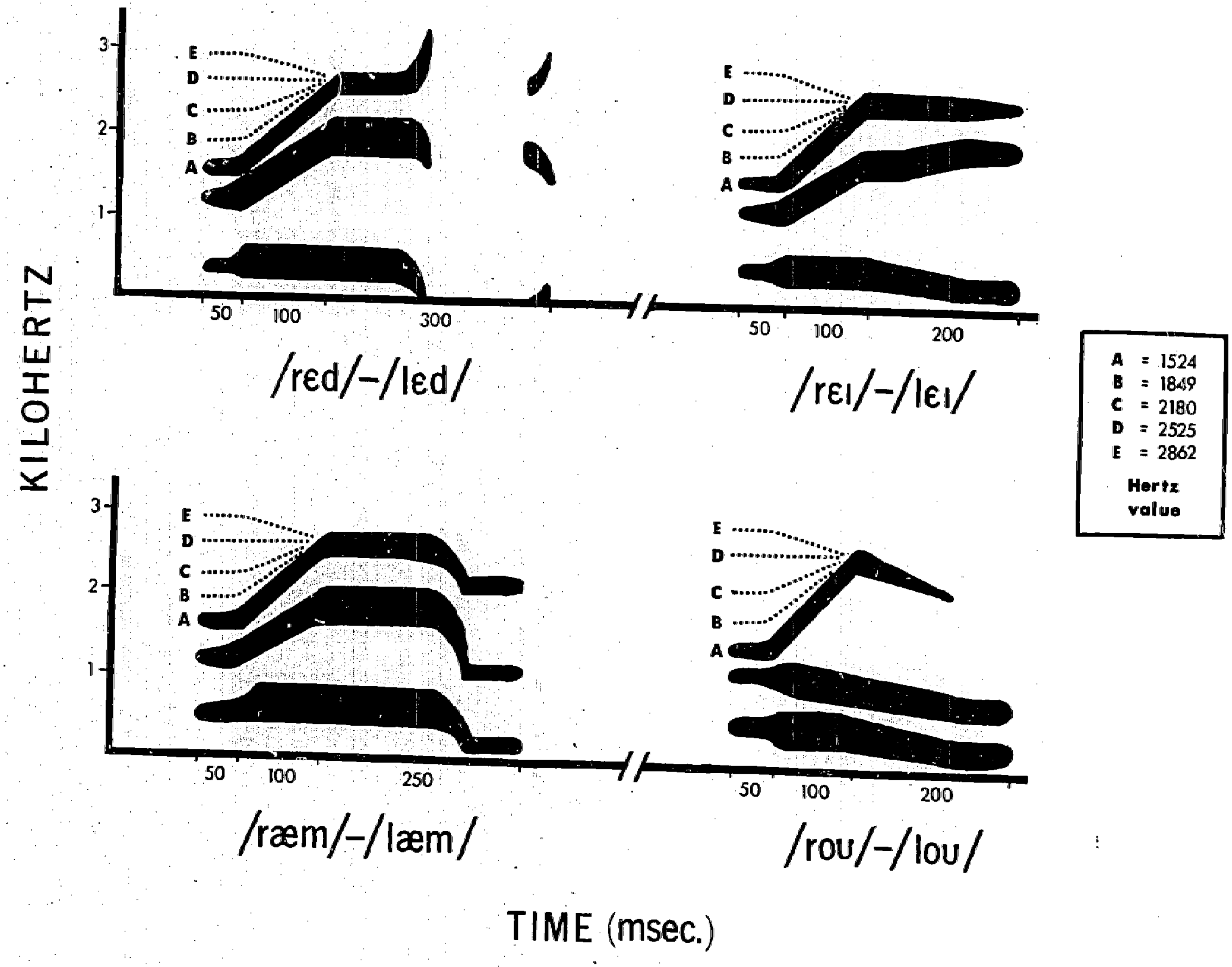


Figure 1: Schematic spectrograms of the liquid stimuli.

array. F_1 began with a 50-msec steady-state portion at 513 Hz followed by an abrupt 20-msec transition to its resting frequency for the following vowel, while F_2 began with a 50-msec steady-state segment at 1155 Hz and was followed by a 100-msec transition to its resting frequency.

Subjects. The same sixteen students served in both experiments. They were between the ages of 16 and 26, were right-handed, native American English speakers, and had no history of hearing trouble. Subjects were tested in groups of four, with stimuli played on an Ampex AG500 tape recorder and sent through a listening station to Grason-Stadler earphones.

Experiment I - Liquid Identification

Tape. In order to evaluate the quality of the individual liquid stimuli, a binaural identification tape was prepared. There were 200 trials: (5 stimuli per array) x (4 arrays) x (10 observations per stimulus). The items occurred in random order with a 1.5-sec interstimulus interval.

Procedure. Subjects were introduced to the stimuli by hearing the endpoints of each array (stimuli A and E). No one had difficulty in perceiving them as /r/ and /l/. Then a forced-choice procedure was used. The binaural identification tape was played, and subjects wrote down an "r" or an "l" to indicate the first phoneme they heard on every presentation.

Results. Perception of the stimuli in each array was categorical as shown in Figure 2. The two stimuli in each array with the lowest and fastest-rising F_3 (stimuli A and B) were perceived as beginning with /r/, while the two stimuli at the other end of the continuum with the highest and nearly steady-state F_3 (stimuli D and E) were perceived as beginning with /l/. The middle stimulus (C) was ambiguous: about half the time it was perceived as beginning with /r/ and half the time with /l/. There were no significant differences among the subjects or among the groups of subjects. Most subjects split their responses evenly for stimulus C. Three out of the four liquid arrays showed the basic symmetry described above. However the /ræm/-/læm/ array showed a slight asymmetry in favor of the /l/ end of the continuum.

Discussion. Many speech sounds are perceived categorically. That is, equal changes in an acoustic parameter do not yield equal changes in perception. Instead there is a quantal change in perception somewhere along the acoustic continuum. Acoustic cues that yield categorical perception for stop consonants are the direction and extent of the second-formant transition (Liberman, 1957) and voice onset time (Lisker and Abramson, 1967). Recently Pisoni (1971) showed that vowels can be perceived categorically if the duration of isolated vowels is short enough. The present experiment showed that liquids can also be perceived categorically when the F_3 transition is varied, thus supporting the earlier work of O'Connor et al. (1957). Categorical perception appears to be unique to speech sounds. Mattingly et al. (1971) demonstrated that when synthetic syllables composed of a stop consonant and a vowel are broken down into bleats (single second formants) or chirps (F_2 transitions) or are reversed, categorical perception disappears.

Since orderly identification functions for liquids were obtained in Experiment I, these stimuli are suitable for pairing with appropriate stop

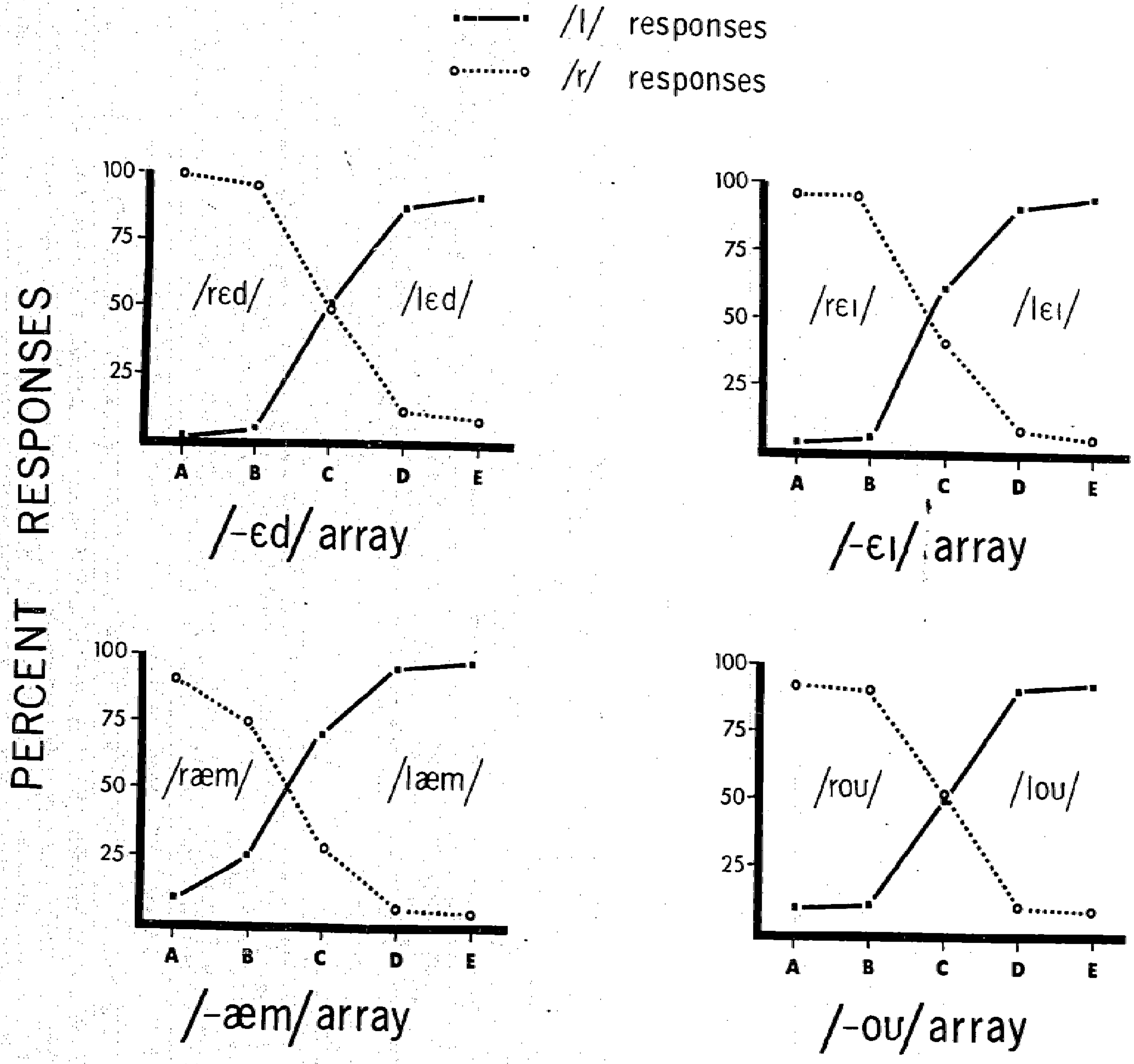


Figure 2: Identification data for the liquid stimuli.

stimuli in order to study dichotic fusion. In particular, we are concerned with the fusion properties of stimuli at each point along the F_3 continuum.

Experiment II - Dichotic Fusion

Tape. An appropriate stop-initial stimulus was prepared for each liquid array. For example, PAY was synthesized for the RAY and LAY stimuli. The pulse code modulation system (Cooper and Mattingly, 1969) was used to digitize each stimulus, store it on a disc, and prepare dichotic tapes. Each stop was paired with each member of the liquid array. The relative onset time was varied for each pair: the stop led by 50 msec, or the liquid led by 50 msec, or both began at the same time. Alignment accuracy was ±1/2 msec. There were 120 pairs on the tape: (5 pairs per array) x (4 arrays) x (3 lead times) x (2 channel arrangements).

Procedure. Subjects were told to report "what they heard" on every trial, no matter whether they heard a "real word," "nonsense word," one word or two.

Results. Fusion occurred readily for all stimulus sets. The F_3 manipulation had no effect on fusion rate: fusions occurred at a comparable rate at each point along the liquid continuum.

Given the particular fusion level for a given stimulus set, when did subjects perceive stop + /r/ clusters and when stop + /l/ clusters? One might expect these results to parallel those obtained in the binaural identification test. For example, when PAY is paired with the most "/r/-like" of its liquids (stimuli A and B), there ought to be a high proportion of PRAY responses; and pairs with the ambiguous stimulus C ought to split more or less evenly into PRAY and PLAY. These predictions are summarized in Figure 3 for each fusion set. Despite the reasonable nature of these predictions, Figure 4 shows a very different pattern of results. Consider the PAY set. Most fusions were PLAY, independent of which stimulus was presented. Even though liquid stimuli A and B were identified as RAY better than 95% of the time on the binaural identification test, when these stimuli were paired with PAY in the dichotic task, 86% of all fusions were PLAY.

The data from all four stimulus sets were pooled and are summarized in Figure 5. Perception of the liquids was categorical in the binaural identification test (Figure 5A). Nevertheless, when these stimuli were paired with the appropriate stop-initial stimuli, 88% of all fusions were stop + /l/ (Figure 5B).

Discussion. Despite the fact that the F_3 transition is crucial for the perception of /r/ vs. /l/, this parameter was not responsible for the phoneme substitutions observed in the dichotic fusion task. We seem to have a perceptual elephant: the /l/-for-/r/ substitution is a large, robust phenomenon. We tried to bring it under control by varying a highly relevant acoustic parameter--and failed. In more recent studies we have attacked it with a whole arsenal of parameters: we have varied the relative fundamental frequency of the dichotic inputs, their relative intensities, the vocal tract configurations of the stimuli, F_2 transitions of the liquids, and the duration of the initial steady-state portion of the liquids. The result is analogous

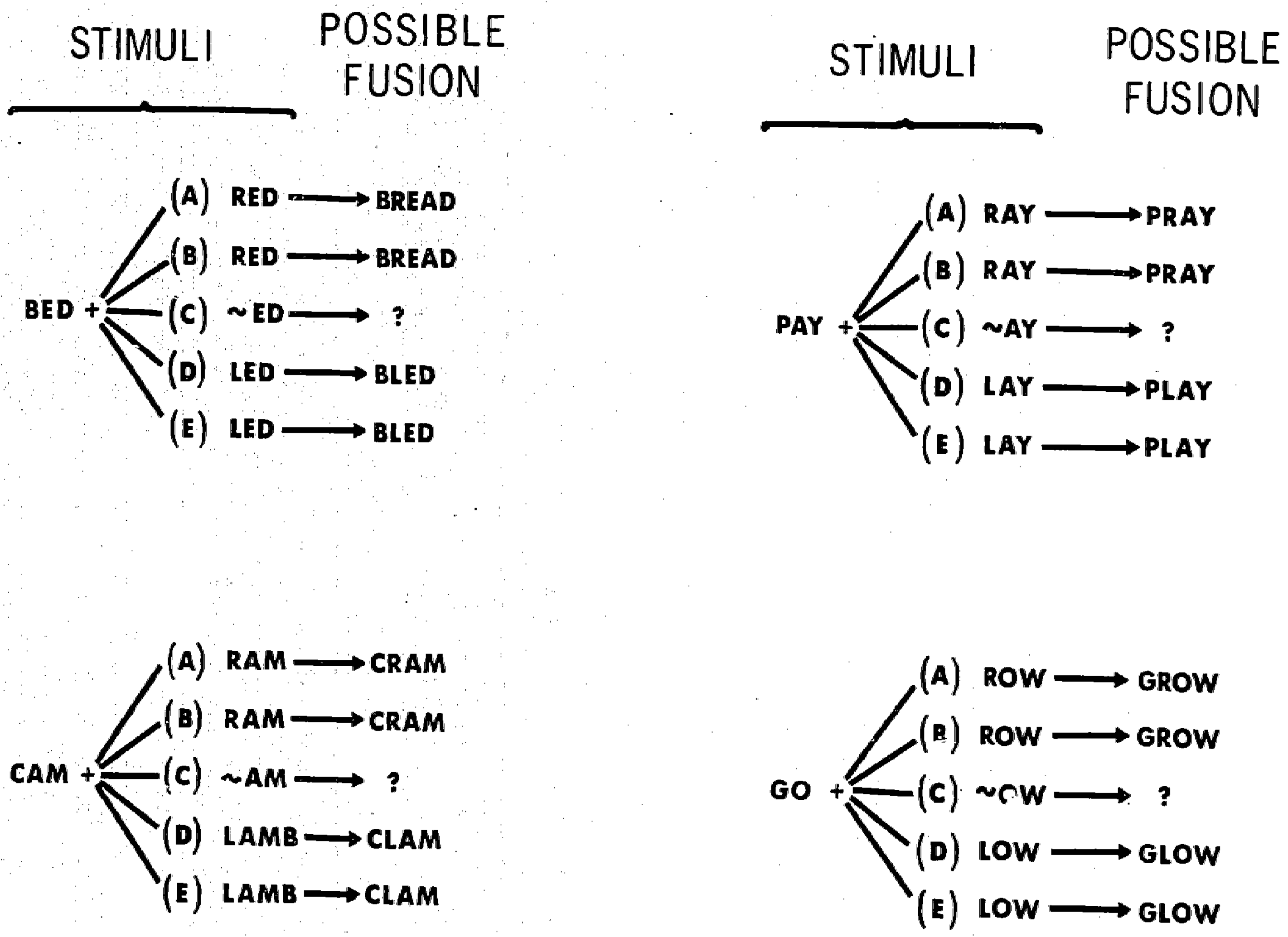


Figure 3: Predicted fusions, given liquids that vary from /r/ to /l/.

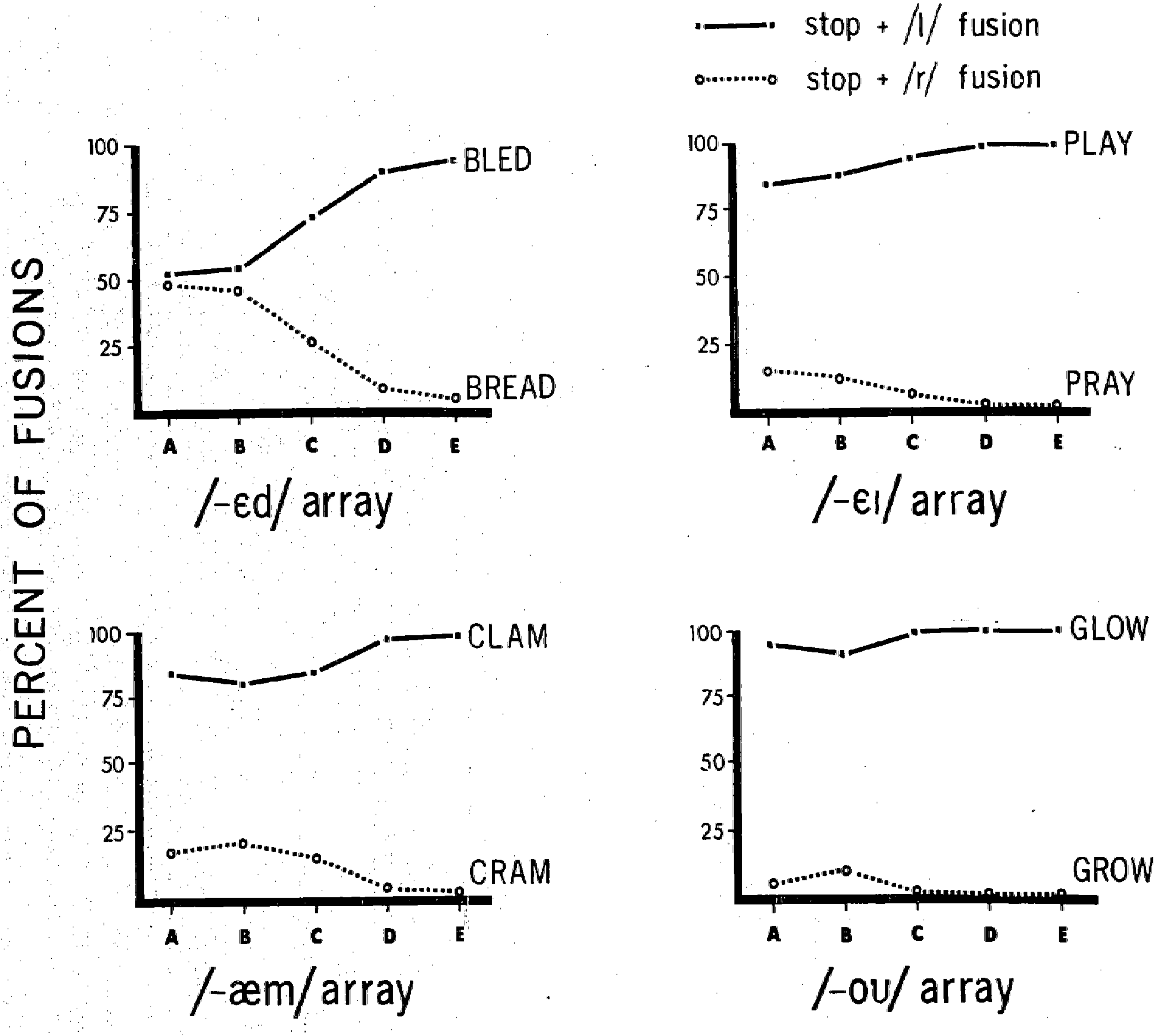


Figure 4: Percent of all fusions reported as stop + /r/ and stop + /l/.

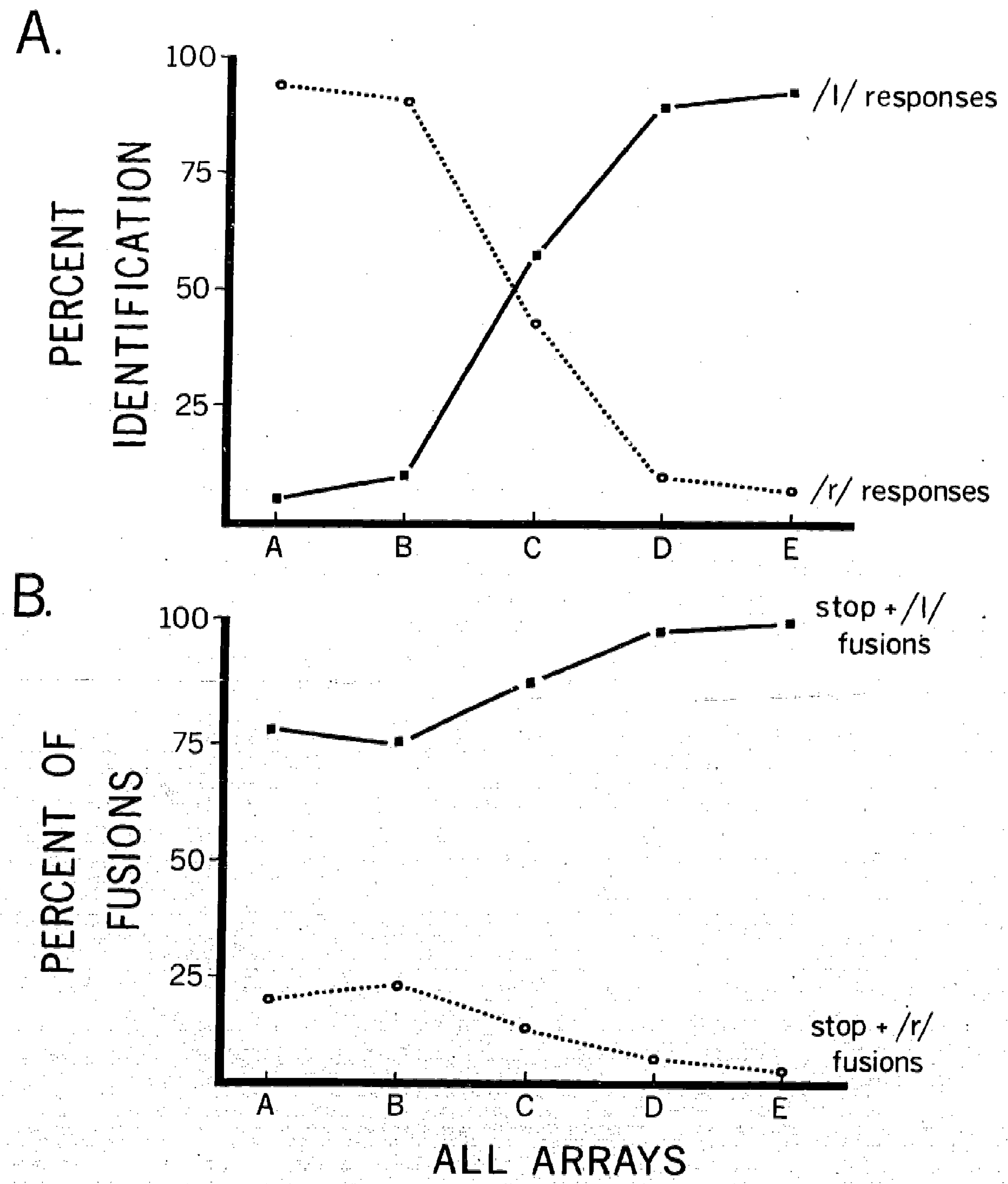


Figure 5: Comparison of identification and fusion results (collapsed over the four liquid arrays).

to shooting a perceptual elephant with acoustic beebees: there has been virtually no change in the /l/-for-/r/ substitution rate.

The development of phoneme production in children shows some interesting distinctions between /r/ and /l/. Most children acquire /l/ before /r/; the average age for /l/ is about 4 years, while that for /r/ is about 5 years (Powers, 1957). Mispronunciation data are of particular interest. One study (Morley, 1957, p. 42a) showed that /r/ errors were almost forty times more frequent than /l/ errors in children aged 3 years 9 months. Failure to produce /r/ resulted in a /w/-substitution about 60% of the time, /l/-substitution about 20%, and /y/ about 15%. On those occasions when the children failed to produce /l/, the target phoneme was replaced by /w/, not /r/. Apparently these differential difficulties with /r/ and /l/ are not readily amenable to therapeutic techniques: speech therapists sometimes comment that it is very difficult (if not impossible) to teach children to produce /r/ (Murray, 1962).

Production errors in adults are also revealing. Within a few days we overheard the following errors: CRIPTIC was produced as CLIPTIC, PRESENT as PLEASANT, and INCREASES as INCLEASES. Freud (1901) has discussed slips of the tongue in terms of psychoanalytic notions. We suspect that such errors as CLOWN PRINCE may have little to do with repressed hostility toward authority and more to do with linguist concerns. (Recently, one of the authors was discussing this example and said CLOWN PLINCE.)

Delayed auditory feedback (DAF) also yields distinctions between /r/ and /l/. When subjects read word lists under DAF, they often reduplicate /r/ but had little difficulty with /l/ (Applegate, 1968).

Deaf people have selective difficulty with /r/ (Rosen, 1962). Stop + /r/ clusters are more difficult to perceive than stop + /l/ clusters. Misperception of initial liquids yields an /l/-for-/r/ substitution rate which is twice as frequent as the reverse substitution.

The number of distinct articulations of /r/ in American English may be greater than those of /l/. There may also be greater cross-linguistic variation in /r/. It is commonly noted that second language learners have difficulty with /r/; for example, English speakers have trouble with the uvular /r/ in French and the trilled /r/ in Spanish.

Perhaps these very diverse observations can be unified in a single concept: that of stability. In both perception and articulation /l/ is relatively stable, while /r/ is relatively unstable. The dichotic fusion experiment may be viewed as a situation of information overload. It is interesting to note that it is the less stable /r/ that suffers in this demanding situation.

REFERENCES

Applegate, R.B. (1968) Segmental analysis of articulatory errors under delayed auditory feedback. University of California (Berkeley): Project on Linguistic Analysis 8, 1-27.

Cooper, F.S. and I.G. Mattingly. (1969) Computer-controlled PCM system for investigation of dichotic speech perception. J. acous. Soc. Amer. 46, 115(A).

Day, R.S. (1968) Fusion in dichotic listening. Unpublished Ph.D. thesis (Psychology), Stanford University.

Day, R.S. (1970a) Temporal order judgments in speech: Are individuals language-bound or stimulus-bound? Haskins Laboratories Status Report on Speech Research SR-21/22, 71-87.

Day, R.S. (1970b) Temporal order perception of a reversible phoneme cluster. Haskins Laboratories Status Report on Speech Research SR-24, 47-56.

Day, R.S. (forthcoming-a) Release from language-bound perception.

Day, R.S. (forthcoming-b) Dichotic fusion despite knowledge of the stimuli.

Day, R.S. and J.E. Cutting. (1970) Levels of processing in speech perception. Paper presented at the 10th annual meeting of the Psychonomic Society, San Antonio, November.

Freud, S. (1901) The Psychopathology of Everyday Life. (London: Hogarth Press).

Liberman, A.M. (1957) Some results of research on speech perception. J. acous. Soc. Amer. 29, 117-123.

Lisker, L. (1957) Minimal cues for separating /w,r,l,y/ in intervocalic position. Word 13, 257-267.

Lisker, L. and A.S. Abramson. (1967) Some effects of context on voice onset time in English stops. Language and Speech 10, 1-28.

Mattingly, I.G., A.M. Liberman, A.K. Syrdal, and T. Halwes. (1971) Discrimination in speech and nonspeech modes. Cog. Psychol. 2, 131-157.

Morley, M.E. (1957) Development of Speech Disorders in Childhood. (London: Livingstone).

Murray, R.S. (1962) The development of /r/ in the speech of preschool children. Unpublished Ph.D. thesis (Speech), Stanford University.

O'Connor, J.D., L.J. Gerstman, A.M. Liberman, P.C. Delattre, and F.S. Cooper. (1957) Acoustic cues for the perception of initial /w,j,r,l/ in English. Word 13, 25-43.

Pisoni, D.B. (1971) On the nature of categorical perception of speech sounds. Unpublished Ph.D. thesis (Psycholinguistics), University of Michigan.

Powers, M.H. (1957) Functional disorder of articulation: Symptomatology and etiology. In Handbook of Speech Pathology, L.E. Travis, ed. (New York: Appleton-Century-Crofts).

Rosen, J. (1962) Phoneme identification in sensorineural deafness. Unpublished Ph.D. thesis (Speech), Stanford University.

Thorndike, E.L. and I. Lorge. (1944) The Teacher's Word Book of 30,000 Words. (New York: Teachers College Press).

The Activity of the Intrinsic Laryngeal Muscles in Voicing Control: An Electromyographic Study*

Hajime Hirose[+] and Thomas Gay[++]
Haskins Laboratories, New Haven

INTRODUCTION

Although a considerable number of laryngeal electromyographic (EMG) studies on the mechanism of phonation have been conducted during the past decade, EMG studies of the intrinsic laryngeal muscles during speech are still in their preliminary stages. This is due, mainly, to the technical difficulties in data acquisition using conventional needle electrodes during complex and rapid movements of the articulators in speech gestures and, also, in extracting subtle changes in muscle activity patterns from raw EMG data. However, recent advances in EMG recording and processing techniques have helped us overcome these technical problems. In particular, the use of double-ended hooked-wire electrodes (Basmajian and Stecko, 1962) give us a combination of electrode stability in the muscle with little discomfort to the subject. Also, the use of a digital computer system to obtain an averaged EMG activity pattern for a number of tokens of a given speech utterance provides a convenient and accurate means for quantifying a pattern of contraction of a given muscle or muscle group.

Most of the previous studies in laryngeal physiology generally support the classical division of the intrinsic laryngeal muscles into three functional groups: abductor, adductor, and tensor. However, there still are many unanswered questions concerning the function of individual laryngeal muscles in speech articulation.

In particular, the participation of the posterior cricoarytenoid muscle (PCA) in speech has not been systematically studied, although the function of the PCA as a respiratory muscle has been well documented (Pressman, 1942; Suzuki and Kirchner, 1969). As far as PCA activity in phonation is concerned, Faaborg-Andersen (1957) reported that EMG activity of the PCA decreased during sustained phonation. Kotby and Haugen (1970), on the other hand, observed increasing activity in the PCA during phonation and postulated that the PCA is not solely an abductor muscle. Dedo (1970) also reported increasing activity in the PCA during phonation in some of his clinical cases. However, the data of these authors are concerned exclusively with sustained vowel phonation, when fundamental frequency is not specified.

---

*This is an expanded version of a paper presented at the 1971 Convention of the American Speech and Hearing Association, Chicago, Illinois, November 1971.

[+] Also Faculty of Medicine, University of Tokyo.

[++] Also University of Connecticut Health Center, Farmington.

Hiroto, Hirano, Toyozumi, and Shin (1967) examined laryngeal muscle activity for some Japanese words containing an intervocalic fricative /s/ and stated that there was a temporary change in the electrical activity of all the intrinsic laryngeal muscles (except for the cricothyroid) corresponding to voiceless consonant articulation. What they observed in their data was an apparent increase in PCA activity accompanied by a decrease in the activity of the adductors for articulation of the intervocalic /s/. Hirano and Ohala (1969) showed one example of a raw EMG record of the PCA, illustrating increasing activity for release of glottal stops with reciprocally decreasing activity in the interarytenoid.

As far as the adductor laryngeal muscles are concerned, there has been, again, no systematic description of their function in speech articulation, although the possibility of functional differentiation of the adductor muscle group was suggested by one of the present authors in a previous report (Hirose, 1971a).

The primary purpose of the present study was to systematically investigate the actions of the intrinsic laryngeal muscles in speech with special reference to the articulation of segmental features of American English. Particular attention was directed to the function of the PCA. An attempt was also made to investigate the temporal aspects of consonant production by studying the timing relationships between laryngeal and supralaryngeal muscle activity patterns.

PROCEDURES

Subjects

The present experiment was performed on two adult male subjects, both native speakers of American English; for one subject, two separate retest recordings were made, thus giving a total of four sets of data. Table I lists the muscles examined in each session for the experiment.

Preparation and Insertion of Electrodes

Hooked-wire electrodes, after the type developed by Basmajian and Stecko (1962), were used in the present study. Briefly, these electrodes are produced by threading a pair of thin wires through the cannula of a hypodermic needle and bending the exposed ends of the wires back over the needle to form a pair of hooks. The entire assembly is inserted into the muscle, after which the needle is withdrawn. This leaves only the hooked ends of the wires anchored into the muscle. Removal of the wires requires only a light tug. In this experiment, a platinum-iridium alloy (90%-10%) wire (.002 inch diameter and polyester enamel coated) was used in conjunction with either a No. 26 or No. 27 gauge needle.

The PCA and the interarytenoid (INT) muscles were reached perorally while the vocalis (VOC),[1] lateral cricoarytenoid (LCA), and cricothyroid (CT) muscles

---

[1] By reason of both past experience and the verification techniques employed, we are confident that we isolated the vocalis portion (vocalis muscle) of the thyroarytenoid. However, since the insertion was not viewed directly, we cannot be virtually certain that the electrode field did not include any potentials from the "external" thyroarytenoid.

TABLE I. Muscle insertions for both subjects.

Subject LL

- posterior cricoarytenoid
- interarytenoid
- vocalis
- cricothyroid

Subject LJR

Series A

- posterior cricoarytenoid
- vocalis
- cricothyroid

Series B

- posterior cricoarytenoid
- lateral cricoarytenoid
- obicularis oris
- genioglossus

Series C

- posterior cricoarytenoid
- interarytenoid

were reached percutaneously, after the procedure described by Hirano and Ohala (1969). Premedication consisted of the administration of 5-10 mg. Valium and 7-10 drops of Tincture of Belladonna by mouth. Subjects were seated in an examining chair throughout the experiment.

For the peroral insertions, an anesthesia procedure utilizing Cetacaine spray and a gargle of 2 ml. of 2% Xylocaine was sufficient to desensitize the pharyngeal and laryngeal areas to a point where indirect laryngoscopy could be easily tolerated. A Xylocaine-soaked cotton swab was then applied to the specific areas selected for electrode insertion. The PCA and the INT were reached by using an L-shaped rod with the carrier needle epoxy-bonded to the shorter arm. The needle was threaded in the conventional manner. The entire assembly was directed to the point of insertion by indirect laryngoscopy (Hirose, Gay, and Strome, 1971).

The percutaneous insertions were preceded by topical administration of 2% Xylocaine through a Pan Jet-70 air jet (Hirose, 1971b) at the site of the needle insertions. The electrode insertion techniques for the VOC, LCA, and

CT are described in detail in previous reports (Hirose, Gay, and Strome, 1971; Gay, Hirose, Strome, and Sawashima, in press).

In all cases, correct electrode placement was verified by monitoring an oscilloscope during various functional maneuvers. At the same time, the muscle signals were amplified and fed to a loudspeaker for auditory monitoring (Hirose, 1971b).

Data Recording and Processing

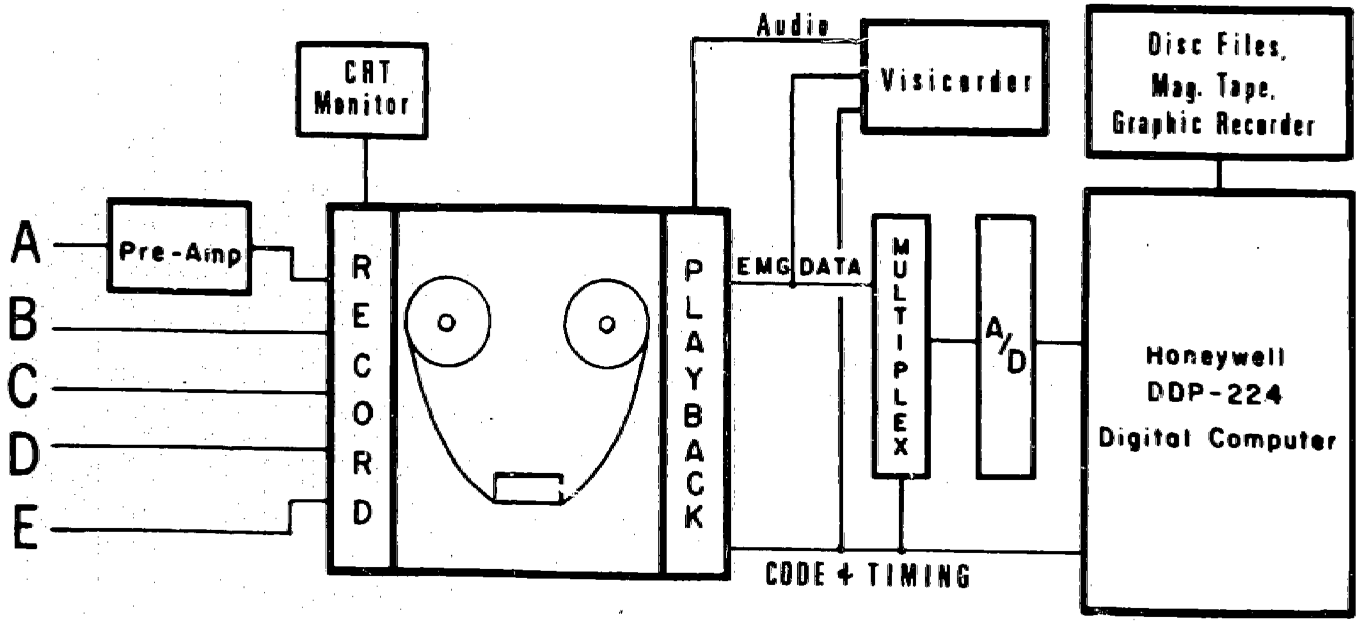
In order to obtain a convenient quantitative record of muscle activity, the raw EMG signal can be easily transformed into a display of amplitude versus time by the process of full-wave rectification and RC smoothing (integration). Generally speaking, the envelope of the integrated curve is an indication of the strength of the muscle contraction. This is only an approximation, however, as the amplitude of the recorded signal varies with the distance between the electrodes and the active motor units of the muscle. Further, since the integrated curve represents the vector sum of a number of asynchronously discharging motor unit potentials and since productions of identical utterances vary from one token to the next, a number of these curves must be averaged before a reasonably stable picture of muscle activity at a given electrode position can be obtained.

The basic data-processing procedure followed in this experiment was to collect EMG data for a number of tokens for each of the test utterances and, using a digital computer, average the integrated EMG signals at each electrode position.

A block diagram of the EMG recording and processing system used in the present study is shown in Figure 1. The system contains fourteen data channels, of which eight are for the recording of EMG signals. The other inputs are for the acoustic signal, air pressure data, a banter channel for the experimenter's comments, and finally, two channels for a clock track and digital code pulse. In addition, a calibration signal alternates with the EMG signals intermittently throughout the run. This signal is used by the computer to calculate the EMG levels in actual microvolts.

The purpose of the digital code pulse (octal format) is to identify each utterance for the computer. This pulse code is laid down on the tape, automatically, at one-second intervals. Before actual processing, the computer receives instructions on how the various tokens of a given utterance are to be superimposed or lined up with each other. This is done by marking the time interval between the nearest code pulse and any preselected line-up point, which can vary for each utterance type. During the data-processing run, all calculating and tabulating operations are done automatically. The averaged output curve is plotted on a strip chart recorder.

Timing measurements were obtained from a Honeywell visicorder optical oscillograph, and fundamental frequency measurements were made from sound spectrograms.



A - EMG (8 Channels)

B - Air Pressure (2)

C - Voice

D - Banter

E - Digital Code & Timing

Figure 1: Block diagram of the EMG data acquisition and processing system.

Experimental Conditions

The subjects were required to read randomized lists of the stimulus words sixteen times each. Stimulus words consisted of disyllabic nonsense words containing voiced and voiceless consonant pairs in both pre- and post-stressed positions. Typical examples of test words are given in Table II. For one subject, only /p/ vs. /b/ and /s/ vs. /z/ contrasts were examined, while pairs of three stops and four fricatives were examined for the other subject.

TABLE II. Test words used for both subjects.

ə C ʌp

b ʌ C ə

əb ʌ C

ʌb ə C

p ʌ p ə

h ʌ p ə

RESULTS

Voiced/Voiceless Contrast in Word-Medial Position

Averaged EMG curves for the voiced/voiceless contrast are shown in Figures 2-5. The curves in Figure 2 represent the averaged muscle activity levels for the PCA, INT, and VOC during the production of /p/ and /b/ in medial prestressed position (in /ə p ʌ p/ and in /ə b ʌ p/).

It is quite obvious that the PCA shows marked activity for production of /p/, while it is suppressed by /b/ as well as for vowel production.

For /ə p ʌ p/, PCA activity starts to decrease approximately 250 msec before the onset of /ə/. The activity then begins to increase 100 msec prior to stop release, after which it immediately begins to decrease again with the vowel production. It then shows another peak for final /p/, followed by a relatively higher level of activity, presumably for inspiration, after completion of the utterance.

For /ə b ʌ p/, on the other hand, PCA activity stays low throughout the voiced period from the initial vowel to the stressed vowel, including intervocalic /b/. It should be noted, however, that the EMG curve ascends slightly about 110 msec prior to the release of /b/, then descends again approximately at the time of the release, and finally rises steeply 40 msec before /p/ closure.

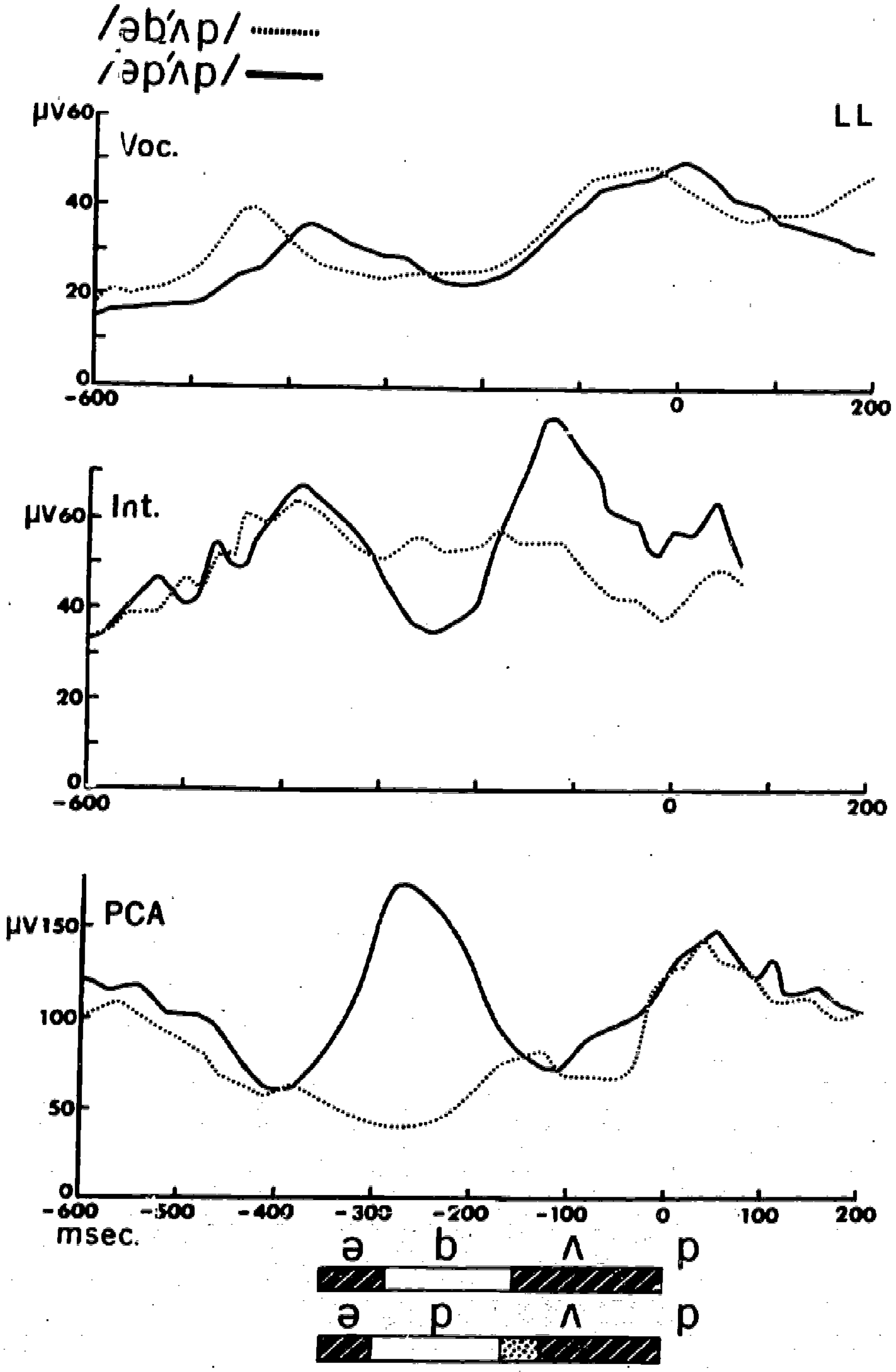


Figure 2: Superimposed averaged EMG curves of the VOC, INT, and PCA of Subject LL for the utterances, /əpʌp/ and /əbʌp/. The line-up point (0 on the abscissa) indicates voice offset of the stressed vowel.

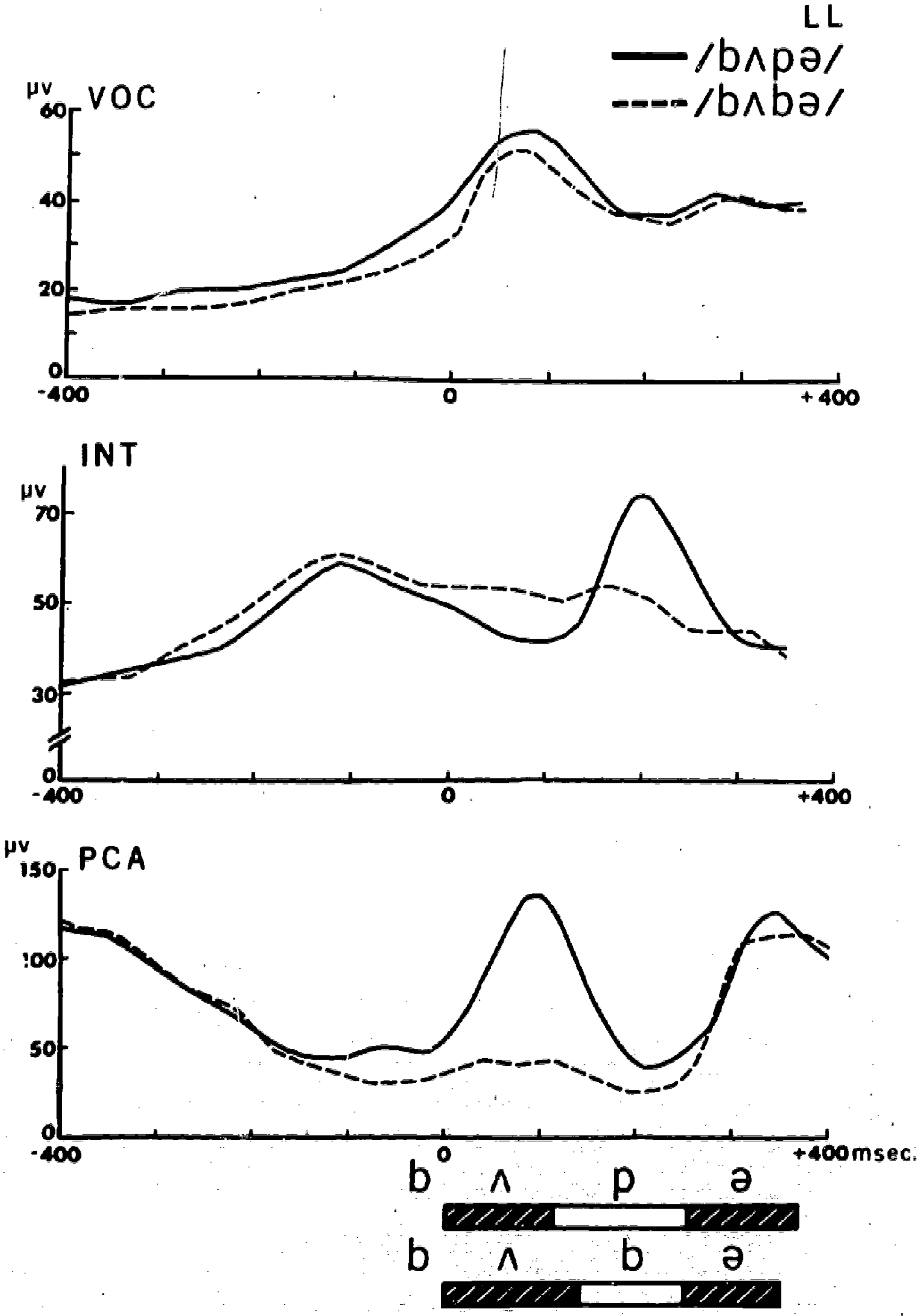


Figure 3: Averaged EMG curves for /bʌpə/ and /bʌbə/ (Subject LL). The line-up point is the onset of the stressed vowel.

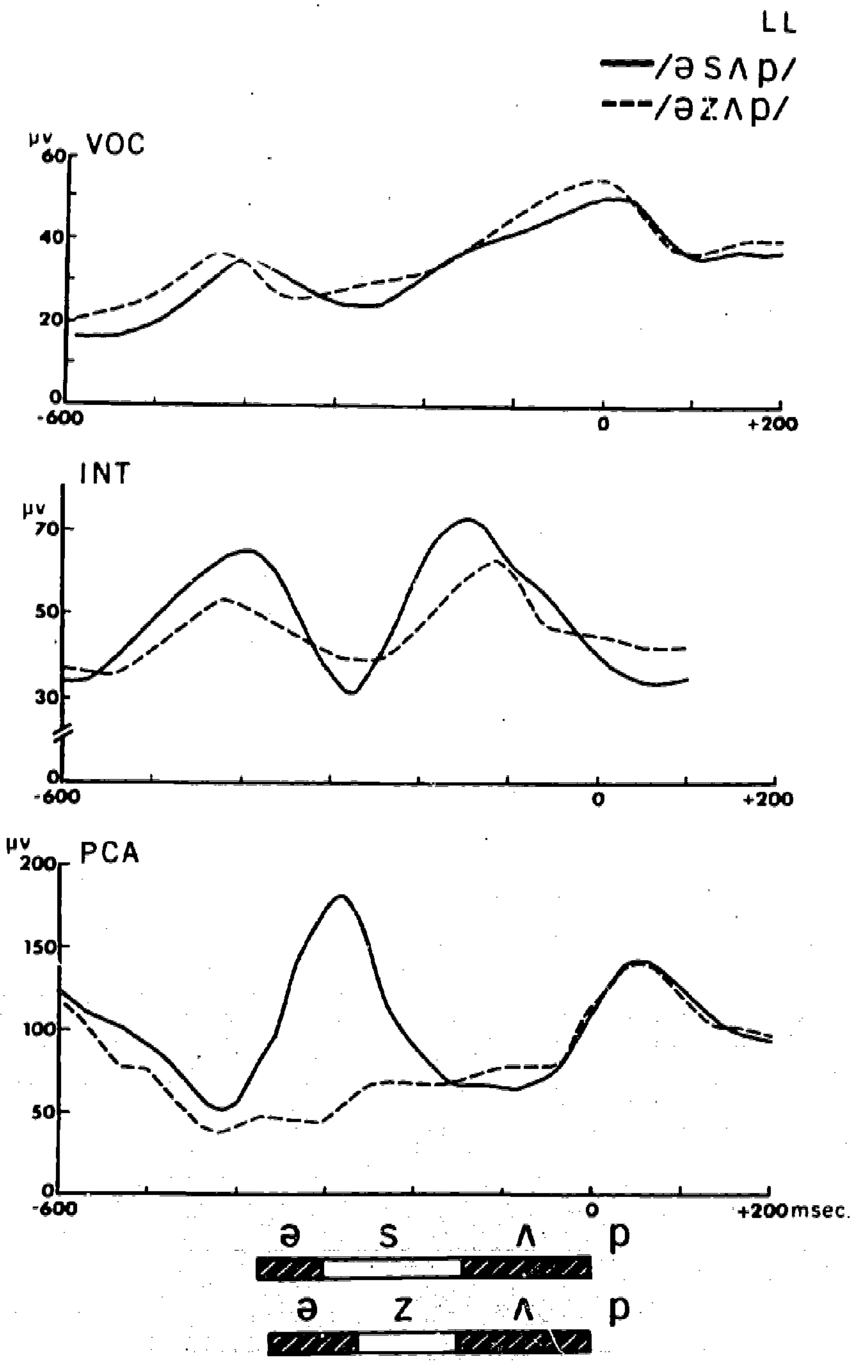


Figure 4: Averaged EMG curves for /əsʌp/ and /əzʌp/ (Subject LL).

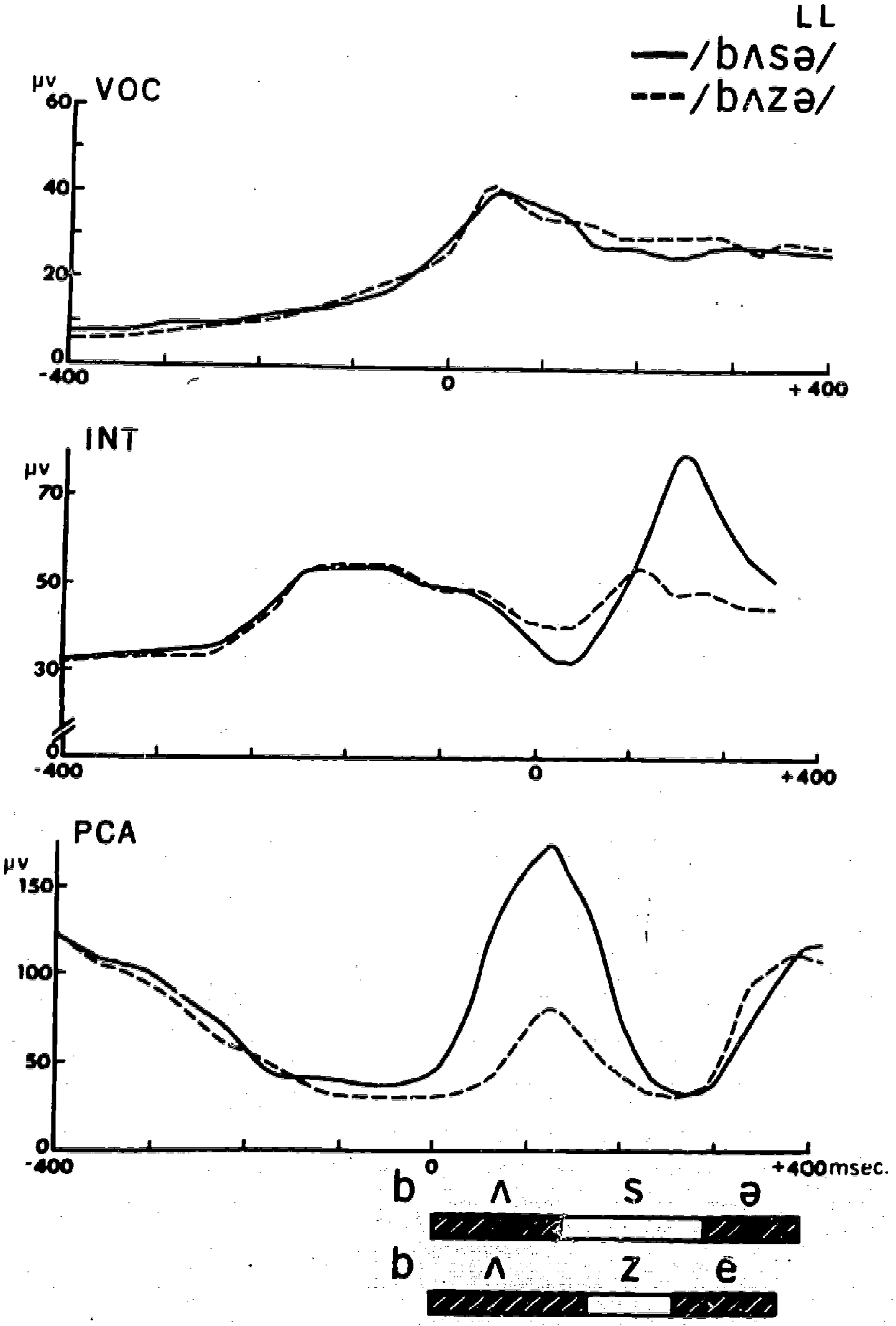


Figure 5: Averaged EMG curves for /bʌsə/ and /bʌzə/ (Subject LL).

For both consonants, the INT shows a sort of reciprocal pattern of activity when compared to the PCA. INT activity begins to increase about 250 msec prior to initial vowel production. For /əpʌp/, activity reaches a peak when the PCA reaches a valley and vice versa. For the articulation of /əbʌp/, the INT shows more or less continuous activity, although there is some decrease in activity for intervocalic /b/ when compared to the preceding vowel segment.

The VOC shows two peaks, each of which appears to correspond to vowel production, with a higher peak for the stressed vowel. This higher peak for the stressed vowel is consistent across all samples regardless of whether the stressed vowel is preceded or followed by the unstressed vowel. Between the two peaks, activity stays low for the intervocalic consonantal segment, regardless of voicing distinction.

Figure 3 shows the medial /p/ vs. /b/ contrast in poststressed position for the same subject. For this condition, each muscle shows essentially the same features as observed in the previous example, that is, the PCA shows increasing activity for the voiceless segment, while INT activity is higher for the voiced segment. The VOC again shows two peaks with the higher one accompanying the stressed vowel.

When we compare the peak EMG values of the PCA for medial /p/ production in two different phonetic conditions (as shown in Figures 2 and 3) activity is higher for prestressed /p/ than poststressed /p/. This is consistent with the findings for another subject in which a comparison was made for three voiceless stops in pre- and poststressed conditions. Here, too, peak PCA activity for the medial voiceless stop production was always higher for the prestressed condition than the poststressed. Further, the duration of PCA activation for voiceless stop production was also found to be longer for prestressed than poststressed conditions.

The data in Figures 2 and 3 also provide some information on the timing relationships between laryngeal and oral articulatory gestures.

In order to compare the timing relationship between the glottal gesture and oral stop closure, three different points on the averaged PCA curve were measured with reference to implosion and release of the stop closure of medial /p/. These points were (1) the point where PCA activity begins to increase for stop production: P_1; (2) the point where the activity reached its peak: P_2; and (3) the point where the activity decreased to its minimum for the production of the postconsonantal vowel: P_3.

Table III shows these time intervals for both the pre- and poststressed conditions. It is shown for both subjects that the time intervals thus specified are always larger for poststressed stops than for prestressed stops. It is worth noting, in particular, that P_3 always occurred earlier than, or synchronously with, stop release for poststressed stops, while it never did so for prestressed stops. In other words, stop closure is released after complete suppression of PCA activity in the case of poststressed stops, while for prestressed stops, stop release occurs before the completion of PCA suppression.

Figure 4 compares the activities of the same three muscles for the prestressed /s/ vs. /z/ contrast in the pair /əsʌp/ vs. /əzʌp/. Here, the

TABLE III. Time intervals for PCA activity in relation to stop closure and release in msec. A negative value indicates stop release preceding complete PCA suppression.

		Interval Between		
		P_1 and Stop Closure	P_2 and Stop Release	P_3 and Stop Release
Subject 1				
/p/	prestressed	110	110	-55
	poststressed	135	165	40
Subject 2				
/p/	prestressed	110	60	-90
	poststressed	150	130	10
/p/	prestressed	70	45	-140
	poststressed	160	95	0
/p/	prestressed	85	40	-165
	poststressed	155	140	30

PCA again shows a large peak for the voiceless consonant, while it is suppressed for the voiced segments. The activity of the INT is, in this case too, higher for the voiced consonant /z/ than for voiceless /s/, but the difference is less marked when compared to that for /b/ and /p/. This is probably because its activity is considerably lower for the consonantal segment of /z/ in comparison to its neighboring vowel segments. This tendency of INT activity to be lower for a voiced fricative than for a vowel is also observed in Figure 5, where the poststressed /s/ vs. /z/ contrast is shown for the pair /bʌsə/ vs. /bʌzə/. It should be further noted in this figure that the PCA shows increasing activity for the segment of /z/ compared to the neighboring vowel segments, the time course of which appears to correspond to a dip in INT activity.

Figures 6 and 7 summarize PCA and VOC activity for the intervocalic voiced/voiceless contrast for Subject LJR. The data points in the middle of each figure indicate the mean of peak EMG values for seven different pairs of voiced and voiceless consonants, while the vertical bar represents the entire range of sample variation. The circles at either end indicate the mean EMG activity at 100 msec prior to and after the peak for each consonant. In Figure 6, it is clearly shown that PCA activity is definitely higher for the production of a voiceless consonant than for a voiced consonant.

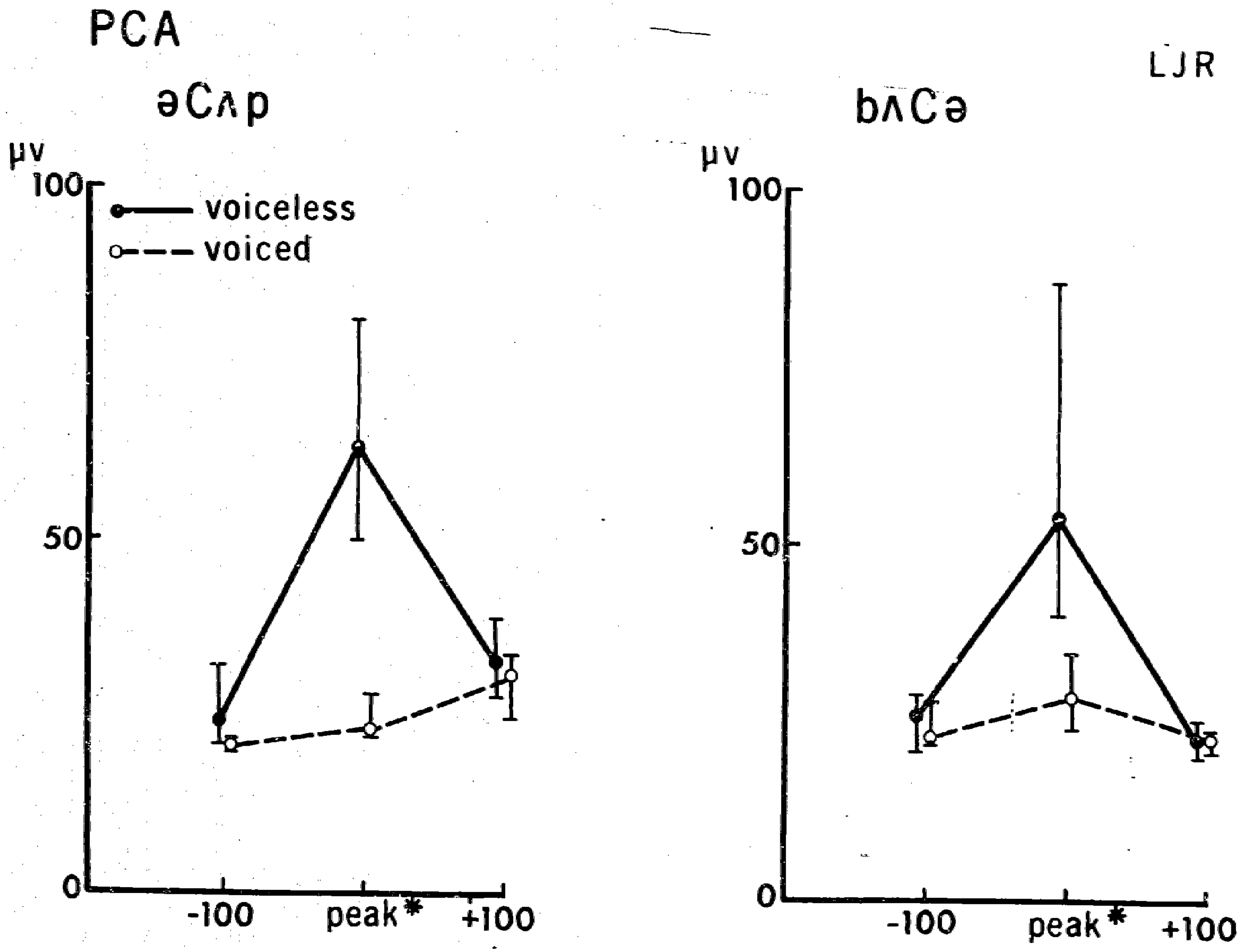


Figure 6. Comparison of PCA activity for the medial voiced and voiceless consonant production (Subject LJR). A peak in activity was clearly present for only the group of voiceless consonants. EMG values for the voiced group were taken at the corresponding time moment.

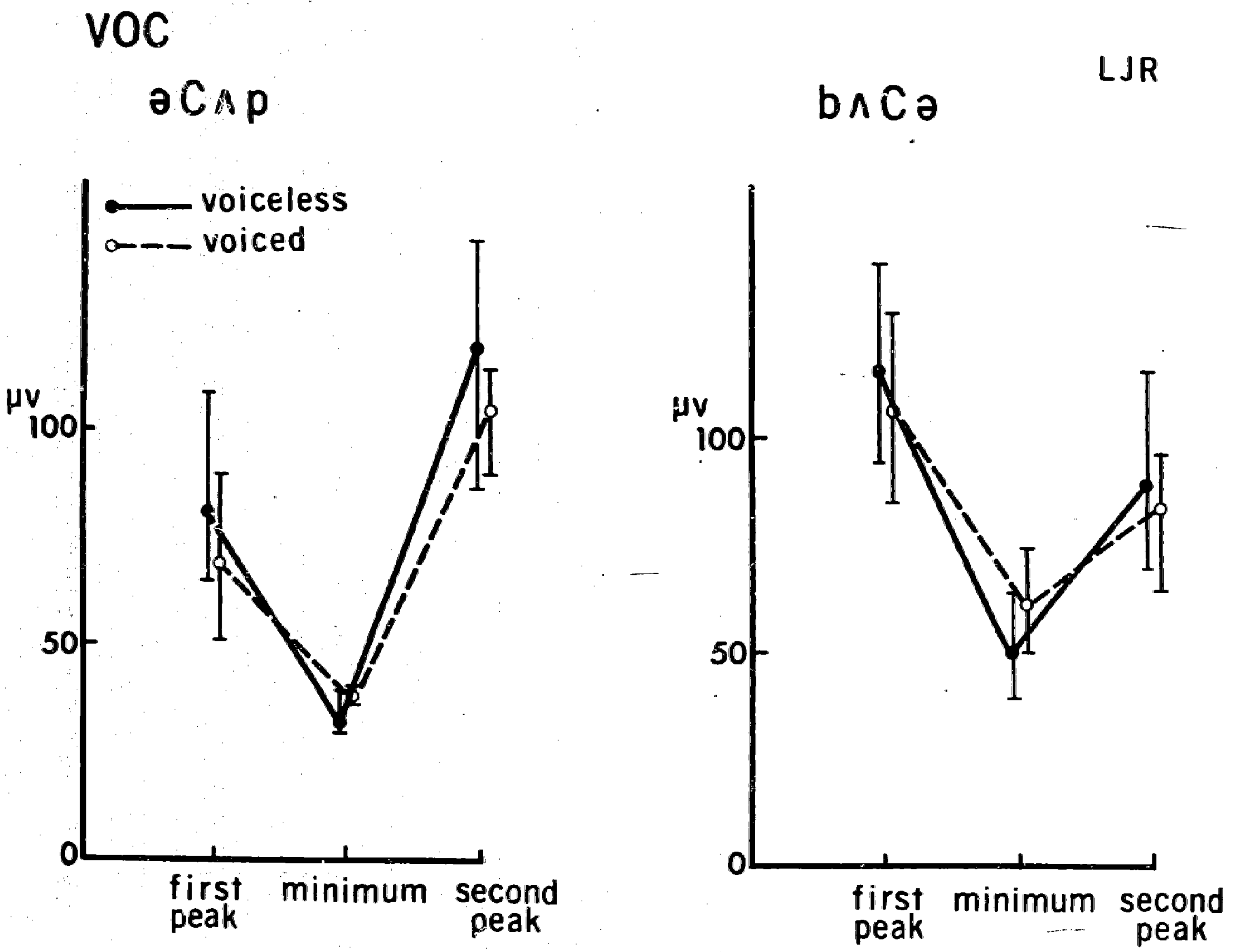


Figure 7: Comparison of VOC activity for medial voiced and voiceless consonant production (Subject LJR).

In the case of the VOC, however, there is no apparent difference in the pattern of activity with respect to the voiced/voiceless contrast. In Figure 7, the end data points indicate the mean of peak EMG values for the vowel segments, while the circles in the middle represent the mean of the minimum values between the two peaks. It is shown that VOC activity is suppressed during the period of consonant production between the two peaks for the vowel segments regardless of the difference in voicing distinction. As far as the peak height for vowel production is concerned, it appears to be higher for a stressed vowel than for an unstressed vowel.

In addition, we have also observed that the pattern of LCA activity appears to be similar to that of the VOC, showing increasing activity only for vowel segments, with a higher peak for the stressed vowel. Its activity decreases for the intervocalic consonantal segment regardless of voiced/voiceless distinction.

Figure 8 shows the averaged activity of the CT for the pairs /əkʌp/ vs. /əgʌp/ and /bʌkə/ vs. /bʌgə/, for Subject LJR. The general pattern of muscle activity is similar for each pair; one large peak is always observed, apparently corresponding to the position of stress in the test word (i.e., where the F_0 contour reaches its peak). There are no discernible differences in the averaged EMG curves with respect to the voiced/voiceless distinction.

Voiced/Voiceless Contrast in Word-Final Position

Figures 9 and 10 show the EMG curves of the PCA and the INT for the /p/ vs. /b/ contrast in the final, poststressed, and postunstressed positions. It is apparent in these figures that the PCA shows high activity for the voiceless consonant, during which time INT activity is suppressed. Conversely, PCA activity is continuously suppressed when the interconsonantal vowel is followed by final /b/, at which time the INT shows higher activity. In addition to the final rise, there is also a slight ascent in PCA curves in both these examples, apparently associated with initiation of the stressed vowel.

In Figure 11, PCA activity for Subject LJR is schematically shown during the time period including the final consonantal segment. As before, averaged EMG values are compared here for voiced and voiceless pairs at three time moments: at the line-up point (time 0) and 100 msec and 200 msec after the line-up, as given on the abscissa. The values in the figures again respresent mean EMG values for seven different kinds of consonants. Both graphs clearly show that PCA activity is higher for the final voiceless consonants.

VOC activity is likewise compared in Figure 12, where averaged EMG values were taken at the time when the EMG curve reaches its second peak[2] and 100 msec and 200 msec thereafter. Both graphs show that VOC activity is higher for the final voiced consonants than for the voiceless conson-

---

[2] The VOC and the LCA show two peaks in the EMG curves for these test words, each of which appears to correspond to vowel production. The second peak thus specifies the EMG peak for the vowel preceding the final consonant.

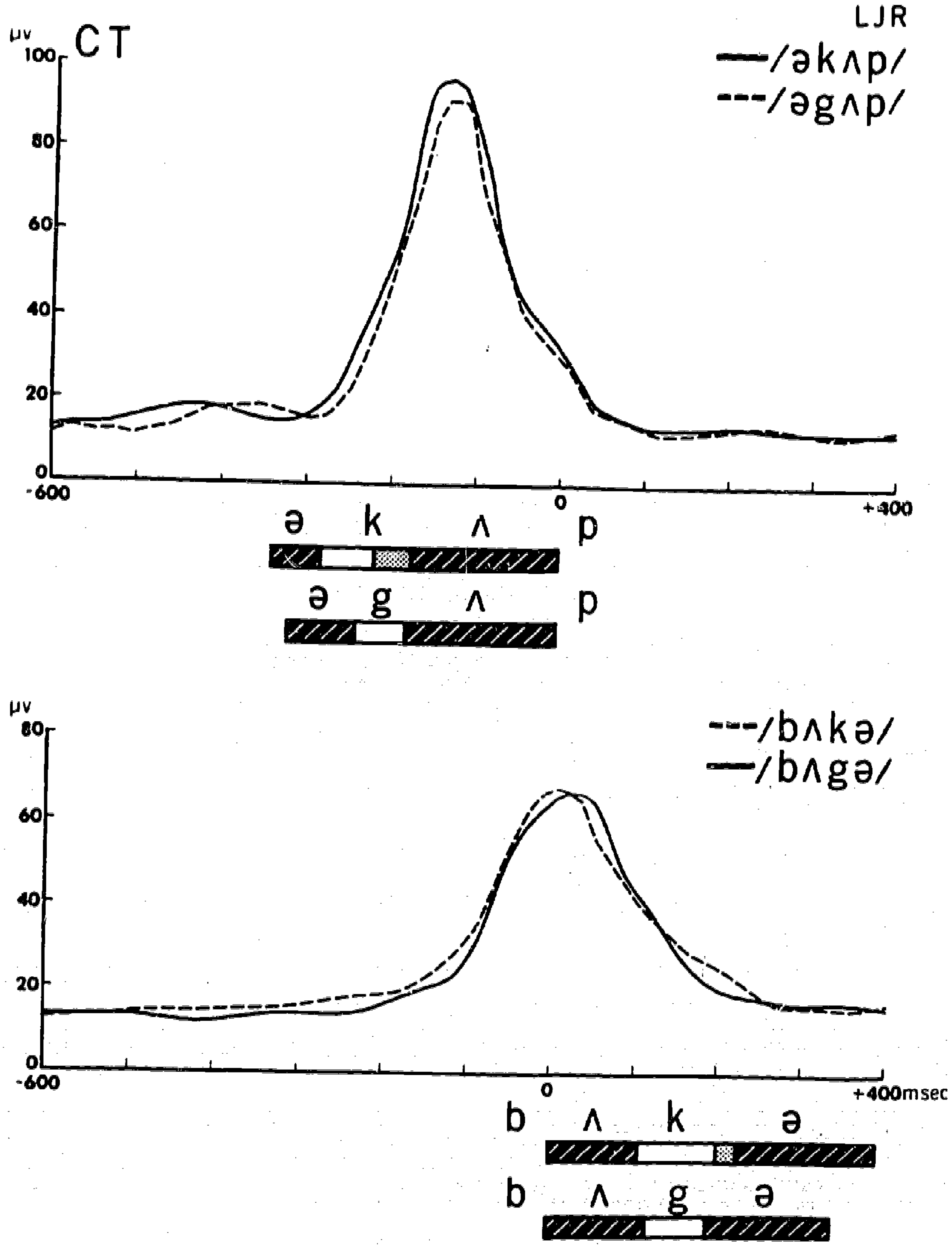


Figure 8: Superimposed averaged EMG values of CT activity for the pairs /əkʌp/ vs. /əgʌp/ and /bʌkə/ vs. /bʌgə/ (Subject LJR).

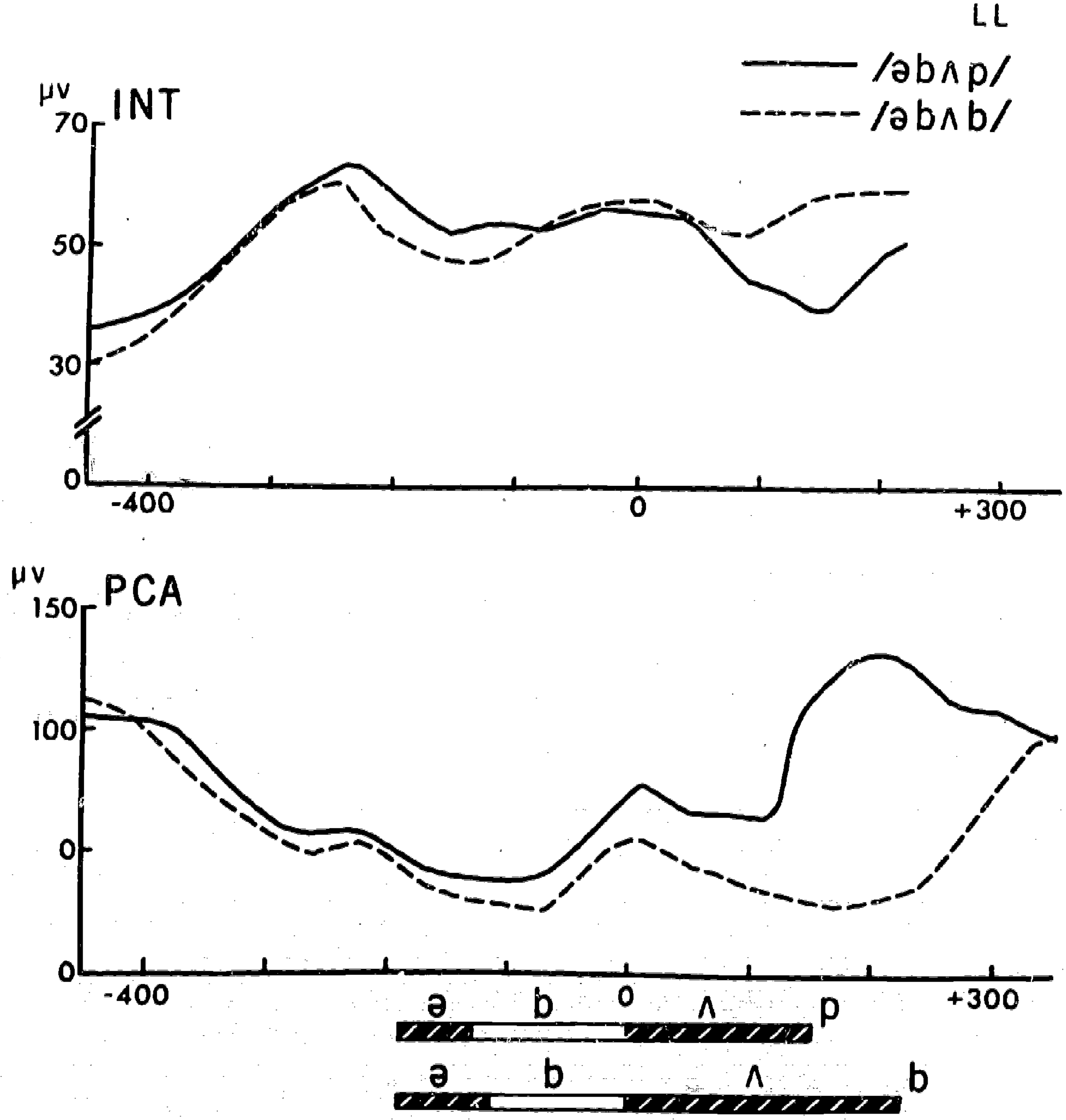


Figure 9: Superimposed averaged EMG curves of INT and PCA activity for the utterances /əbʌp/ and /əbʌb/ (Subject LL). The line-up point is the onset of the stressed vowel.

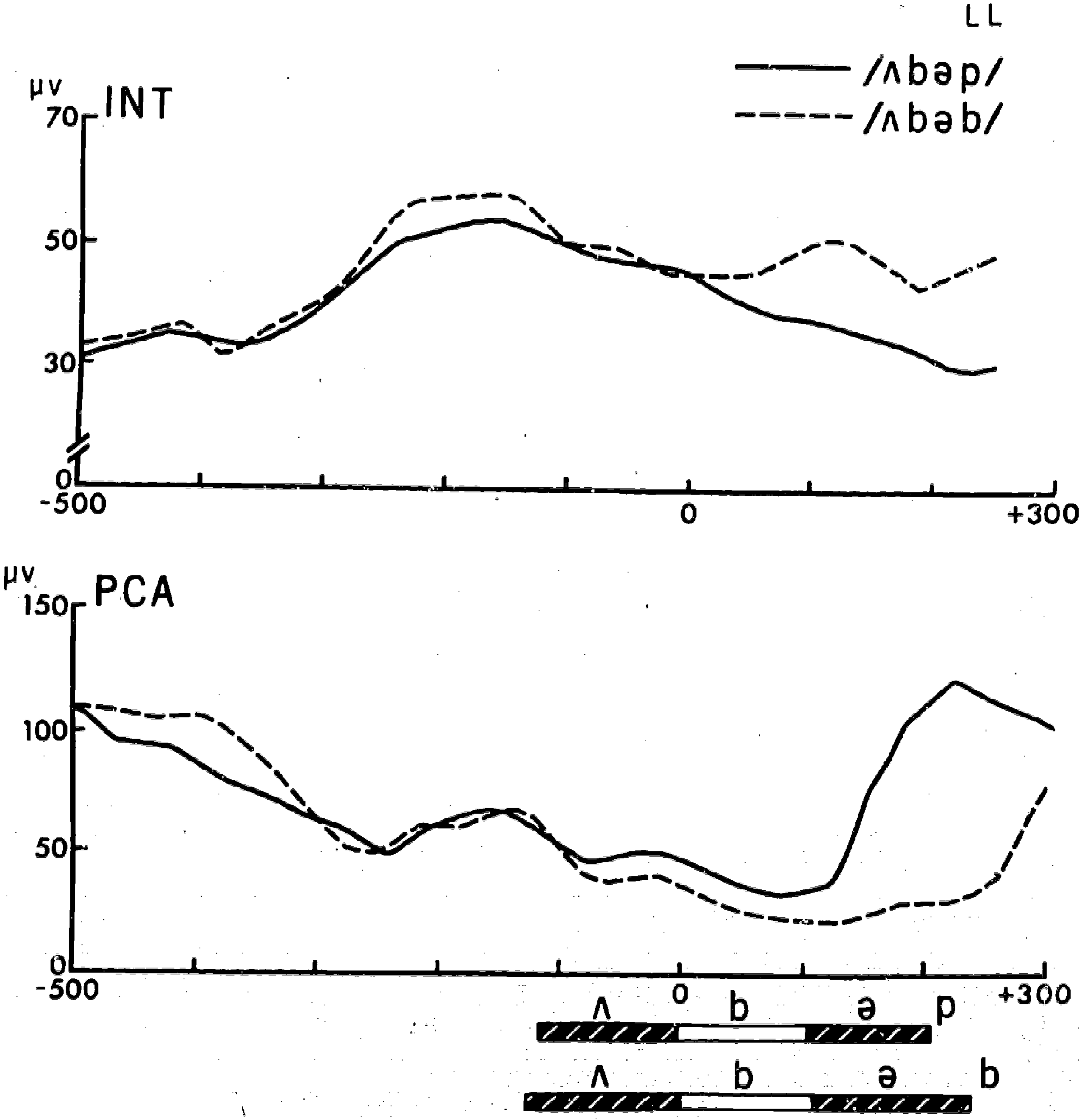


Figure 10: Averaged EMG curves for /ʌbəp/ and /ʌbəb/ (Subject LL).

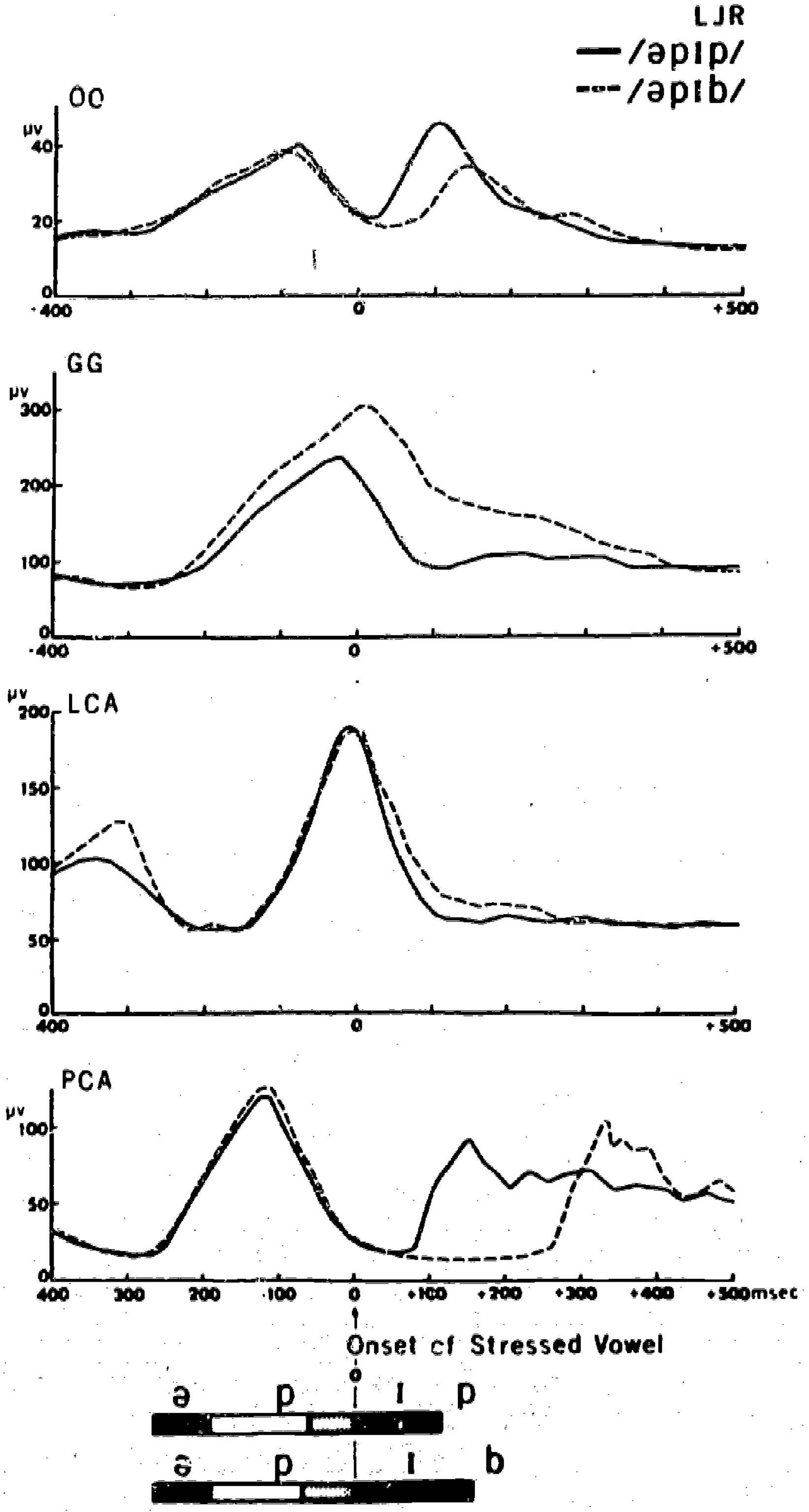


Figure 11: Comparison of PCA activity for final voiced and voiceless consonant production (Subject LJR). "0" on the abscissa indicates the line-up point.

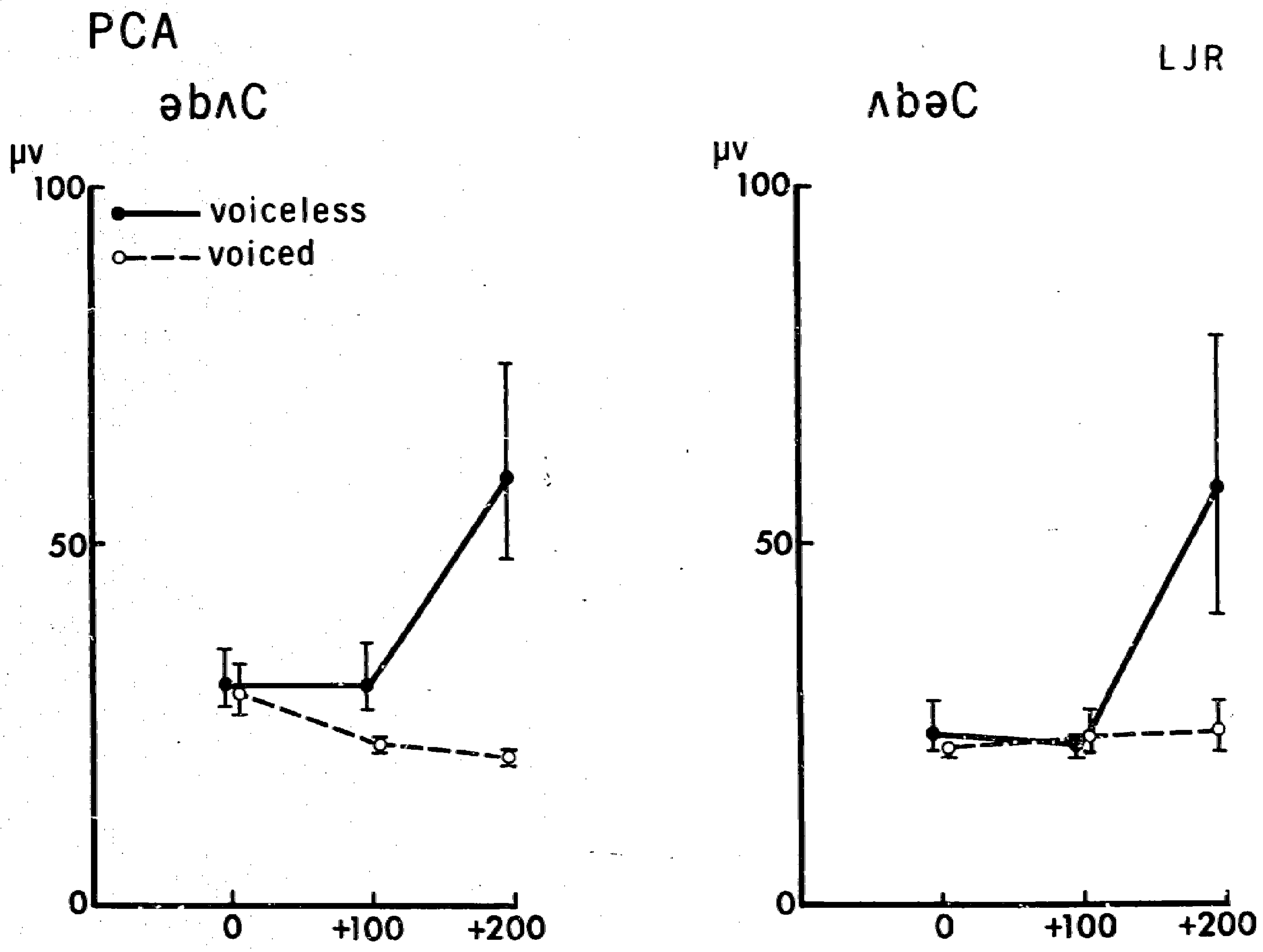


Figure 12: Comparison of VOC activity for final voiced and voiceless consonant production (Subject LJR).

ants. This same tendency was observed for LCA activity, which also appeared to be higher for the voiced pairs.

It has often been observed that English vowels are of greater duration before voiced than before voiceless consonants. Thus, there is a possibility that the higher VOC or LCA activity for the final voiced consonant is an effect of the preceding vowel segment. In other words, the higher VOC or LCA activity levels might be associated with greater vowel duration rather than with any distinctive consonant feature.

In order to examine this possibility, the activity of other articulatory muscles (the genioglossus and the orbicularis oris) were later recorded simultaneously with LCA and PCA activity. The genioglossus (GG) is one of the extrinsic lingual muscles responsible for /ɪ/ production, while the orbicularis oris (OO) is important for lip closure. Data for all four muscles are shown in Figure 13. It is clearly shown in this figure that the duration of the vowel /ɪ/ preceding the final consonant is greater for /əpɪb/ than for /əpɪp/ and that GG activity stays higher for the former than for the latter. The OO shows two peaks for the medial and the final bilabial stops and the interval between the two peaks indicates the duration of /ɪ/, which is longer for /əpɪb/. These findings suggest that the duration of muscle activity for /ɪ/ is longer for /əpɪb/ than for /əpɪp/. If we attempt to slide the EMG curve of the LCA for /əpɪb/ to the left on the abscissa in order to synchronize the end of the vowel /ɪ/ with that of /əpɪp/, the descending portions of the two LCA curves will be almost superimposed together. Thus, it seems reasonable to consider that the apparently higher LCA activity near the end of the test words for /əpɪb/ in Figure 13 corresponds to the vowel /ɪ/ preceding the final /b/. However, PCA activity stays higher for /əpɪp/ and is suppressed for /əpɪb/ near the end of the test words even when the sliding of the EMG curves is attempted as above. Therefore, it can still be concluded that PCA activity is higher for a voiceless consonant than for a voiced consonant, even in final position.

Voiced/Voiceless Contrast in Word-Initial Position

Comparisons of EMG activity levels for a voiced/voiceless consonant pair in initial position were made only for the pair /pʌpə/ vs. /bʌpə/, the results of which are shown in Figure 14.

For /pʌpə/, PCA activity stays higher before lip closure and then decreases steeply approximately 110 msec before the onset of /ʌ/. An increase then follows for the medial /p/. INT activity shows a steep rise when the PCA shows the steep fall. The same tendency is seen in VOC activity, which also shows a steep rise but which starts somewhat later than the INT.

DISCUSSION

Functional Characteristics of the Individual Laryngeal Muscles in Articulatory Adjustments

The posterior cricoarytenoid (PCA). It was revealed in the present study that the PCA actively participates in laryngeal articulatory adjustments, particularly for the voiced/voiceless distinction. There is a consistent

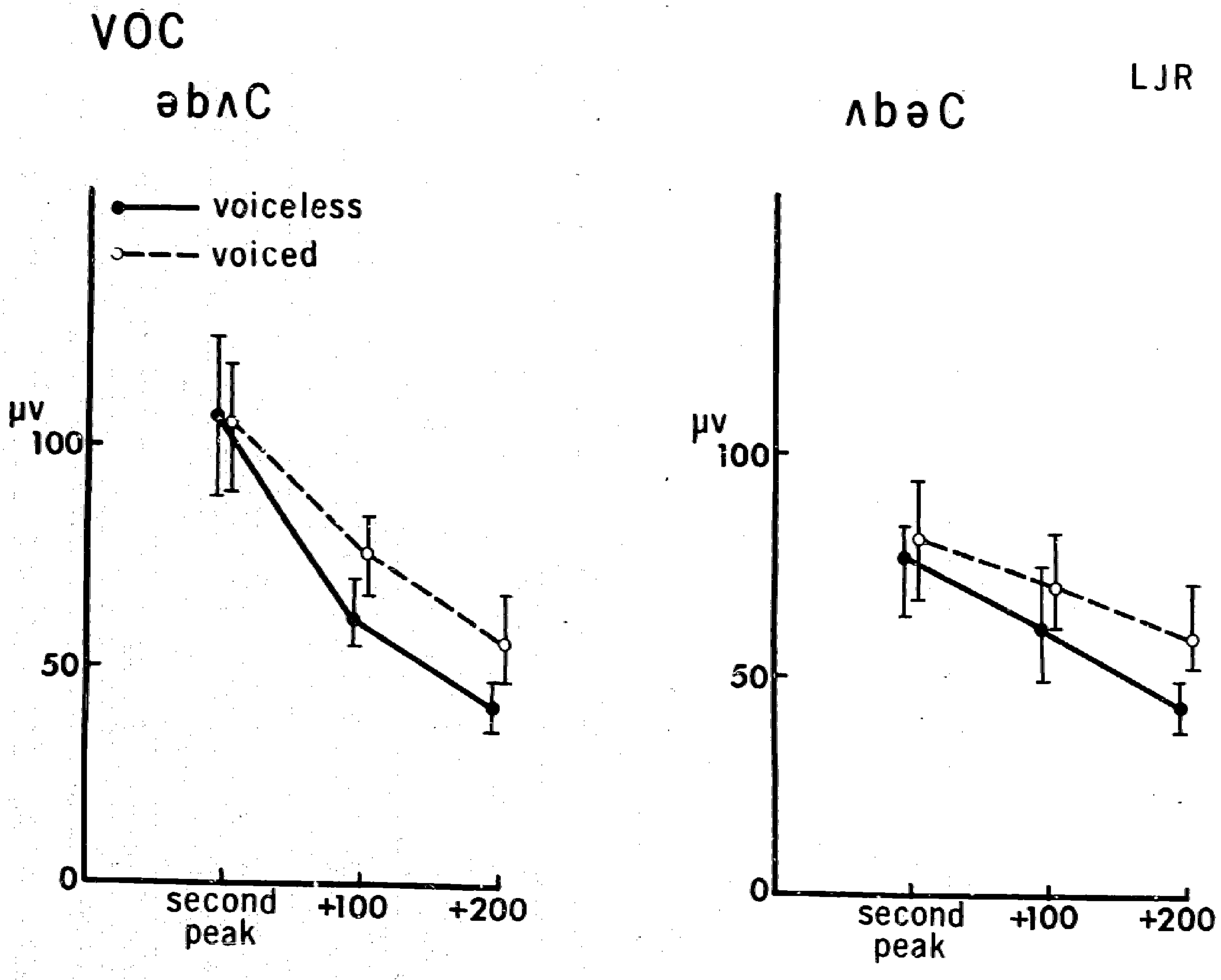


Figure 13: Superimposed averaged EMG curves of the OO, GG, LCA and PCA for /əpɪp/ and /əpɪb/ (Subject LJR).

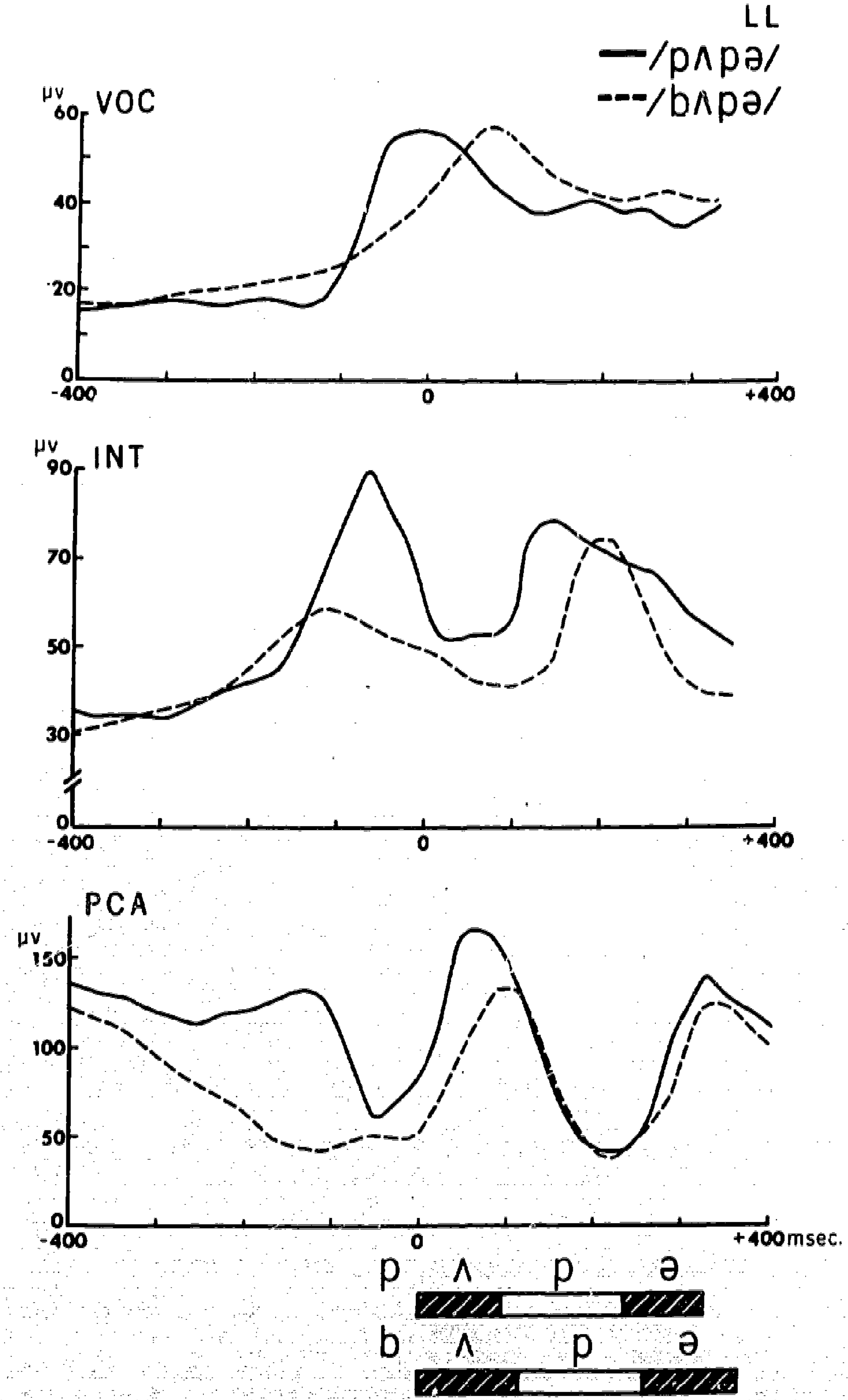


Figure 14: Superimposed averaged EMG curves of the VOC, INT, and PCA for /pʌpə/ and /bʌpə/ (Subject LL). The line-up point is the onset of the stressed vowel.

increase in PCA activity for voiceless consonant production regardless of phoneme environment. For all utterance types PCA activity shows a transient increase before the onset of phonation, presumably for prephonatory inspiration. Its activity then starts to decrease for initial vowel production, unless a voiceless consonant is placed in the absolute initial position of the test utterance. If a voiceless consonant is in initial position, PCA activity stays at about the level of the initial prephonatory rise or even higher (Ref. Figure 14: /pʌpə/). For the production of a voiceless consonant in the medial or final position, the PCA always shows a marked increase in activity before the onset of consonant closure. As far as the voiced/voiceless distinction in final position is concerned, PCA activity appears to be significantly higher for a voiceless cognate, even if differences in the duration of the preceding vowel are taken into consideration as another possible cause of prolonged PCA suppression.

It should also be mentioned that the PCA does not show only a simple all-or-none pattern of activity but rather shows a pattern of fine adjustment. As seen in Figure 5, the PCA shows partial activation for the production of the poststressed voiced fricative /z/, which seems to indicate a less complete glottal closure. In a transillumination study of the larynx, Lisker, Abramson, Cooper, and Schvey (1969) found that a high percentage of voiced fricatives were produced with at least a partially open glottis. Incomplete closure of the glottis during voiced fricative production can be obtained either by partial activation of the PCA or slight suppression of the adductors, particularly the INT. In the case of the poststressed /z/ mentioned above, both factors appear to work together, while in the case of the prestressed /z/ (as in Figure 4), suppression of INT activity is more manifest.

Another interesting finding is the small PCA peak just before the onset of an initial or medial stressed vowel (Figures 2, 4, 9, and 10). Interpretation of this transient PCA activity is not perfectly clear as yet, but it is conceivable that the PCA acts to counterbalance the strong contraction of the adductors at the onset of the stressed vowel. In a study of the EMG activity of the laryngeal muscles in phonation (Gay, Hirose, Strome, and Sawashima, in press), we observed that PCA activity is generally suppressed for sustained phonation, except for an increase at the highest range in chest register. The increasing PCA activity in that extreme condition may reflect the counterbalancing function of the abductor for the strong contraction of the adductors, as suggested in the literature (Pressman, 1942; Suzuki and Kirchner, 1969). Another possibility is that functionally different motor units are participating in the execution of muscle contractions during different types of phonation, since there is evidence, at least in animal experiments, that the PCA contains several kinds of motor units (Suzuki and Kirchner, 1969).

Although the function of the PCA, particularly during sustained phonation, should be a subject for further investigation, the role of the PCA as a pure adductor in speech articulation is well demonstrated in the present study.

<u>The interarytenoid (INT)</u>. The present data indicate that there is apparent reciprocal activity between the PCA and the INT. In this sense, the INT can be considered to be a pure adductor of the vocal folds.

In general, there is an apparent difference in the degree of INT activity for vowel segments depending on the preceding consonant. More specifically, INT activity for the production of a postconsonantal vowel appears to be higher after a voiceless consonant than after a voiced consonant (Figures 2, 3, and 5). Since EMG activity represents the muscle activity necessary for obtaining effective force and/or displacement, the degree of the activity of a given muscle can also be higher if, for example, the displacement is greater. Since glottal width is larger during the articulation of a voiceless consonant than for a voiced consonant (Sawashima, 1968; Sawashima, Abramson, Cooper, and Lisker, 1970), it is reasonable to assume that the activity of the INT, which is responsible for adducting the vocal folds, should be greater after a voiceless consonant.

As seen in Figure 4, INT activity is apparently lower for voiced consonants, namely fricatives, than for vowels. This would also indicate that glottal closure is likewise less tight for voiced consonants than for vowels.

The vocalis (VOC) and the lateral cricoarytenoid (LCA). The VOC and the LCA are considered to have complex functions in laryngeal articulatory adjustments. Both muscles appear to be activated for the vowel segment of the test words but rather suppressed for the consonantal segment, regardless of the voiced/voiceless distinction. It is conceivable, therefore, that the apparent glottal closure usually observed during the production of voiced obstruents can be achieved without increasing the activity of either the VOC or the LCA. Or, one can also argue that glottal closure during voiced obstruent production is less tight because of the absence of increased VOC and LCA activity. In any case, the function of the VOC and LCA as adductors seems different from that of the INT.

For the articulation of the vowel segment, both the VOC and LCA show higher activity for a stressed vowel than for an unstressed vowel, regardless of whether the stressed vowel is preceded or followed by the unstressed. This finding suggests that these two muscles participate in the control of the suprasegmental features as well, possibly in pitch raising. It has been reported that the VOC and the LCA participate in the mechanism of pitch rise (Hirano, Vennard, and Ohala, 1970), particularly when the activity of these two muscles increases simultaneously with the cricothyroid. In this sense, the VOC and the LCA can also be considered to function as tensors of the larynx, although in the case of singing, these two muscles do not seem to be contributing equally to pitch regulation (Gay, Hirose, Strome, and Sawashima, in press).

In an EMG study of vowel devoicing in Japanese, Hirose (1971a) postulated the possibility of functional differentiation between the VOC and the LCA. The present study, however, does not seem to substantiate this differentiation but rather shows fairly similar patterns of EMG activity for these two muscles, at least for those utterance types examined.

The cricothryoid (CT). The CT shows a temporary increase in activity for a stressed vowel but does not seem to participate in the voiced/voiceless distinction. This was not unexpected, as the CT is universally considered as a prime pitch raiser (Arnold, 1961; Gårding, Fujimura, and Hirose, 1970; Simada and Hirose, 1971).

Coordination and Timing of Muscle Activities

It is conceivable in the living human that most of the articulatory muscles are activated in a well-coordinated fashion during normal speech production. More specifically, some muscles behave in reciprocal fashion, while others are synergetic, depending on the particular articulatory condition.

As far as segmental features of the present test words are concerned, the PCA and the INT show consistent reciprocity in both the level of EMG activity and the timing of activation.

It is also worthy to note that timing relationships between laryngeal muscle activity and supraglottal articulatory events vary, depending on phoneme environment (Table III). It has also been reported that in the case of unaspirated voiceless stops, the arytenoids resume a closed position just after oral release, while for aspirated stops, arytenoid closure is completed well after oral release (Lisker, Sawashima, Abramson, and Cooper, 1970; Sawashima, 1970). This is coherent with the present EMG data where suppression of PCA activity is not yet complete at the moment of oral release in the case of prestressed voiceless stop production (suggesting that the glottis remains at least partially open at that moment), while in post-stressed stops, PCA suppression is complete before oral release.

The timing relationships found here are also relevant to more general questions concerning the nature of timing control in speech articulation, i.e., are the observable differences in voice onset times the consequence of other physical and physiological features such as subglottal pressure, glottal aperture, etc. (Chomsky and Halle, 1968; Kim, 1970) or a separate independent physiological mechanism (Abramson and Lisker, 1970; Lisker and Abramson, 1971)?

If timing differences are responses of the system to forces other than direct muscular control, we would expect that the timing of muscle activity patterns would be the same across various contrasts. In other words, the gestures would be organized in the same way but differentially modified according to prevailing glottal conditions.

Our data, though, do not support this concept but rather show differences in the relative timing of muscle activity patterns and, thus, active muscular control of glottal configuration. In other words, our data would suggest the ubiquity of an independent timing control mechanism. At the same time, however, there is the possibility that other laryngeal features are, themselves, independently controlled.

The degree of overall activity of the PCA appears to be higher for prestressed than for poststressed voiceless stops. This finding agrees with both fiberoptic (Sawashima, Abramson, Cooper, and Lisker, 1970) and transillumination data (Lisker, Abramson, Cooper, and Schvey, 1969), which indicate that the degree of glottal opening is greater for the prestressed voiceless stops than for the poststressed.

Based on the acoustical and mechanical aspects of vocal cord vibration, Halle and Stevens (1971) proposed a scheme of laryngeal features to classify certain obstruents, glides, and vowels. They postulated that there are two independently controlled parameters: the stiffness of the vocal cords (adjusted by the thyroarytenoid and the cricothyroid) and statis glottal opening. These two parameters yield four features: spread glottis, constricted glottis, stiff vocal cords, and slack vocal cords. In addition, nine distinct phonetic categories can result from combinations of these four features.

Although an EMG model might not be the only analog of such a feature system, muscle contraction properties are certainly important correlates. Assuming, then, a relationship between "stiffness" and muscle activity, our present data do not support their system with respect to certain points. For example, Halle and Stevens postulated the [+ stiff] feature for both the voiceless unaspirated stop [p] and the voiceless aspirated stop [p^h]. However, the present data show that the CT, VOC, and LCA are suppressed for the production of these consonants. Thus, there is no EMG evidence, in the form of increased CT, VOC, or LCA activity, to support the concept of [+ stiff] vocal cords for the production of voiceless obstruents. Further, the proposed feature of [- spread] glottis for the voiceless unaspirated stop [p] is not supported by the present data either, since this consonant is associated with high PCA activity and suppressed INT activity for an open glottis.

Although the present data are quite straightforward, it is obvious that more extensive experiments, including a combined EMG-fiberoptic approach are needed to provide further information on the relationships among muscle activity, glottal configuration, and distinctive features.

REFERENCES

Abramson, A.S. and L. Lisker. (1970) Laryngeal behavior, the speech signal and phonological simplicity. Actes du X^{e} Congrès International des Linguistes, IV (Bucarest: Academie de la Republique Socialiste de Roumanie) 124-128.

Arnold, G.E. (1961) Physiology and pathology of the cricothyroid muscle. Laryngoscope 71, 687-753.

Basmajian, J.V. and G. Stecko. (1962) A new bipolar electrode for electromyography. J. appl. Physiol. 17, 849.

Chomsky, N. and M. Halle. (1968) The Sound Pattern of English. (Harper & Row, New York).

Dedo, H. (1970) The paralyzed larynx: An electromyographic study in dogs and humans. Laryngoscope 80, 1455-1517.

Faaborg-Andersen, K. (1957) Electromyographic investigation of intrinsic laryngeal muscles in humans. Acta physiol. Scand. 41, Suppl. 140.

Gårding, E., O. Fujimura, and H. Hirose. (1970) Laryngeal control of Swedish word tones: A preliminary report on an EMG study. Annual Bulletin (Research Institute of Logopedics and Phoniatrics, University of Tokyo) No. 4, 45-54.

Gay, T., H. Hirose, M. Strome, M. Sawashima. (In press) Electromyography of the intrinsic laryngeal muscles during phonation. Annals of Otology, Rhinology and Laryngology.

Halle, M. and K.N. Stevens. (1971) A note on laryngeal features. QPR MIT 101, 198-213.

Hirano, M. and J. Ohala. (1969) Use of hooked-wire electrodes for electromyography of the intrinsic laryngeal muscles. JSHR 12, 362-373.

Hirano, M., W. Vennard and J. Ohala. (1970) Regulation of register, pitch and intensity of voice. Folia phoniat. 22, 1-20.

Hirose, H. (1971a) The activity of the adductor laryngeal muscles in respect to vowel devoicing in Japanese. Phonetica 23, 156-170.

Hirose, H. (1971b) Electromyography of the articulatory muscles: Current instrumentation and technique. Haskins Laboratories Status Report on Speech Research SR-25/26, 73-86.

Hirose, H., T. Gay, and M. Strome. (1971) Electrode insertion technique for laryngeal electromyography. J. acous. Soc. Amer. 50, 1449-1450.

Hiroto, I., M. Hirano, Y. Toyozumi, and T. Shin. (1967) Electromyographic investigation of the intrinsic laryngeal muscles related to speech sounds. Ann. Otol. Rhinol. Laryng. 76, 861-872.

Kim. C.W. (1970) A theory of aspiration. Phonetica 21, 107-116.

Kotby, M.N. and L.K. Haugen. (1970) Critical evaluation of the action of the posterior crico-arytenoid muscle, utilizing direct EMG-study. Acta Otolaryng. 70, 260-268.

Lisker, L., A.S. Abramson, F.S. Cooper, and M. Schvey. (1969) Transillumination of the larynx in running speech. J. acous. Soc. Amer. 45, 1544-1546.

Lisker, L., M. Sawashima, A.S. Abramson, and F.S. Cooper. (1970) Cinegraphic observations of the larynx during voiced and voiceless stops. Haskins Laboratories Status Report on Speech Research SR-21/22, 201-210.

Lisker, L. and A.S. Abramson. (1971) Distinctive features and laryngeal control. Language 47, 767-785.

Pressman, J.J. (1942) Physiology of the vocal cords in phonation and respiration. Arch. otoloaryng. 35, 335-398.

Sawashima, M. (1968) Movements of the larynx in articulation of Japanese consonants. Annual Bulletin (Research Institute of Logopedics and Phoniatrics, University of Tokyo) No. 2, 11-20.

Sawashima, M. (1970) Glottal adjustments for English obstruents. Haskins Laboratories Status Report on Speech Research SR-21/22, 187-200.

Sawashima, M., A.S. Abramson, F.S. Cooper, and L. Lisker. (1970) Observing laryngeal adjustments during running speech by use of a fiberoptic system. Phonetica 22, 193-201.

Simada, Z. and H. Hirose. (1971) Physiological correlates of Japanese accent patterns. Annual Bulletin (Research Institute of Logopedics and Phoniatrics, University of Tokyo) No. 5, 41-49.

Suzuki, M. and J.A. Kirchner. (1969) The posterior cricoarytenoid as an inspiratory muscle. Ann. Otol. Rhinol. Laryng. 78, 849-864.

Velopharyngeal Function in Oral/Nasal Articulation and Voicing Gestures[*]

Fredericka B. Berti[+] and Hajime Hirose[++]
Haskins Laboratories, New Haven

The aim of the electromyographic study reported here is to describe oral and nasal articulatory gestures and patterns of velopharyngeal activity which accompany voicing distinctions among stop consonants.

Nonsense disyllables were designed to place maximum stress on the mechanisms of oral and nasal articulation, that is, to have strongly oral consonants preceded by nasals and nasal consonants preceded by strongly oral consonants, with both conditions observed in varying vowel environments. In addition, voicing contrasts were included. Figure 1 presents the format of the stimuli (e.g., /fimkip/, /futmup/, /faŋbap/).

METHODS

Peroral insertions of bipolar, hooked-wire electrodes were made into the dimple of the levator palatini, the superior constrictor at the estimated level of velopharyngeal closure, the middle constrictor at the level of the epiglottis, the palatopharyngeus (which is considered to be the muscular component of the posterior faucal pillar), and the palatoglossus (which is the muscular component of the anterior faucal pillar). Percutaneous insertions were made into the sternohyoid at the level of the thyroid lamina and the orbicularis oris upper at the vermillion border (Hirose, 1971).

The EMG potentials, along with the audio signal and automatic timing markers, were recorded onto magnetic tape. The potentials were rectified, integrated, and computer averaged, using the data-processing system described by Port (1971). Ten to sixteen tokens of each utterance type were averaged for each subject. The line-up point selected for averaging was the termination of /m/ or /ŋ/ when it occurs as C_1 and the initiation of /m/ when it occurs as C_2. This point is labeled "0" on the abscissa; voice onset of V_1 and offset of V_2 are indicated by arrows in the figures.

RESULTS

The Oral/Nasal Distinction

The same general pattern of activity is found in the levator palatini, superior and middle constrictors, and the palatopharyngeus for oral

---

[*] Paper presented at the 1971 Convention of the American Speech and Hearing Association, Chicago, Illinois, November 1971.

[+] Also The Graduate Center, The City University of New York.

[++] Also Faculty of Medicine, University of Tokyo.

$$f\ V_1C_1C_2V_2p$$

$V_1 = V_2 =$ /i/, /a/, or /u/

and

1. $C_1 =$ /m,ŋ/ and $C_2 =$ /p,t,k,b,d,g/

<u>or</u>

2. $C_1 =$ /p,t,k,b,d,g/ and $C_2 =$ /m/

Fig. I

articulation. We will use the palatopharyngeus as a representative of this group of muscles in the succeeding discussion.

Inspection of the averaged EMG curves for the palatopharyngeus reveals peaks of activity which correspond to stop consonant production (Figure 2). The peak is more distinct when the stop immediately follows a nasal consonant, as in the utterance /famdap/, than when it precedes a nasal consonant, as in the utterance /fadmap/. That is, greater EMG activity is recorded when a strong oral gesture follows a nasal one (/famdap/), than when a strong oral gesture follows the gesture of another oral phone (/fadmap/).

A decrease in activity may be seen accompanying nasal gestures. The lowest point for /famdap/ occurs at about -175 msec, or well before the zero point that indicates the end of the acoustic signal of the nasal consonant. The equivalent depression in /fadmap/ is found near zero, which is the beginning of the acoustic signal of the nasal consonant.

Electromyographic activity peaks are similarly found for oral gestures in the levator palatini (Figure 3), superior constrictor (Figure 4), and middle constrictor (Figure 5), while nasal gestures are accompanied by reductions in the electromyographic signals recorded from these muscles. This description of the oral gesture is in general agreement with that of Fritzell (1969), who found similar patterns for the levator palatini and superior constrictor. Fritzell, however, described the nasal gesture differently: he found reductions in levator palatini and superior constrictor activity, as we have, but he also described increased palatoglossus activity for the nasal gesture. Our data, on the other hand, indicate that the palatoglossus does not participate in the nasal gesture but rather is active for tongue-backing and tongue-raising maneuvers.

Examination of the averaged curves of Figure 6 reveals peaks for each /u/ produced in the two utterance types displayed, /fumkup/ and /fukmup/. There is a separate peak for /k/ in /fumkup/. This middle peak is not associated with the nasal gesture, a point which is made clear by inspection of the curve for /fukmup/. Here we see that the palatoglossus activity which occurs for /u/ continues to increase into /k/, and that the activity then drops off abruptly at the time of nasal articulation, only to begin rising again 100 msec later, for the production of the second /u/. In addition, and contrary to the findings of Fritzell, little or no palatoglossus activity has been found for the production of /ŋ/ for either of the two subjects for whom data are available.

There are several conclusions to be drawn from the data presented. First, the levator palatini, superior constrictor, middle constrictor, and palatopharyngeus are all active for oral gestures and show decreased activity for the production of nasal consonants. Palatoglossus activity appears to be correlated with both tongue backing and tongue raising, with no evidence of activity for nasal or oral gestures. This latter conclusion differs from the conclusion of Fritzell, and also of Lubker, Fritzell, and Lindqvist (1970) that nasalization is accomplished by active lowering of the soft palate by the palatoglossus muscle. Rather, it appears from our data that the nasal gesture is a passive one, with palatal lowering resulting from a combination of reduced contraction of the levator palatini and,

Palatopharyngeus

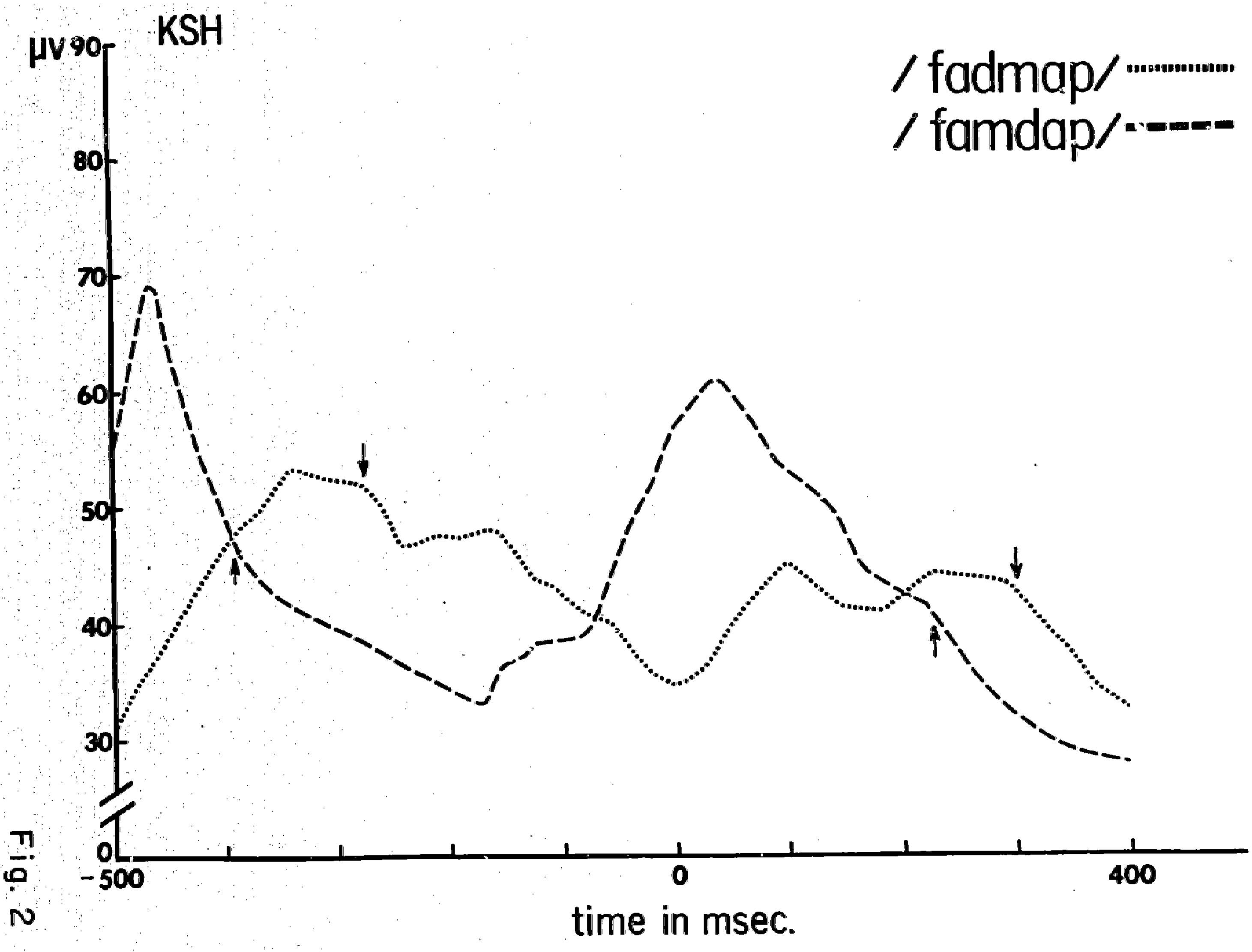


Fig. 2

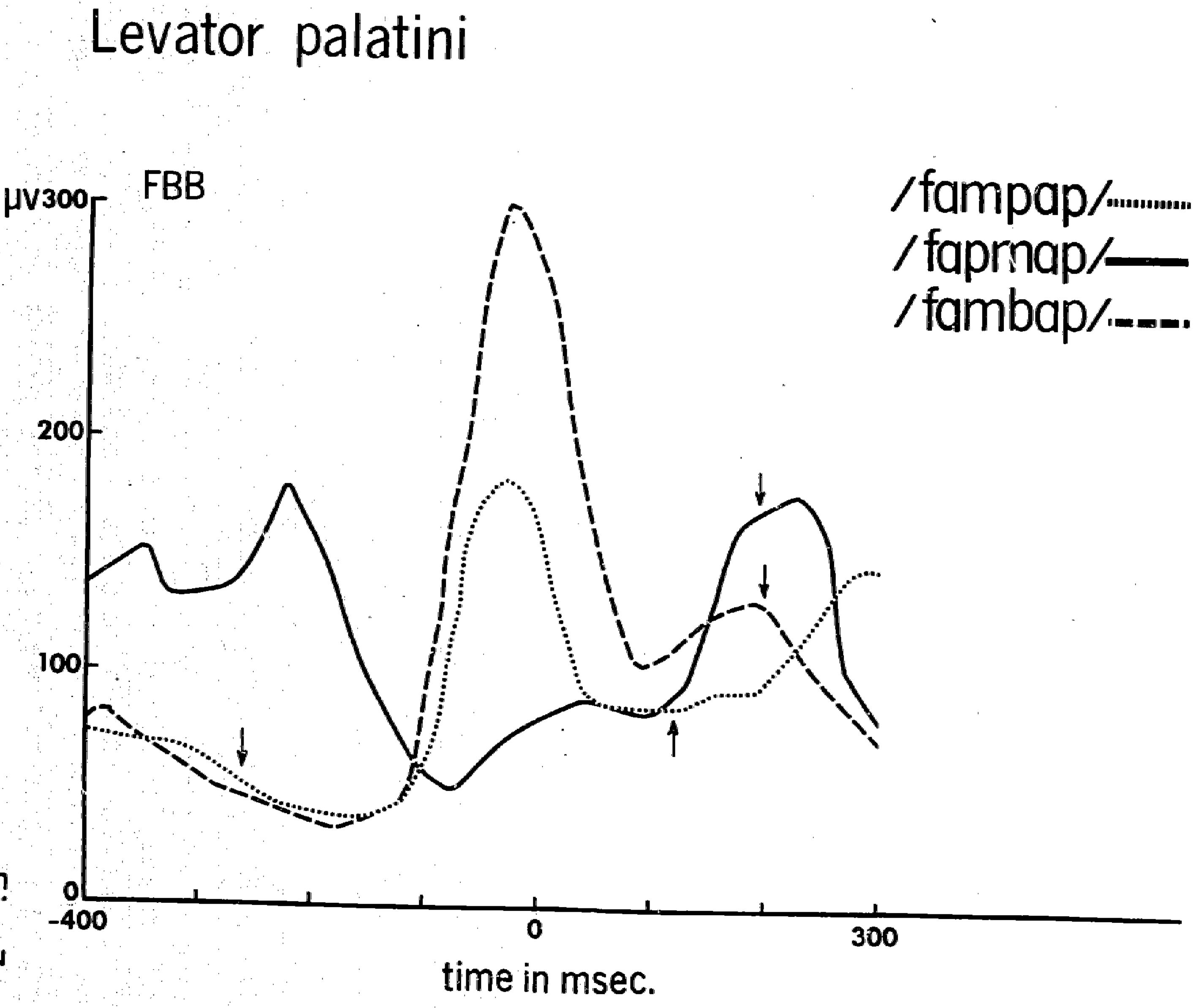

Levator palatini
FBB
μV300
200
100
0
-400
0
300
time in msec.
/fampap/
/fapmap/
/fambap/

Fig. 3

Superior Constrictor

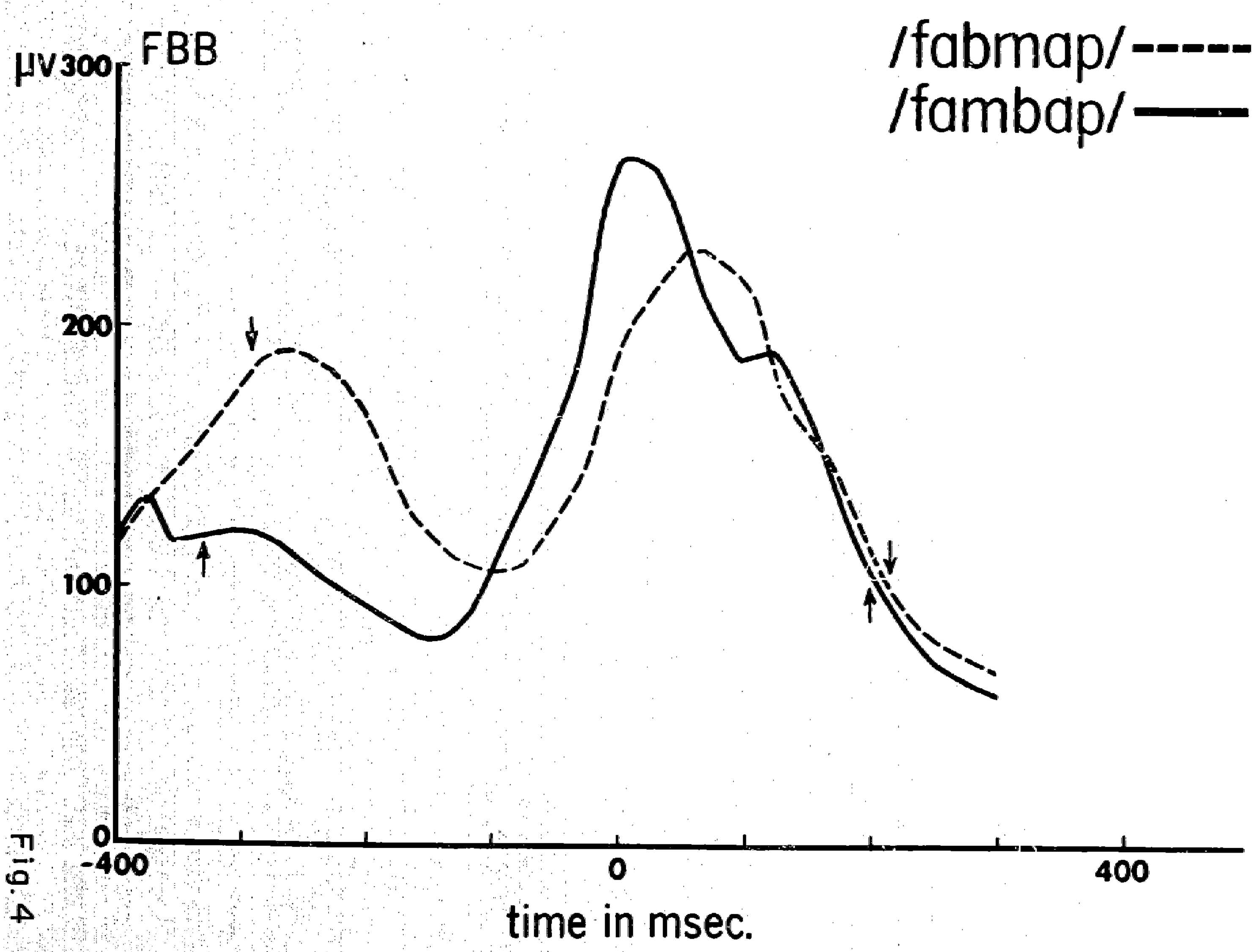


Fig. 4

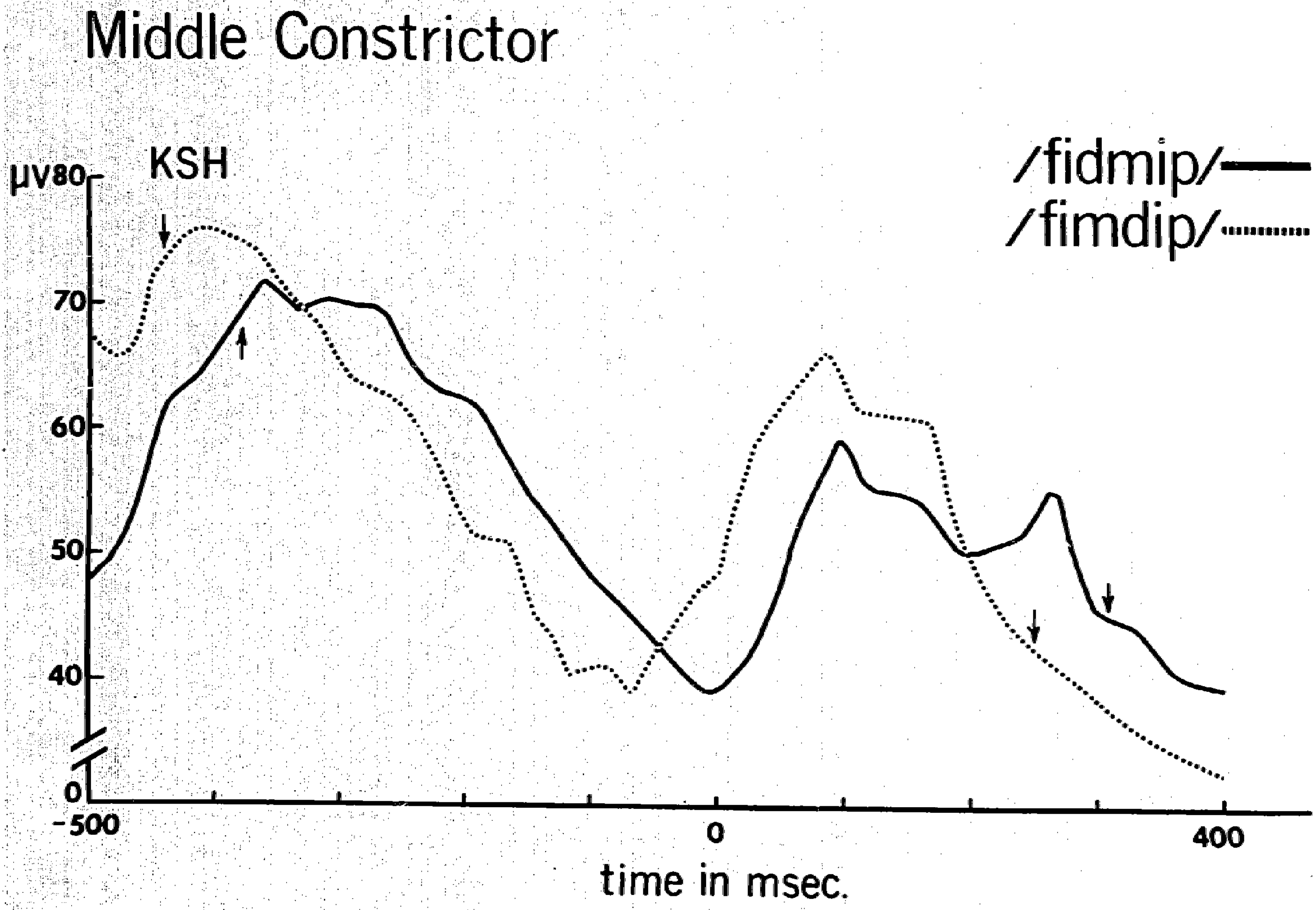

Middle Constrictor
KSH
μv80
70
60
50
40
0
-500
0
400
time in msec.
/fidmip/
/fimdip/

Fig. 5

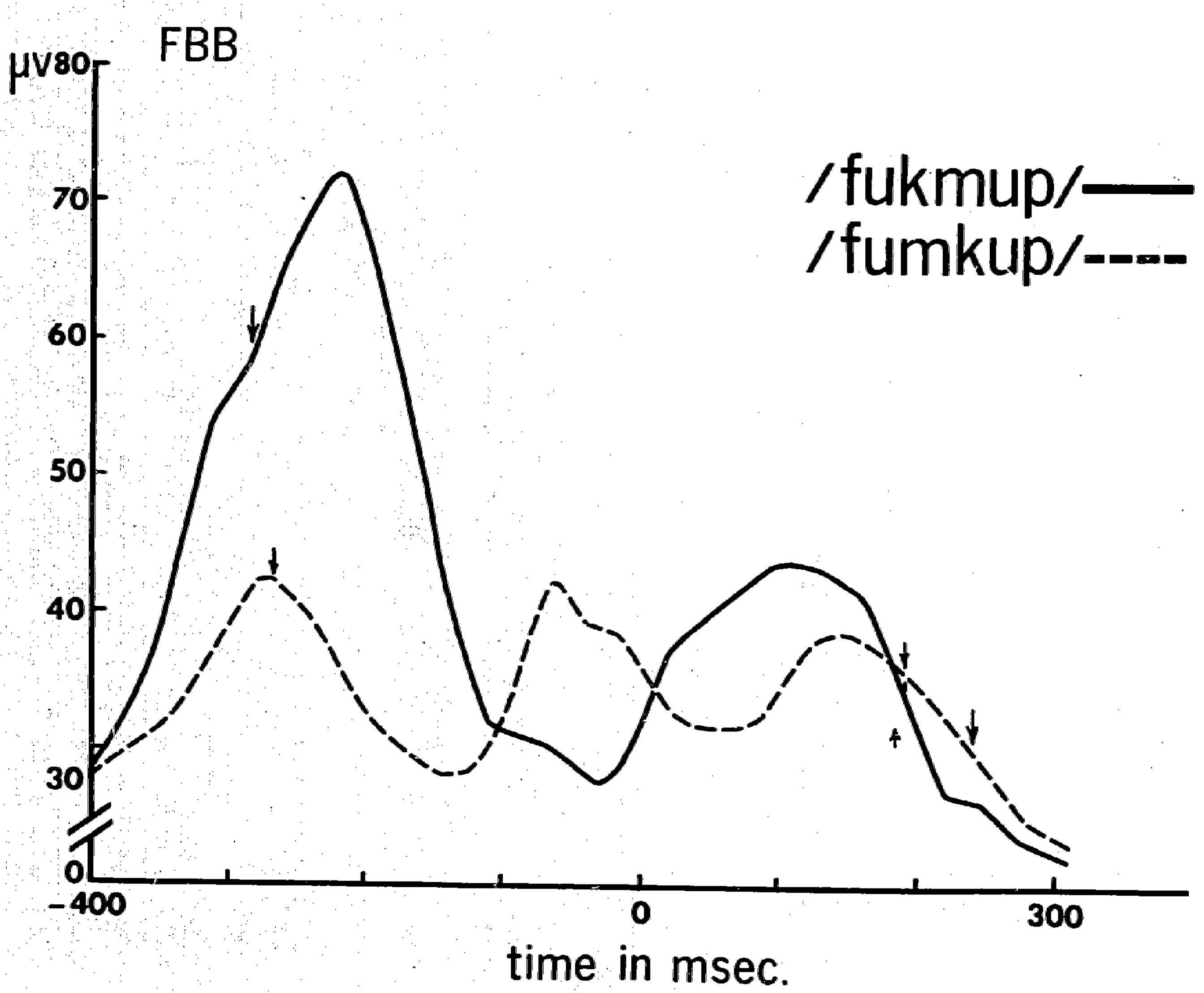

Palatoglossus

Fig. 6

to some degree, the muscles of the lateral and posterior pharyngeal walls (and also, presumably, the force of gravity acting upon the palate).

The Voicing Distinction

The stop consonant phones are distinguished on the basis of whether or not glottal pulsing occurs during the period of upper vocal tract occlusion. The continuation of glottal pulsing during occlusion of the upper vocal tract requires the maintenance of a transglottal pressure differential. One means of accomplishing this is increasing pharyngeal cavity volume, creating a pressure differential which is sufficient to allow continued glottal pulsing. All of the muscles associated with velopharyngeal closure effect pharyngeal cavity size.

There are two possible modes of pharyngeal enlargement. One mode is passive, that is, decreased activity of the pharyngeal wall muscles will result in pharyngeal cavity enlargement for voiced stops. The other mode is an active one, with greater muscle activity accompanying pharyngeal cavity enlargement for voiced stops.

Perkell (1969) and Chomsky and Halle (1968) postulate that pharyngeal wall tension is lower for voiced than for voiceless stop consonants, and therefore, that one essential quality of "voiced" stops is that they are "lax." Perkell's data on pharynx width are cited as supporting this distinction. Kent and Moll (1969) argue that a more plausible hypothesis is that pharyngeal cavity enlargement is the result of an active mechanism. Their data revealed a depression of the hyoid bone accompanied by a depression of the larynx for voiced stop consonants, causing "active" pharyngeal enlargement.

In addition to this active enlargement of the pharynx by hyoid bone and larynx depression, it is theoretically possible to increase pharyngeal volume by increasing velar height for voiced stops as compared with that achieved for voiceless stops.

Inspecting Perkell's cineradiographic data for the measurements related to these points (Figure 7), we observe that in each case the difference between voiced and voiceless stops supports the hypothesis of greater pharynx size for the voiced stop.

Both upper and lower pharynx width (Perkell's D and E) are greater for voiced than for voiceless stops. Velum height (Perkell's P) is only slightly affected, but the difference nevertheless supports the notion that pharyngeal cavity height may be increased cranially. Larynx height (Perkell's I) and hyoid height (Perkell's H) are both greater for /t/ than for /d/, that is, the larynx and the hyoid are depressed for the voiced stop consonant production. These arguments lead directly to our hypotheses.

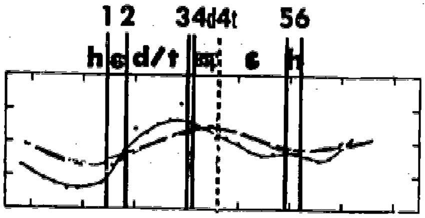
The first hypothesis proposes that there is active enlargement of the pharyngeal cavity for voiced stop consonant production. The muscles of this study which will have this effect are the levator palatini and the sternohyoid. The levator palatini is hypothesized to increase velum height for voiced stops, while the sternohyoid should act to depress the hyoid bone and the

(AFTER PERKELL)

If pharynx larger
Then for 2-3:

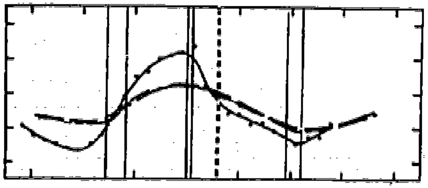
d ——
t - - -

d>t



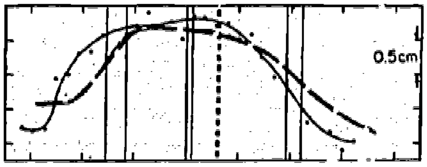
Upper pharynx width (D)

d>t



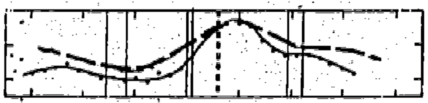
Lower pharynx width (E)

d>t



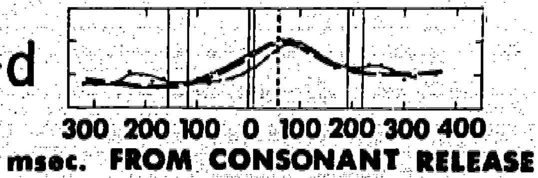
Velum height (P)

t>d



Larynx height (I)

t>d



Hyoid height (H)

Fig. 7

larynx. For these muscles, EMG levels for the voiced stop consonants should be greater than those for the voiceless stops (Figure 8).

The second hypothesis proposes that there is passive enlargement of the pharyngeal cavity for voiced stops. The muscles which might have this effect are the superior and middle constrictors, the palatopharyngeus, and the palatoglossus. Relaxation of these muscles should cause retraction of the lateral and posterior pharyngeal walls. If this hypothesis is pertinent, EMG levels for the voiced stop consonants should be lower than those for the voiceless stops (Figure 8).

Inspection of EMG levels for each of the muscles in this study, save the orbicularis oris, at the time of peak leavator palatini activity associated with stop production was performed for all minimal stimulus pairs for each subject (for instance, /fambap/ and /fampap/). There is a total of seventy-four possible minimal comparisons, across three subjects, for each muscle studied. When the difference in potential for a given contrast supported the hypothesis a value of "1" was assigned to that muscle for that utterance pair. When there was no difference, a value of "1/2" was assigned. When the difference failed to support the hypothesis, a value of "0" was assigned.

All of the supporting instances for the active hypothesis were pooled; the cases of levator palatini and sternohyoid activity which support the active hypothesis were added and then divided by total comparisons for the active hypothesis for that subject. The same analysis was performed for the muscles involved in the passive hypothesis.

The results of this analysis indicate three patterns of EMG activity accompanying production of voiced, as compared with voiceless, stop consonants (Figure 9). A speaker who uses relatively little active enlargement (Subject LJR: 52% $p > .05$) uses a considerable amoumt of passive enlargement (81% $p < .01$). A speaker who uses a great deal of active pharyngeal enlargement (Subject FBB: 90% $p < .01$) uses relatively little passive enlargement (50% $p > .10$). A speaker whose use of active enlargement falls midway between the more extreme cases (Subject KSH: 65% $p < .05$) also makes use of a "middling" amount of passive enlargement (73% $p < .01$). Overall hypothesis support was 72% ($p < .01$) for LJR and 67% ($p < .01$) each for KSH and FBB.

It appears from these data that an adequate description of pharyngeal cavity enlargement for voiced stop consonants is neither exclusively active nor exclusively passive. Each speaker uses both modes, though some prefer one to the other. It is also apparent that the description "tense-nontense" is inadequate for describing the activities of the pharyngeal cavity concomitant with voicing distinctions, as such a description at best explains the larger portion of some speakers' pharyngeal enlargements and never explains the full measure of enlargement.

It is not known at the present time whether these different modes are related to anatomical differences among subjects, dialectal differences, or learned patterns based on other articulatory balances.

In summary, the muscles of the velopharynx participate in oral and nasal articulation and in adjustments of pharyngeal cavity size. Oral gestures are

Hypotheses:	EMG
Active Enlargement	
levator palatini (LEV)	voiced > voiceless
sternohyoid (SH)	
Passive Enlargement	
superior pharyngeal constrictor + middle pharyngeal constrictor (CON)	voiced < voiceless
palatoglossus + palatopharyngeous (FP)	

Fig. 8

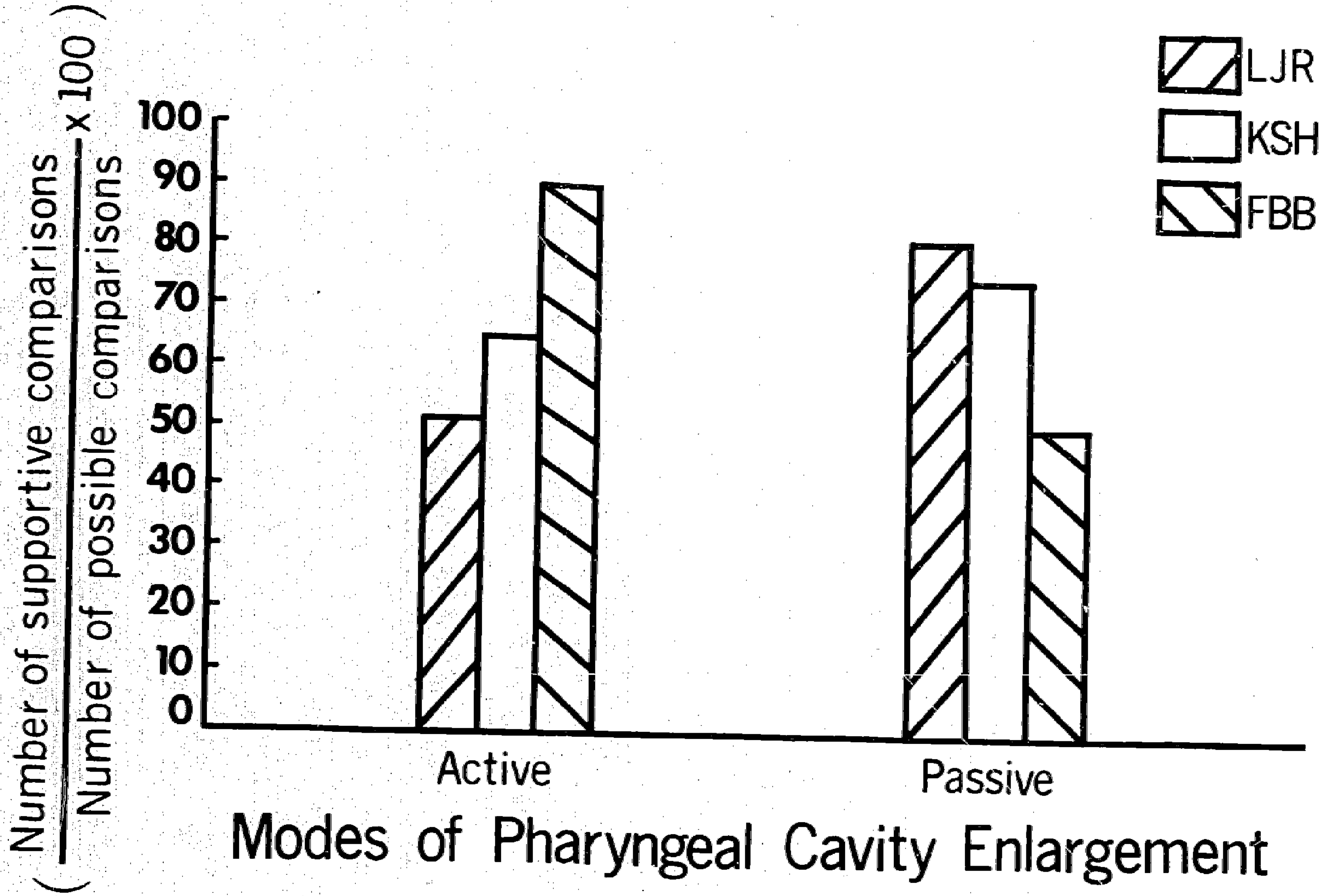

LJR
KSH
FBB
100
90
80
70
60
50
40
30
20
10
0
Active
Passive
Modes of Pharyngeal Cavity Enlargement
(Number of supportive comparisons / Number of possible comparisons x 100)

Fig. 9

accompanied by peaks of EMG activity in the levator palatini, superior and middle constrictors, and palatopharyngeus. Nasal gestures are accompanied by decreased activity in the aforementioned muscles, with no evidence of an active palate-lowering muscle. Palatoglossus activity peaks for tongue-backing and -raising gestures. Pharyngeal cavity enlargement may be effected by varying combinations of increased levator palatini and sternohyoid activity and decreased pharyngeal constrictor and faucal pillar activity.

REFERENCES

Chomsky, Noam and Morris Halle (1968) The Sound Pattern of English. New York: Harper and Row.

Fritzell, Bjorn (1969) The velopharyngeal muscles in speech: An electromyographic and cineradiographic study. Acta Otolaryngologica, Suppl. 250.

Hirose, Hajime (1971) Electromyography of the articulatory muscles: Current instrumentation and technique. Haskins Laboratories Status Report on Speech Research 25/26, 73-86.

Kent, Raymond D. and Kenneth L. Moll (1969) Vocal tract characteristics of the stop cognates. J. acoust. Soc. Amer. 46, 1549-1555.

Lubker, James F., Bjorn Fritzell, and Jan Lindqvist (1970) Velopharyngeal function: An electromyographic study. Royal Institute of Technology Speech Transmission Laboratory, Quarterly Progress and Status Report 4, 9-20.

Perkell, Joseph S. (1969) Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study. Cambridge, Mass.: M.I.T. Press.

Port, Diane K. (1971) The EMG data system. Haskins Laboratories Status Report on Speech Research 25/26, 67-72.

Laryngeal Adjustments for Vowel Devoicing in Japanese: An Electromyographic Study

Hajime Hirose*
Haskins Laboratories, New Haven

It is well known that high vowels between voiceless consonants are often devoiced in many dialects of Japanese including Tokyo dialects (Bloch, 1950; Han, 1962). Previous studies with a fiberscope revealed that the glottis remained open for the devoiced vowel segments (Hirose, 1971a; Sawashima, 1969, 1971). Based on an electromyographic study of the activity of the vocalis muscle in articulation, the present author reported that devoicing of Japanese vowels appears to be a matter concerning the neural process that determines the motor commands to the larynx (Hirose, 1971a). In the present study, electromyographic activities of selected intrinsic laryngeal muscles were examined with special reference to vowel devoicing in Japanese in comparison with the production of voiceless consonants.

METHOD

A speaker of the Tokyo dialect served as the subject in the present study and read randomized lists of test sentences sixteen times each. Each sentence embedded a test word in a frame "soreo -- to ju:" (That we call --). Table I lists the types of test words used in the experiment. They are all meaningful Japanese words. No accent kernel is attached to those words except for the last four pairs in the table, in which the position of the accent kernel is indicated by the mark "⌝." Devoicing typically occurs for all [i]'s between voiceless consonants as indicated in the table.

Electromyographic recordings were made using hooked-wire electrodes. The wires used were insulated platinum-iridium alloy, the outer diameters of which were approximately 50 microns. The electrodes were inserted perorally using a curved probe into the posterior cricoarytenoid (PCA) and the interarytenoid (INT) by indirect laryngoscopy, while percutaneous approach was employed for insertion into the vocalis (VOC) and the cricothyroid (CT). Further description of the insertion techniques may be found in previous reports (Hirose, 1971a, b).

The electromyographic signals were recorded on a multichannel data recorder together with acoustic signals and automatic timing markers. The signals were reproduced, high-pass filtered, and fed into a computer after appropriate rectification and integration. The electromyographic signals were averaged for more than fourteen selected utterances of each test sentence with reference to a line-up point on the time axis representing a predetermined

---

*Also Faculty of Medicine, University of Tokyo

(1) Words with no accent kernel

[sesse:]	[seki̥se:]	
[sekke:]	[seki̥ke:]	[seki'e:]
[sette:]	[seki̥te:]	
[s$_j$i̥se:]	[z$_j$ise:]	[s$_j$ize:]
[s$_j$i̥ke:]	[z$_j$ike:]	[s$_j$ige:]
[s$_j$i̥te:]	[z$_j$ite:]	[s$_j$ide:]
[s$_j$i̥he:]	[z$_j$ihe:]	
[kiri]	[giri]	
[tenko:]	[denko:]	

initial [s] 7

initial [s$_j$] 3

initial [z$_j$] 4

medial [k] 4

One for each, otherwise

initial .... k, g, t, d,
si̥s, si̥k, si̥t, si̥h

medial .... s, z, g, t, d, h
ki̥s, ki̥t, ki̥k,
ss, kk, tt,

(2) Words with accent kernel

[se⌉:ri]	[ze⌉:ri]	[ke⌉:ri]	[ge⌉:ri]
[te⌉:t i]	[de⌉:t i]	[pa⌉su]	[ba⌉su]

Table I: List of test words used in the present study.

speech event. In the present experiment, voice onset following [t] in the frame " --- to yuu" in each sentence was taken for a line-up point. The data recording and the computer-processing system employed in the present experiment as described in more detail by Port (1971).

RESULTS

The laryngeal adjustments in terms of the opening and closing gestures of the glottis for the voiced/voiceless distinction appeared to be executed by reciprocal activities of the abductor and adductor muscle groups of the larynx. In particular, the PCA consistently showed increasing activity for the voiceless portions of test utterances, while its activity was suppressed for the production of voiced segments. Conversely, INT activity appeared to be suppressed for voiceless portions and increased for voiced ones, thus presenting a sort of inversion of the pattern of PCA activity throughout the utterance.

Figure 1 illustrates an example of the averaged EMG curves of, from bottom to top, the PCA, the INT, and the VOC, for the utterance of [soreo $z_J$ike: to ju:] and of [soreo $s_J$ige: to ju:], thus comparing the patterns of the muscle activity in respect to the [z_J] vs. [s_J] and [k] vs. [g] contrasts. It is clearly demonstrated in both cases that there is a reciprocal pattern of activity between the PCA and the INT.

In the case of [soreo $s_J$ige: to ju:], for example, the PCA shows increasing activity for the production of voiceless [s_J] and [t] and remains suppressed for the rest of the test utterance. On the contrary, the INT shows a rapid decrease in activity for [s_J] and [t], while it stays at high level for the rest. The timing of the peak PCA activity approximately coincides with that of the maximum suppression of INT activity. There is a shallow dip in the INT curve, apparently corresponding to [g] production.

For the utterance of [soreo $z_J$ike: to ju:], the PCA shows increasing activity for [k] and [t] and suppressed activity for the rest. The INT shows a gradual decrease in activity for the sequence [z_Ji], followed by further suppression corresponding to increasing PCA activity.

The activity of the VOC generally stays at a high level for the vowel portion of the utterance, while it becomes low for consonant segments regardless of the voiced/voiceless distinction, although the activity is usually, but with some exceptions, somewhat higher for a voiced consonant than for a voiceless consonant if we compare the averaged EMG values for a given set of voiced/voiceless consonant pairs.[1]

Figure 2 compares the averaged EMG curves for the sentences embedding [sette:] vs. [seki̥te:], where the interconsonantal [i] is devoiced in the latter.

It is shown that PCA activity increases for the sequence [ki̥t] as well as for the geminate [tt] and initial [s], while the INT is markedly suppressed for these sequences.

---

[1] In the examples in Figure 1, VOC activity is higher for [g] than for [k] but lower for [z_J] than for [s_J].

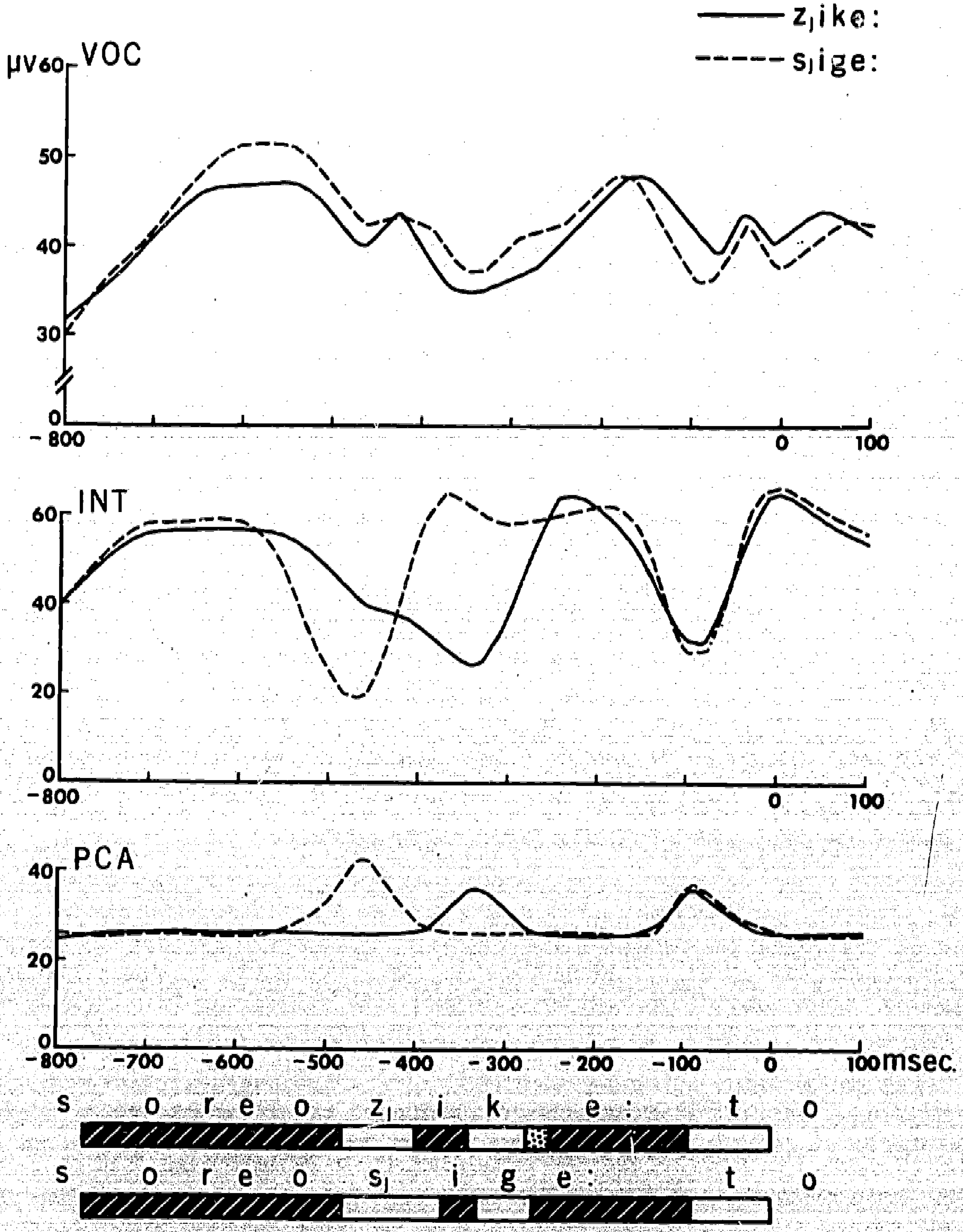


Figure 1: Superimposed averaged EMG signals of, from top to bottom, the vocalis, the interarytenoid, and the posterior cricoarytenoid for the utterances [soreo $z_J$ike: to ju:] and [soreo $s_J$ige: to ju:].

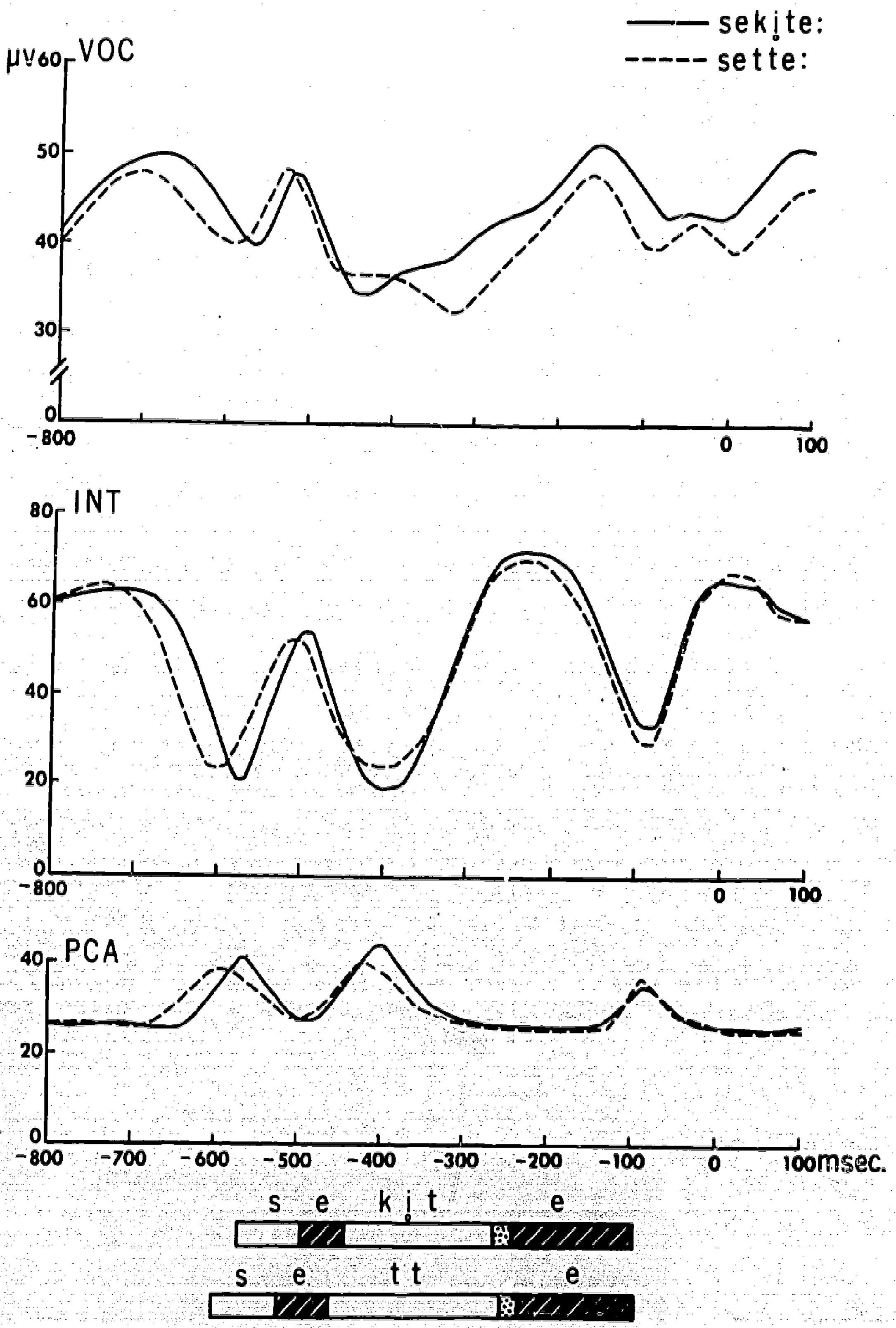


Figure 2: Superimposed averaged EMG signals of the same set of the muscles as given in Figure 1, for the utterances embedding [sekite:] and [sette:].

The VOC appears to be suppressed for consonant segments in these examples, too, as well as for [ki̥t] and for the geminate.

The CT, for which data is not given here, did not give a consistent difference in the pattern of activity with respect to the voiced/voiceless distinction. It was revealed, however, that the CT showed increasing activity for the production of a syllable with an accent kernel.

The present data thus indicate that the PCA and the INT are most likely to play principal roles in the voiced/voiceless distinction as possible physiological correlates for opening vs. closing gestures of the glottis.

An attempt was made to estimate the degree of PCA activation and INT suppression for different phonetic representations by measuring maximum EMG values for the PCA and minimum values for the INT for a given voiceless sequence of the test words. For estimation of PCA activation, maximum EMG values were simply taken. For the INT, the minimum EMG value for a given voiceless portion was subtracted from a predetermined value of 50,[2] and the remainder was taken to indicate the degree of INT suppression.

It was revealed that PCA activation thus specified is highest for initial [si̥t] and that INT suppression is most marked for initial [si̥s].

Figure 3 presents the degree of PCA activation and INT suppression for the voiceless portion of each test word, being normalized and illustrated on an arbitrary scale, where the value for initial [si̥t] for the PCA and that for initial [si̥s] for the INT are taken as standard values of 100. In most cases, the timing for maximum PCA activation and INT suppression were found to coincide. Therefore, normalized values for a given voiceless portion are superimposed in Figure 3. Although there is certain discrepancy between apparent PCA activation and INT suppression thus specified for a given voiceless sequence, we can approximately compare overall muscle activities which are most likely to be responsible for an opening gesture of the glottis for each voiceless portion of the test words. In this figure, the value for initial [s] represents the mean for seven different kinds of test words having initial [s], while that for [s⌉] for three, and that for medial [k] for four. Those consonants in the syllable with an accent kernel are eliminated from the data.

It is found that the initial /si̥C/ sequences generally give the greatest values as noted in Figure 3, while the medial stops show the smallest. The values for geminate stops are higher than those for the medial stops but lower than for medial /kiC/ sequences.

COMMENT

It has been suggested that the intrinsic laryngeal muscles play an essential role in laryngeal articulatory adjustments. It was reported that the PCA

---

[2] INT activity for voiced segments almost always exceeds the level of 50 μv. It was assumed, therefore, that the EMG values lower than that particular level can be regarded as an indication of INT suppression.

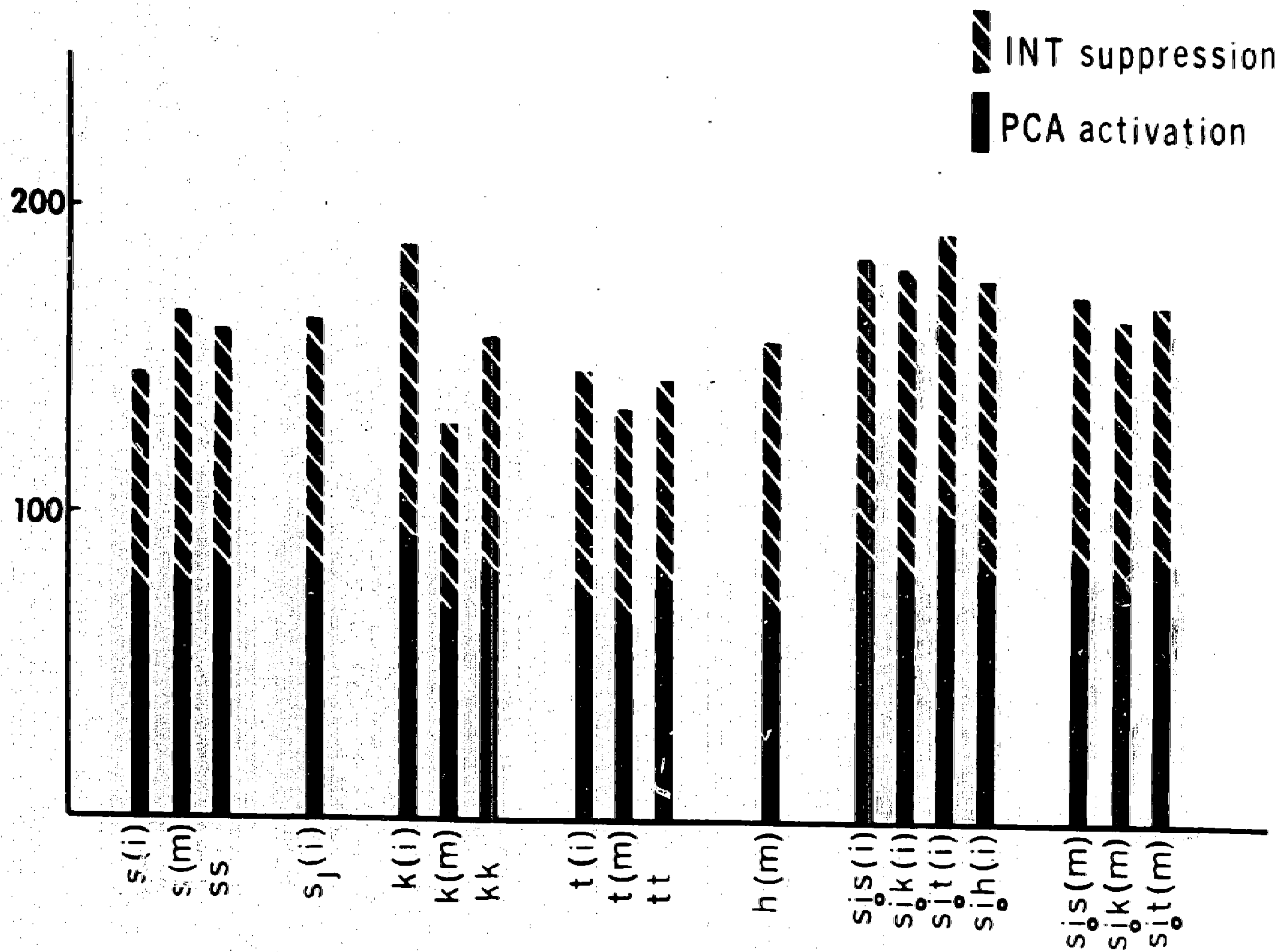


Figure 3: Relative values representing PCA activation plus INT suppression for each voiceless portion of the test word, where the value for the initial [si̥k] for the PCA and that for the initial [si̥s] for the INT are each taken as 100. (i) indicates the word initial position, and (m), the word medial position.

and the INT showed consistent reciprocal activities for articulatory adjustments in terms of the opening and closing gestures of the glottis, while the VOC appeared to participate particularly in vowel production (Hirose, 1971c). It has also been assumed that there is some relationship between VOC activity and suprasegmental features such as pitch.

The present study generally supports the above concepts. It was also revealed that the abductor of the glottis, the PCA, showed marked activity for the sequence containing a devoiced vowel, while the adductors were reciprocally suppressed. These findings are in agreement with our previous studies in which we claimed that devoicing of Japanese vowels is a matter under active motor control of the larynx. The data further suggest that PCA activation associated with INT and VOC suppression is essential for the control of the opening gesture of the glottis in devoicing.

It has been claimed that the integrated electromyogram parallels tension in human muscles contracting isometrically (Inman et al., 1952). Since the laryngeal muscles execute neither purely isometric nor purely isotonic contraction, it is not feasible simply to correlate the averaged EMG values of a given laryngeal muscle to the tension of the muscle or displacement of the effector, such as the vocal cord. However, it would be reasonable to assume that a given value of PCA activation and INT suppression as presented in Figure 3 may, to some extent, represent the degree of glottal opening. If this assumption is correct, it should be of interest to compare the EMG results to the glottal gestures directly observed by means of a fiberscope.

Sawashima (1971) measured glottal width during the production of voiceless consonants, geminates, and voiceless sequences containing a devoiced vowel in Japanese by means of a fiberscope and reported that there were certain differences in glottal width depending on different phonetic representations. In his study of one subject, he observed that the maximum glottal width for a given voiceless portion was largest for the initial voiceless sequence containing a devoiced vowel and smallest for medial stops. These results are apparently coherent with the present data, which indicate that the highest PCA activity is associated with the lowest INT activity for initial voiceless sequences containing a devoiced vowel, and the lowest PCA activity is associated with highest INT activity for medial stops. Sawashima also found that glottal width for geminate stops was significantly smaller than that for medial voiceless sequences containing a devoiced vowel. In the present data, however, the difference in degree of PCA activation and INT suppression between these two conditions does not appear to be very marked. It seems that we need more data in order to specify the physiological basis of these two different phonetic conditions in more detail, although possible individual variation has to be taken into consideration. A combined study of simultaneous fiberscopic observation of the glottis with laryngeal EMG data acquisition is expected to give further information on laryngeal adjustments in speech.

REFERENCES

Bloch, B. (1950) Studies in colloquial Japanese IV: Phonemics. Language 26, 86-125 (Reprinted in Readings in Linguistics, ed. by M. Joos (Chicago: University of Chicago Press) p. 337.)

Han, M.S. (1962) Unvoicing of vowels in Japanese. The Study of Sounds *10*, 81-100.

Hirose, H. (1971a) The activity of the adductor laryngeal muscles in respect to vowel devoicing in Japanese. Phonetica *23*, 156-170.

Hirose, H. (1971b) Electromyography of the articulatory muscles: Current instrumentation and technique. Haskins Laboratories Status Report on Speech Research *SR 25/26*, 73-86.

Hirose, H. (1971c) An electromyographic study of laryngeal adjustments during speech articulation: A preliminary report. Haskins Laboratories Status Report on Speech Research *SR 25/26*, 107-116.

Inman, V.T, H.J. Ralston, J.B. Saunders, and B. Feinstein (1952) Relation of human electromyogram to muscular tension. EEG and Clinical Neurophysiol. *4*, 187-194.

Port, D. (1971) The EMG data system. Haskins Laboratories Status Report on Speech Research *SR 25/26*, 67-72.

Sawashima, M. (1969) Devoiced syllables in Japanese--A preliminary study of photoelectric glottography. Annual Bulletin (Research Institute of Logopedics and Phoniatrics, University of Tokyo) *4*, 9-25.

Sawashima, M. (1971) Devoicing of vowels. Annual Bulletin (Research Institute of Logopedics and Phoniatrics, University of Tokyo) *6*, 7-13.

Vowel Stress and Articulatory Reorganization*

Katherine S. Harris+
Haskins Laboratories, New Haven

If a speaker is asked to produce a word which contains a particular speech sound, it can be shown that there will be a great deal of variability in what is produced. Some of this variability depends on the immediately neighboring speech sounds; some depends on the stress and intonation pattern in which the word is imbedded. A principal thrust of recent physiological investigation has been towards showing that at least part of this variability can be accounted for by relatively low-level rules. One formulation of this sort is the suggestion that a shape template, or target, for a speech sound is stored in the nervous system, and that the effects of coarticulation can be described as due to the overlapping effects of several targets at any moment in time (MacNeilage, 1970; Ohman, 1967). The most careful working out of this sort of formulation is probably Lindblom's (1963) ingenious theory of vowel neutralization.

This theory was developed to account for the changes in vowel color which accompany changes in stress. If a vowel is destressed, it will tend to be of shorter duration and to move in vowel color towards the neutral schwa; the latter phenomenon is called vowel neutralization. Lindblom's proposal is that the neutralization is a consequence of the accompanying shortening. Briefly, in a CVC sequence, although the signals sent to the articulators are constant, the response of the articulators is sluggish. If signals arrive at the muscles too fast, the articulators will start towards the vowel target but will be deflected towards the subsequent consonant target--that is, there will be undershoot. Lindblom tested his theory by having subjects produce sentences containing CVC monosyllables. The effect of rearranging the sentences was to change the stress on one "word" and consequently to change the vowel duration. He made careful measurements of the most extreme positions of the first and second formants, as a function of the vowel length. He found that as vowels lengthened, the formants tended towards a target frequency which could be described as a target articulation.

Lindblom's theory seemed to us to be elegant and testable, if one substitutes for "signals" the more specific "muscle contractions." A reformulation in electromyographic terms would then perhaps be: "Under conditions of changing stress (or rate of articulation) the electromyographic signals associated with any vowel will remain constant. Only the spacing between them will change."

---

[*] This paper is a somewhat rewritten version of a paper, "The Organization of Articulation Schema," presented at the 1971 Convention of the American Speech and Hearing Association, Chicago, Illinois, November 1971.

[+] Also The Graduate Center of the City University of New York.

Some time ago, we performed an experiment on consonants which is relevant here (Harris, Gay, Sholes, and Lieberman, 1968). Subjects produced sentences with one word containing /p/; this word was either heavily contrastively stressed or not. Thus, we would compare "It's the keeper" with "It's the <u>keeper</u>." An electrode in the orbicularis oris muscle measured the strength of the closure contraction, or "signal" in Lindblom's terms.

The results showed a contrast in the amplitude of the EMG signal for the two conditions. However, the effect was quite small--about 20 percent difference between conditions. Furthermore, even this difference was obtained only under conditions of very strong contrastive stress, perhaps stronger than we would observe in ordinary running speech.

We wanted to repeat the experiment with vowels, using stress contrasts more like those in ordinary running speech. The genioglossus muscle, which is active for high vowels, seemed suitable for examination (Harris, 1971; Raphael, 1971; Smith, 1971).

Figure 1 shows the genioglossus muscle. It is a large, fanshaped muscle, which is generally described as bunching and fronting the tongue. The arrow shows the general direction of electrode insertion into the muscle body. Electrode preparation and insertion procedures are described in detail elsewhere (Hirose, 1971).

We constructed a set of nonsense trisyllables, with stress on either the first or the second syllable. The vowel in one syllable was always /i/, while the vowel in the other syllable was /ɔ/ or /u/; /i/ appeared equally often in the first or second syllable and was equally often stressed and unstressed. All conditions were repeated with /p/ as an intervocalic consonant and with /k/. Typical trisyllables, then, would be /ṕikupə/ and /puṕipə/. The subject read sixteen lists in which these nonsense words appeared in random order. The resulting electromyographic signals were recorded and averaged by the usual techniques (Port, 1971).

To return to Lindblom's model, it would lead us to expect a constant muscle signal for the vowel, /i/, with changes in timing of adjacent signals, depending on stress context.

Figure 2 shows the utterances /ṕikupə/ and /piḱupə/. As usual, time runs along the abscissa and the ordinate indicates amplitude of muscle signal. Zero is the point corresponding to the end of voicing in the first syllable. The pair of utterances contrast in whether the first or the second syllable is stressed. If /i/ in the first syllable is stressed, the amplitude for /i/ increases. If /u/ in the second syllable is stressed, /i/ amplitude will decrease. (The vowel /u/ also shows some genioglossus activity, since it is a high vowel.) The amplitude of the stressed syllable is greater than the amplitude of a corresponding unstressed syllable. Of course, we see changes in timing, as well.

Peak heights of the genioglossus activity, averaged over various conditions, are shown in tabular form in Figure 3. This slide shows mean peak height values for four conditions--when /i/ is stressed and unstressed, in the first syllable and in the second. Overall, stress produces greater a activity.

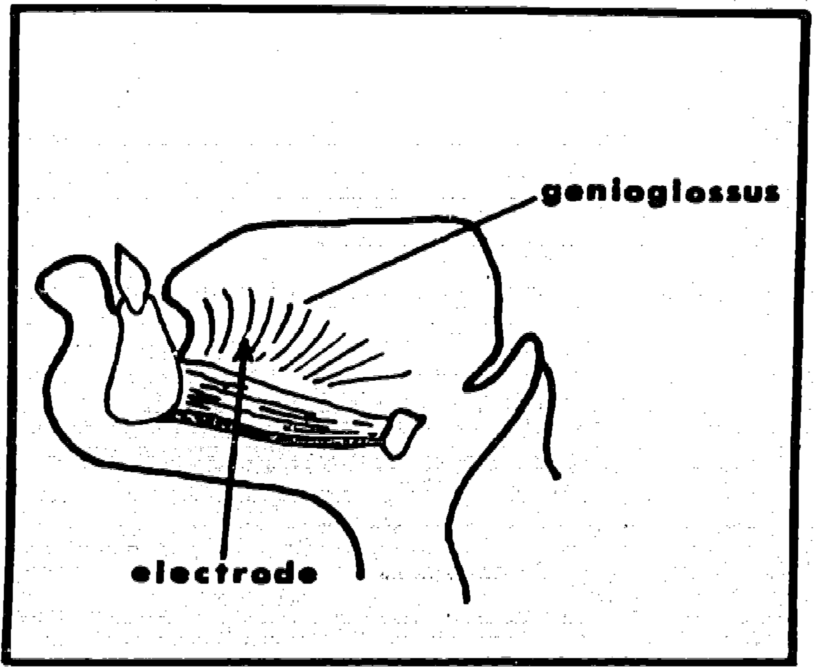


Figure 1: Electrode insertion into the genioglossus muscle.

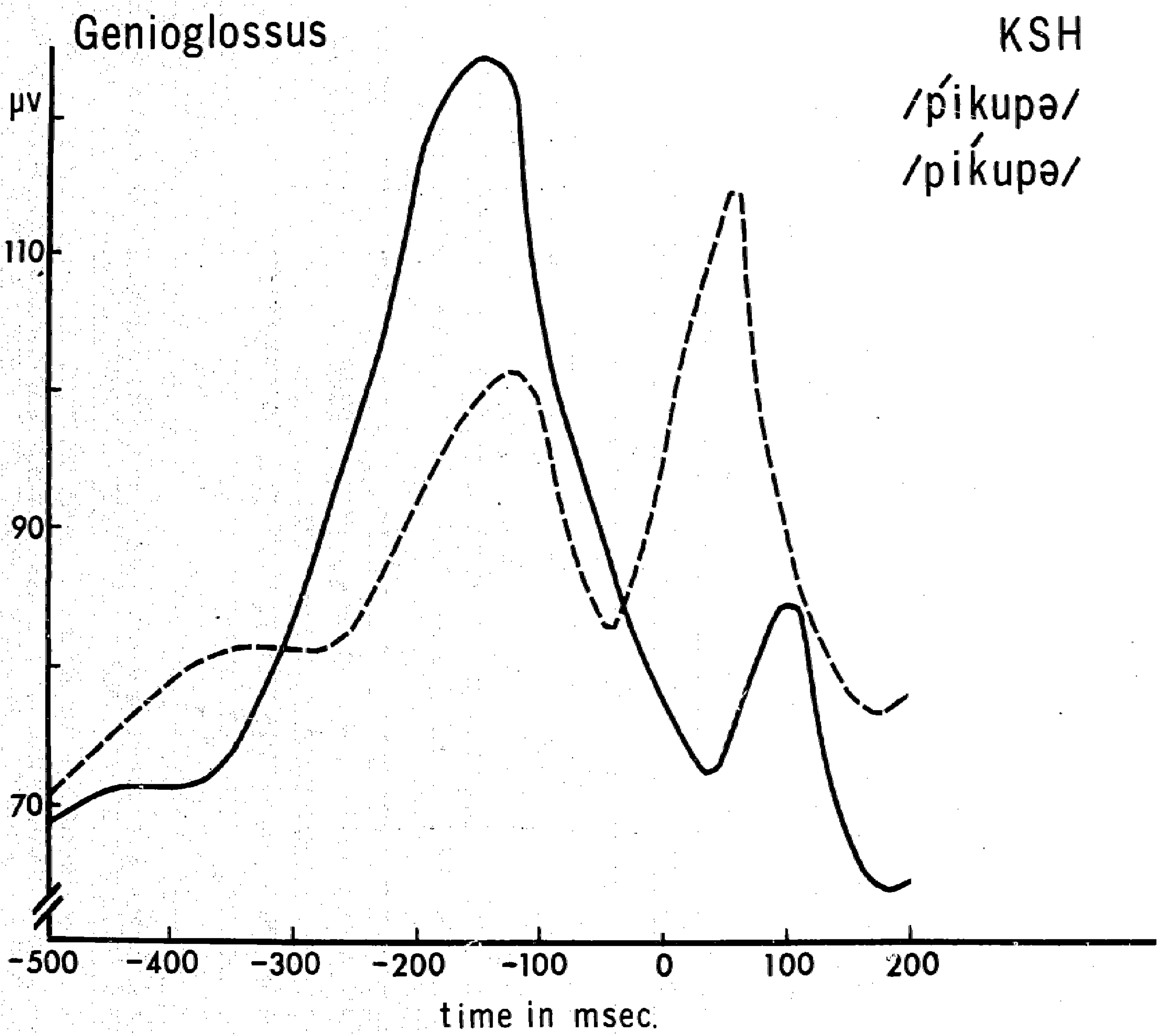


Figure 2: Averaged EMG signals from the genioglossus muscle for tri-syllables which are stressed on the first and second syllable.

KSH Genioglossus

		Stress Condition	
		Stressed	Unstressed
Syllable Position	S_1	137	119
	S_2	169	110

Figure 3: Peak heights of genioglossus activity under various stress conditions.

One further question may be asked. If differences in signal size contribute to stressing differences, is there any evidence that the duration mechanism works as well? Lindblom's model says, in essence, that the longer the vowel, the less neutral. To consider this question, we must extend the model to yet another situation, the vowel duration differences which accompany the shift from voiceless to voiced terminal consonants.

This phenomenon is extremely well known. Briefly, the vowel before a voiceless stop or fricative is shorter than before the corresponding voiced consonant. Now, let us assume that Lindblom's mechanism is at work in running speech. If the time distance between the vowel signal and the consonant signal is shorter for voiceless than for voiced consonants, then one of two things must happen: either the vowel must be more neutralized before voiceless stops, or alternatively, there should be an adjustment of peak activity to compensate for the duration difference. There is no evidence, either in our own work or, so far as I know, in the extensive literature on the voicing effect, that the vowel before a voiceless consonant is more neutralized than before a voiced stop, although we should, of course, check spectrograms, which has not yet been done. Some data collected by Raphael (1971) allow us to examine the second possibility.

Figure 4 shows genioglossus activity for four high front vowels in the frame /pVp/. There is substantial genioglossus activity for /i/ and /e/ before /p/ for this subject but relatively little activity for their so-called "lax" counterparts. Since the genioglossus is apparently a chief determiner of vowel color for /i/ and /e/, we would expect an adjustment in peak height to compensate for the difference in vowel length before voiced and voiceless consonants. On the other hand, we have no such anticipation with respect to /ɛ/ and /ɪ/, since they show very little activity.

Figure 5 shows peak heights for the four vowels before a series of voiced and voiceless consonant pairs. Overall, peak activity is lower for the voiced member of the pair, although there is one case of approximate equality. The situation is reversed before the lax vowels--I have no idea why. For long vowels this result can be interpreted as a tendency to compensate for duration differences, with peak size changes, for "essential" muscles. This compensation anticipates the duration difference, that is, the speaker seems to make some sort of anticipatory calculation.

Figure 6 shows peak values for a second subject, who used relatively high values of genioglossus activity for all four vowels (though notice that /e/ is strongly diphthongized for this speaker, so that only the second peak corresponding to /i/ or /ɪ/, is high). We would, therefore, expect compensation for voicing distinctions in all four vowels.

Figure 7 shows peak heights for the four vowels. We looked only at two sets of voiced/voiceless pairs for this subject. There are two entries for /e/, the diphthongized vowel, one for each peak. We would expect greater activity for the voiceless member of the pair for all four vowels, and indeed, this is about what we get, though there is one case of approximate equality. For /e/, only the second peak shows voicing compensation.

Let me summarize at this point. We have produced some rather preliminary evidence that stressing may affect the size of the contraction signals to

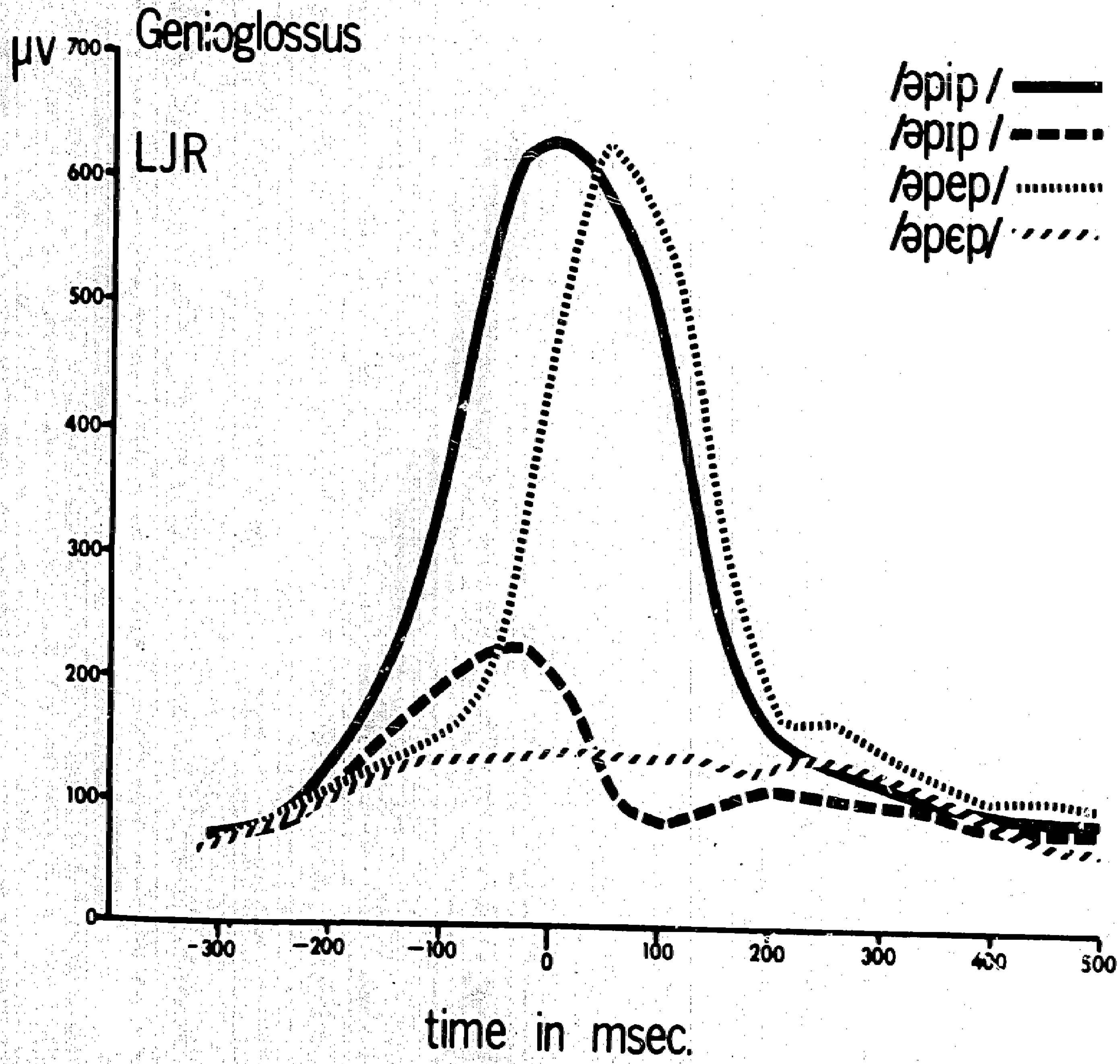


Figure 4: Averaged EMG signals from the genioglossus muscle for four high front vowels; speaker LJR.

LJR

Genioglossus

Terminal Consonant

Vowel	p	b	t	d	k	g	s	z
i	653	606	702	644	694	625	532	537
ɪ	236	319	316	360	268	494	223	294
e	643	533	710	538	643	571	530	500
ɛ	140	154	179	331	206	481	172	195

Figure 5: Peak heights of genioglossus activity for four high front vowels before voiced/voiceless pairs; speaker LJR.

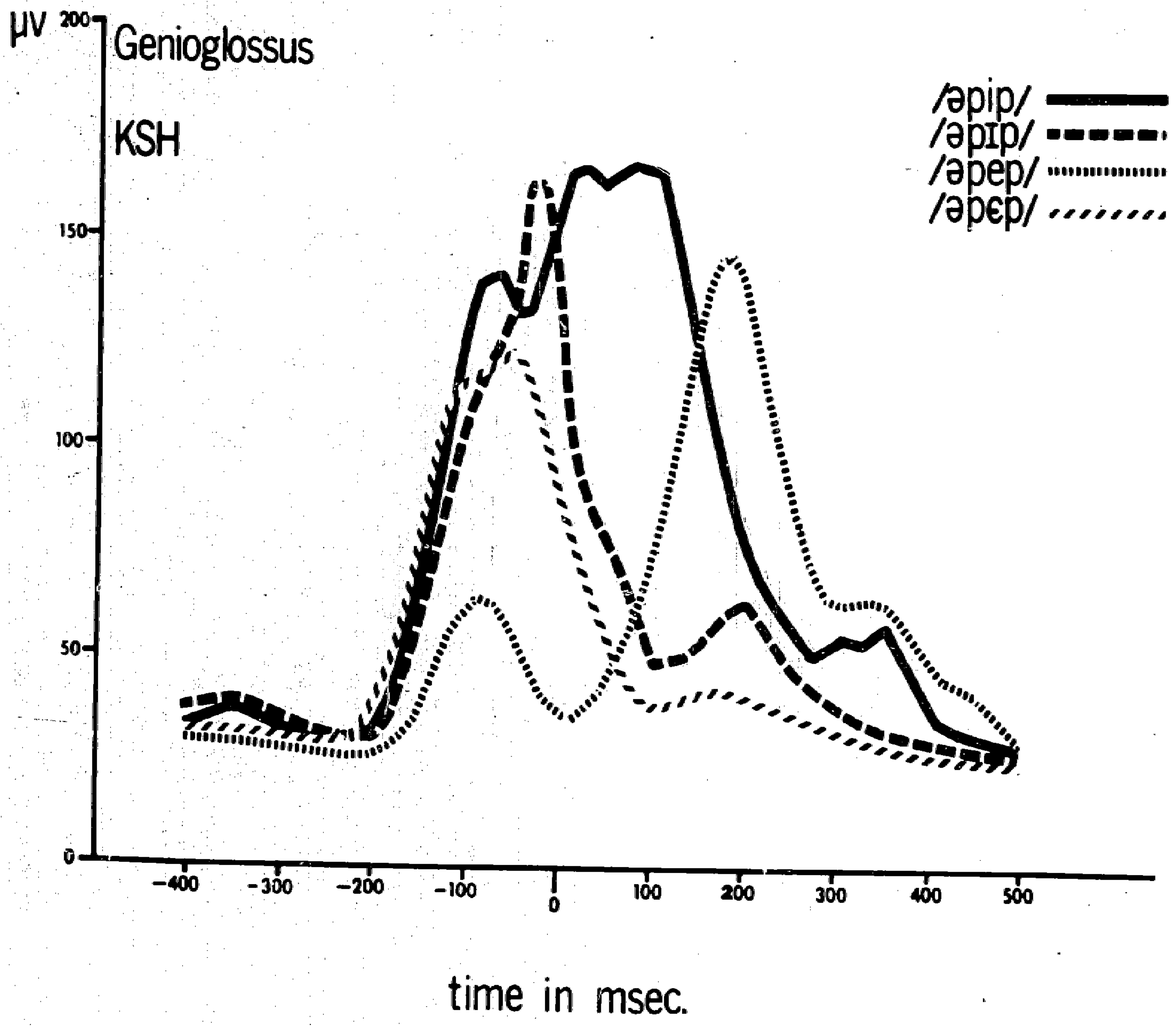


Figure 6: Averaged EMG signals from the genioglossus muscle for four high front vowels; speaker KSH.

KSH Genioglossus

	Terminal Consonant			
Vowel	p	b	k	g
i	177	158	199	111
ɪ	190	148	110	111
e	95/179	135/136	103/116	95/93
ɛ	163	127	102	95

Figure 7: Peak heights of genioglossus activity for four high front vowels before voiced/voiceless pairs; speaker KSH.

muscles, as well as their timing, although, by generous overinterpretation of the data, we can find some evidence for the effectiveness of a timing change mechanism, as well. However, if we presume that the "extra energy" mechanism works at all, it really originates more problems than it solves, since it leaves the question of what is invariant about a vowel under two stress conditions. Presumably, each vowel would be characterized by a pattern of contractions; however, if the size of one member of the pattern changes, what happens to the others?

Vowel height can be shown to be a joint product of tongue height and jaw opening. If genioglossus activity changes under stress, does the activity of the anterior belly, and the other muscles which open the jaw, increase proportionately? It seems far more likely that, for any vowel, only a selected group of muscles increase activity under stress. If this is indeed so, then the pattern of activity for any vowel becomes different, not only in "size" but in configuration, for changes in stress.

How does all this affect our views of speech mechanisms? The most common model for afferent feedback is that there is, for any phone, a "target" articulation, which is represented either as a position in the mouth or, more specifically, as a set of muscle lengths of each phone. These two hypotheses differ in their specificity. In the second case, not only is a target required, but the target must be reached by the same set of muscle adjustments each time. A recent observation (Lindblom and Sundberg, 1971) shows that if a speaker must attain a given tongue height with a jaw opening that is constrained by a block holding the jaws open at a fixed distance, he will use a compensating adjustment of the muscles to raise the tongue. Some data of Borden's (1972) can be interpreted to mean that if one of a set of muscles is partially paralyzed, other muscles will attempt to compensate by more than normal activity. These observations seem to me to indicate that a target representation in muscle length terms is probably not a sensible one. The simple continuous gamma loop correction models, depending on attainment of a set of lengths, would seem to fall with this evidence. "Targets" must somehow be specified in position coordinates which allow for configuration flexibility. The study we reported here seems to indicate that a given vowel must be represented as a series of targets which differ from some neutral point, and which are arrived at by different muscle action patterns. Single loop correction does not seem capable of operating successfully on targets which change in this fashion.

REFERENCES

Borden, Gloria J. (1972) Some effects of oral anesthesia on speech: A perceptual and electromyographic analysis. Ph.D. dissertation, City University of New York.

Harris, Katherine S. (1971) Action of the extrinsic musculature in the control of tongue position. Haskins Laboratories Status Report on Speech Research SR 25/26, 87-96.

Harris, Katherine S., Thomas Gay, George N. Sholes, and Philip Lieberman. (1968) Some stress effects on electromyographic measures of consonant articulations. Haskins Laboratories Status Report on Speech Research SR 13/14, 137-151.

Hirose, Hajime. (1971) Electromyography of the articulatory muscles: Current instrumentation and technique. Haskins Laboratories Status Report on Speech Research SR 25/26, 73-86.

Lindblom, Bjorn E.F. (1963) Spectrographic study of vowel reduction. J. acoust. Soc. Amer. 35, 1773-1781.

Lindblom, Bjorn E.F. and Johan E.F. Sundberg. (1971) Acoustical consequences of lip, tongue, jaw and larynx movement. J. acoust. Soc. Amer. 50, 1166-1179.

MacNeilage, Peter F. (1970) Motor control of serial ordering of speech. Psychological Review 77, 182-195.

Öhman, Sven E.G. (1967) Numerical model of coarticulation. J. acoust. Soc. Amer. 41, 310-320.

Port, Diane K. (1971) The EMG data system. Haskins Laboratories Status Report on Speech Research SR 25/26, 67-72.

Raphael, Lawrence J. (1971) An electromyographic investigation of the feature of tension in some American English vowels. Haskins Laboratories Status Report on Speech Research SR 28, (see this volume).

Smith, Timothy. (1971) A phonetic study of the function of the extrinsic tongue muscles. University of California, Los Angeles, Working Papers in Phonetics 18.

An Electromyographic Investigation of the Feature of Tension in Some American English Vowels*

Lawrence J. Raphael+
Haskins Laboratories, New Haven

Much of the traditional phonetic literature presents us with a picture of vowel articulation usually referred to as the vowel triangle or quadrilateral. In part, this vowel triangle appears as in Figure 1. Among the notable features of the arrangement of the six vowels shown on this part of the triangle are

(1) The pairing of the vowels [i] and [ɪ] (both high front), [e] and [ɛ] (both mid front), and [u] and [ʊ] (both high back).

(2) The relatively higher position for [i,e,u] within each pair.

(3) The more central position of [ɪ,ɛ,ʊ] in each pair.

A variety of features has been put forth to explain the difference in production between the members of these pairs of vowels in English speech. Among them are the following:

(1) Tongue tension. In this view the tongue muscles (though which ones are usually unspecified) are tenser for [i,e,u] than for [ɪ,ɛ,ʊ], thus giving a tense-lax opposition in production between the members of each pair.

(2) Duration (independent of diphthongization). In this view [i,e,u] are of greater duration than [ɪ,ɛ,ʊ], thus giving a long-short opposition between the members of each pair.

(3) Quality change (independent or not of duration). In this view [i,e,u] are characterized by up and forward or back gliding movements of the tongue, while [ɪ,ɛ,ʊ] have no such movements (or movements of insignificant moment). This yields a complex-simple or diphthong-monophthong opposition between the members of each pair.

(4) Jaw opening. In this view the jaw opening for [i,e,u] is less than that for [ɪ,ɛ,ʊ], thus giving a close-open opposition between the members of each pair.

---

*Talk given at the 82nd meeting of the Acoustical Society of America, Denver, Colorado, October 19, 1971.

+Also Herbert H. Lehman College of the City University of New York.

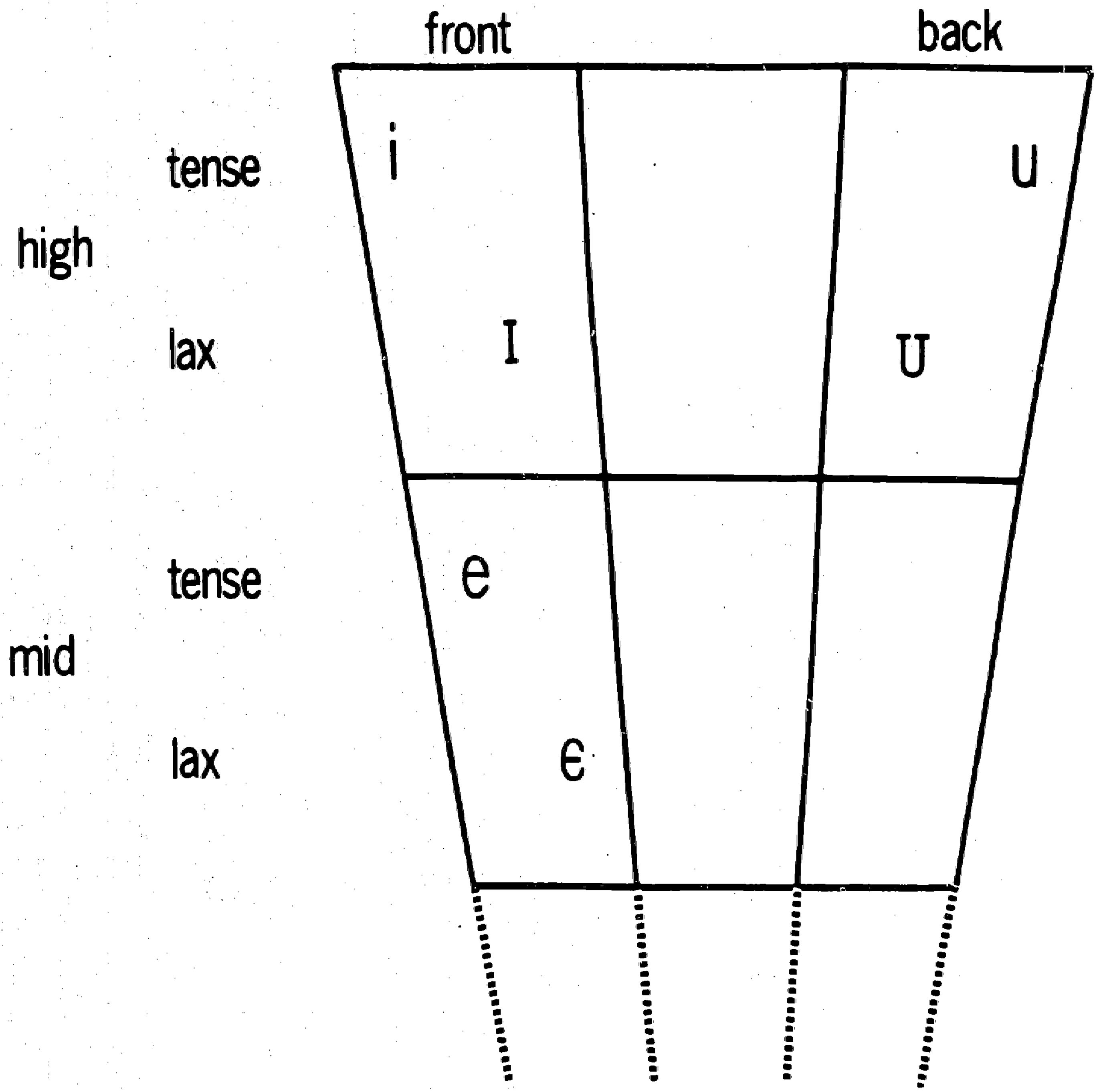


Fig. 1

(5) Simple height difference. In this view the tongue is simply higher for [i,e,u] than for [ɪ,ɛ,ʊ], thus giving a high-low opposition between the members of each pair.

There is no clear agreement as to which of these features (or combination of them) is crucial to the production and perception of the members of the pairs of vowel sounds as different sounds, although any writer who maintains the primacy of one of the features is quite likely to recognize the presence of any or all of the others as redundant. For example, if one were to propose the tense-lax feature as the essential one, he might simply point out that the natural result of such an opposition would be to cause a difference in duration (as the tongue reached the added degree of tension necessary for [i,e,u]); a difference in height and therefore quality (as the added tongue tension caused the muscles to bulge upward, changing the shape of the resonant cavity); and a difference in jaw opening (as the jaw moved up for the tense vowel in tandem with the tensing and rising hump of the tongue). The argument for the primacy of another feature would follow analogous lines.

If we assume a direct correlation between muscular tension and muscular activity, then the feature above most suitably tested by electromyographic techniques is tension. Other features (for example, duration, tongue height, jaw opening) may be inferred from the EMG signal, but muscular activity is more or less directly ascertainable. Furthermore, within reasonable limits of accuracy, we can specify which muscle is being tested, thus bringing a more objective meaning to the term "tongue tension."

There is, of course, a question as to how "muscular activity" is to be interpreted. That is, there are two measures, peak activity and total activity (which would be the integral of the area beneath an EMG curve), both of which are measures of activity. Thus it is possible, if the latter measure, total activity, is taken as primary, that an EMG signal with a relatively low peak but with relatively great duration might be described as indicating more muscular activity than a signal with a relatively high peak but of short duration. In the experiment described here both measures will be referred to separately.

The utterances used in this experiment consisted of the six vowels shown in Figure 1, produced as the vowel in a CVC syllable which was preceded by [ə]. The syllable-initial consonant was always [p] and the final consonants were [p,b,k,g]. Each vowel was paired with each of the consonants, yielding twenty-four utterance types. The syllables were randomized in ten different lists and each list was read twice by the subjects. Approximately seventeen utterances for each item were averaged to produce the EMG curves.

The activity of the genioglossus was chosen as the principle object of investigation in this experiment. The genioglossus is an extrinsic tongue muscle (the largest) which "originates at the point of the jaw...and fans out into the whole anterior-posterior extent of the tongue" (MacNeilage and Sholes, 1964). Earlier experiments have shown it to be active for both front and back vowels, with more activity for high than for low vowels and for front than for back vowels (Hirano and Smith, 1967; Smith, 1970). The data for this muscle (and for others mentioned here) is derived from the EMG signal transmitted by hooked-wire electrodes inserted into the muscle by means of a hypodermic needle. The results for two subjects are reported here.

For both subjects, in all syllables, there was greater genioglossus activity for [i] than for [ɪ], for [e] than for [ɛ], and for [u] than for [ʊ]. Figure 2 shows a typical set of curves for the front series of vowels for one subject. The zero point in time on this and all other figures is the onset of voicing of the stressed vowel of the utterance. The highest EMG peak of activity is for [i], the next highest for [ɪ], a lower one for [e], and the lowest for [ɛ]. The duration of activity for [i] and [e] far exceeds that for [ɪ] and [ɛ], by about 200 msec. Thus both peak height and total activity agree with what the vowel triangle would predict in its tense-lax distinction. It must be noted, however, that there are two peaks for [i] and [e], the first of which is lower than the single peaks for [ɪ] and [ɛ], and that the higher [i]-[e] peaks are reached at a point in time which corresponds to the greater duration of these vowels. Figure 3 shows the same series of vowels for the same subject but in the syllable frame ending in [k]. Again the peak heights are in the same order (although relatively depressed) and in the same durational relation as in the labial syllable, although all the vowels here have only one peak of activity.[1]

Figure 4 shows the two back vowels for the same subject. Both the greater peak height and the greater total activity are evidenced by [u] as opposed to [ʊ]. Analogously with the front vowels, the greater peak for [u] occurs well after the lesser peak for [ʊ].

Figure 5 shows the EMG curves for [i] and [ɪ] in the syllables ending in [p] and [b] for the second subject. In this case all the curves are unimodal, with their peaks occurring very close together in time. The peaks and durations for [i] are again greater than those for [ɪ], so that both measures of muscular activity fulfill the expectations derived from the vowel triangle.

Figure 6 shows the data for [e] and [ɛ] for the same subject and in the same syllable frames as those in Figure 5. The results are much the same as in the case of [i] and [ɪ]. The activity for [ɛ] is remarkably small, neither of the curves for this vowel displaying any clear-cut peak.

The back vowel pair shown in Figure 7 displays the same results as the front vowel pairs: unimodal curves with higher peaks and greater total activity for [u] as opposed to [ʊ]. The curves for [ʊ] are less prominent than for any of the other vowels for this subject.[2]

Despite the uniformly greater activity found for [i,e,u] as opposed to [ɪ,ɛ,ʊ] in both subjects, it appears that the subjects are employing different articulatory strategies in producing some of these vowels. The differences in the ordering of tongue height among the members of the front vowel series

---

[1] It is presumed that the second peak is associated with the final velar consonant.

[2] Incidentally, one can note that this figure and the two before it clearly show the greater duration of muscular activity for the vowels before the voiced than before the voiceless consonants. This greater durational relationship was found without exception in all minimally paired syllables for both subjects.

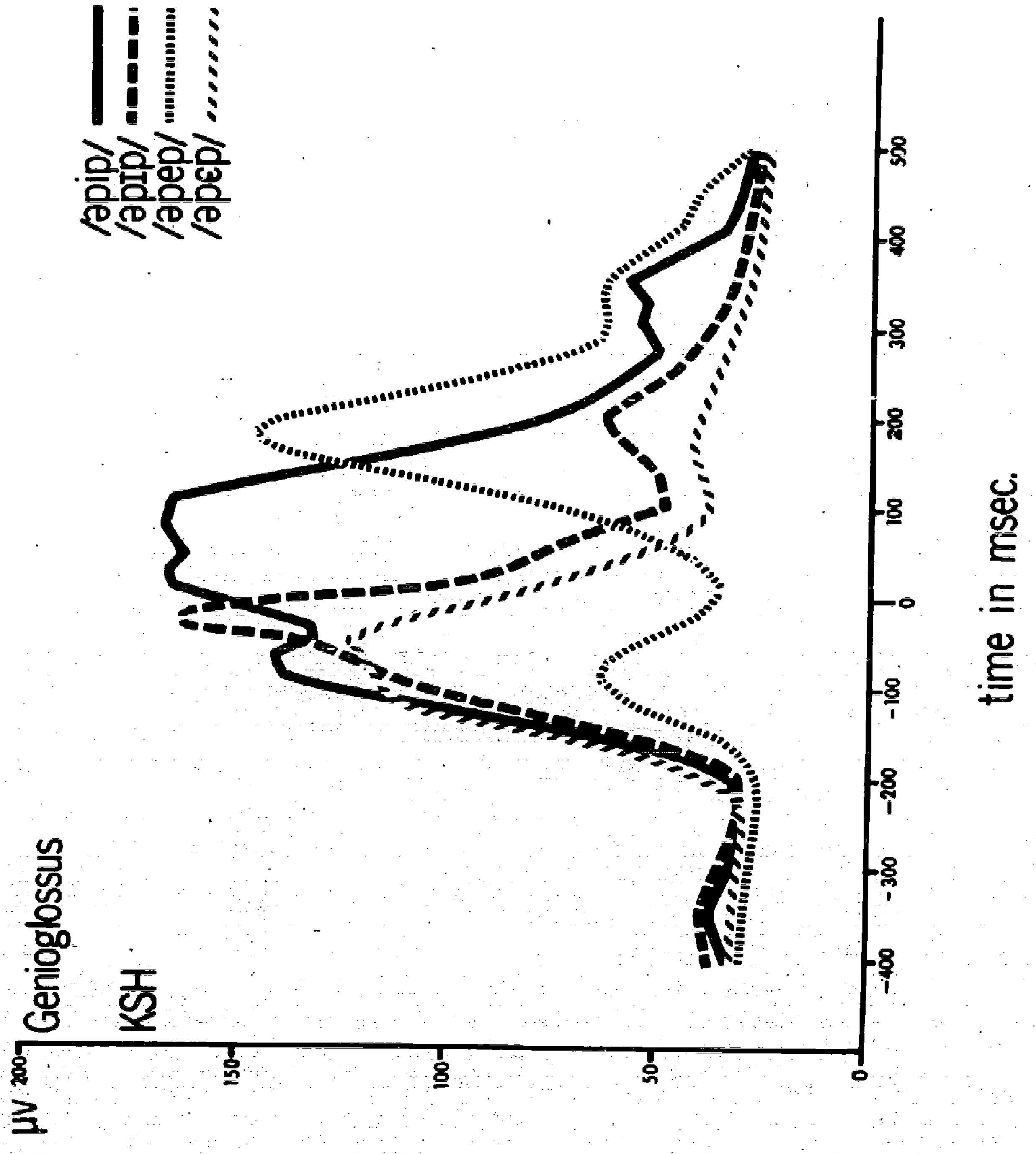

µV
200
150
100
50
0
Genioglossus
KSH
/dide/
/dɪde/
/dede/
/dɛde/
-400
-300
-200
-100
0
100
200
300
400
500
time in msec.

Fig. 2

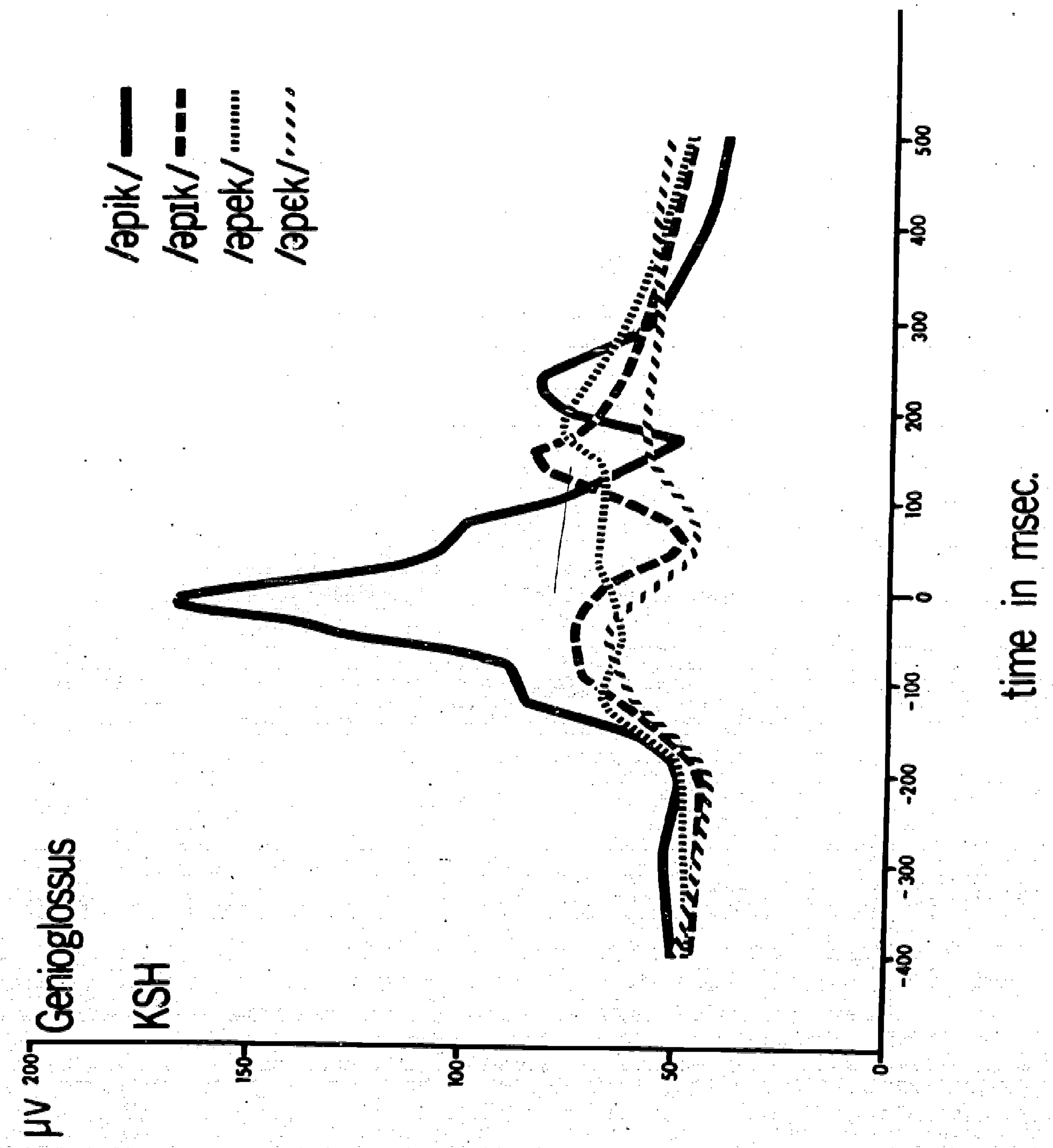

Genioglossus
KSH
μV
200
150
100
50
0
-400
-300
-200
-100
0
100
200
300
400
500
time in msec.
/əpik/
/əpɪk/
/əpek/
/əpɛk/

Fig. 3

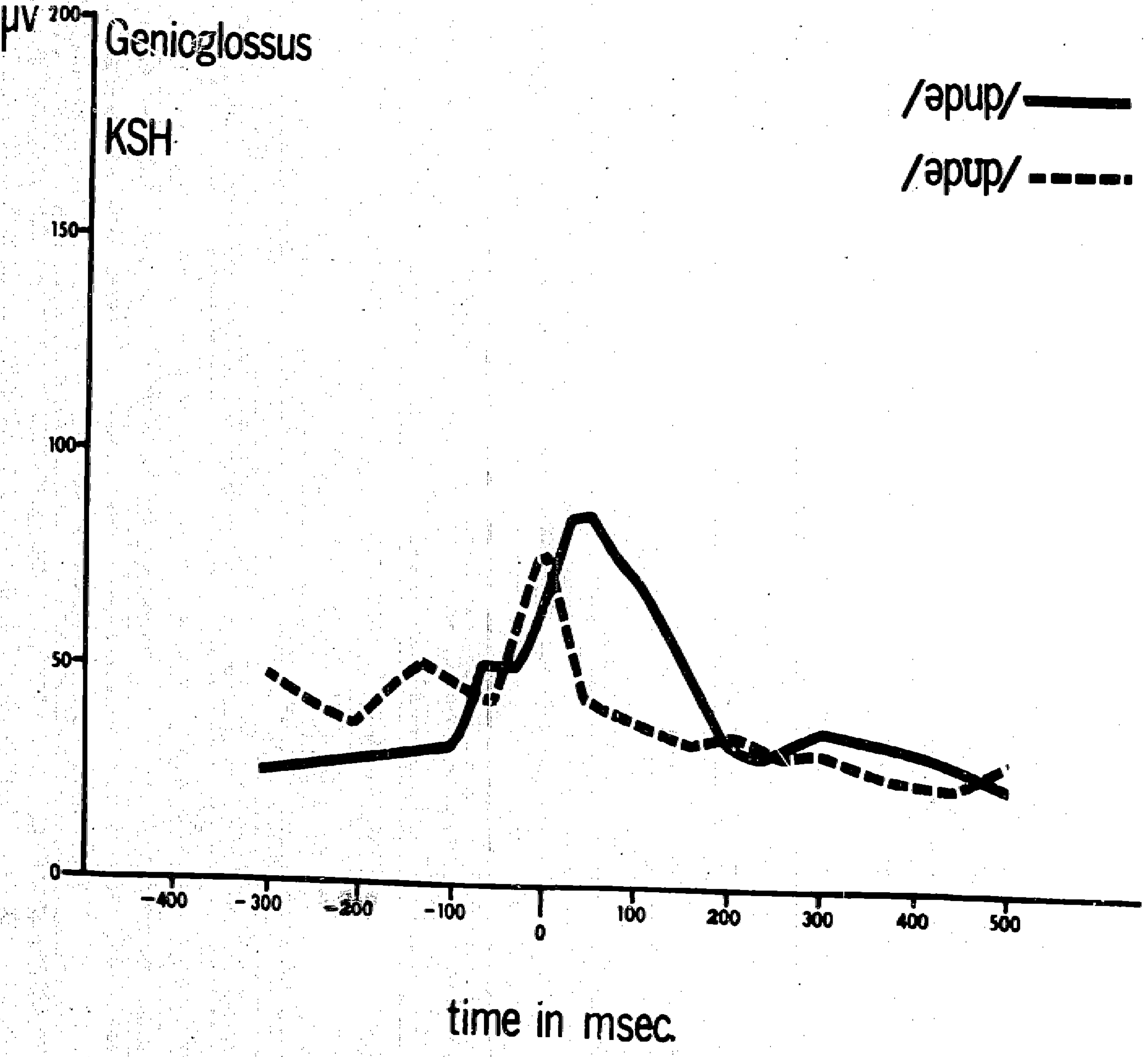

µV
200
150
100
50
0
Genioglossus
KSH
/əpup/
/əpʊp/
-400
-300
-200
-100
0
100
200
300
400
500
time in msec.

Fig. 4

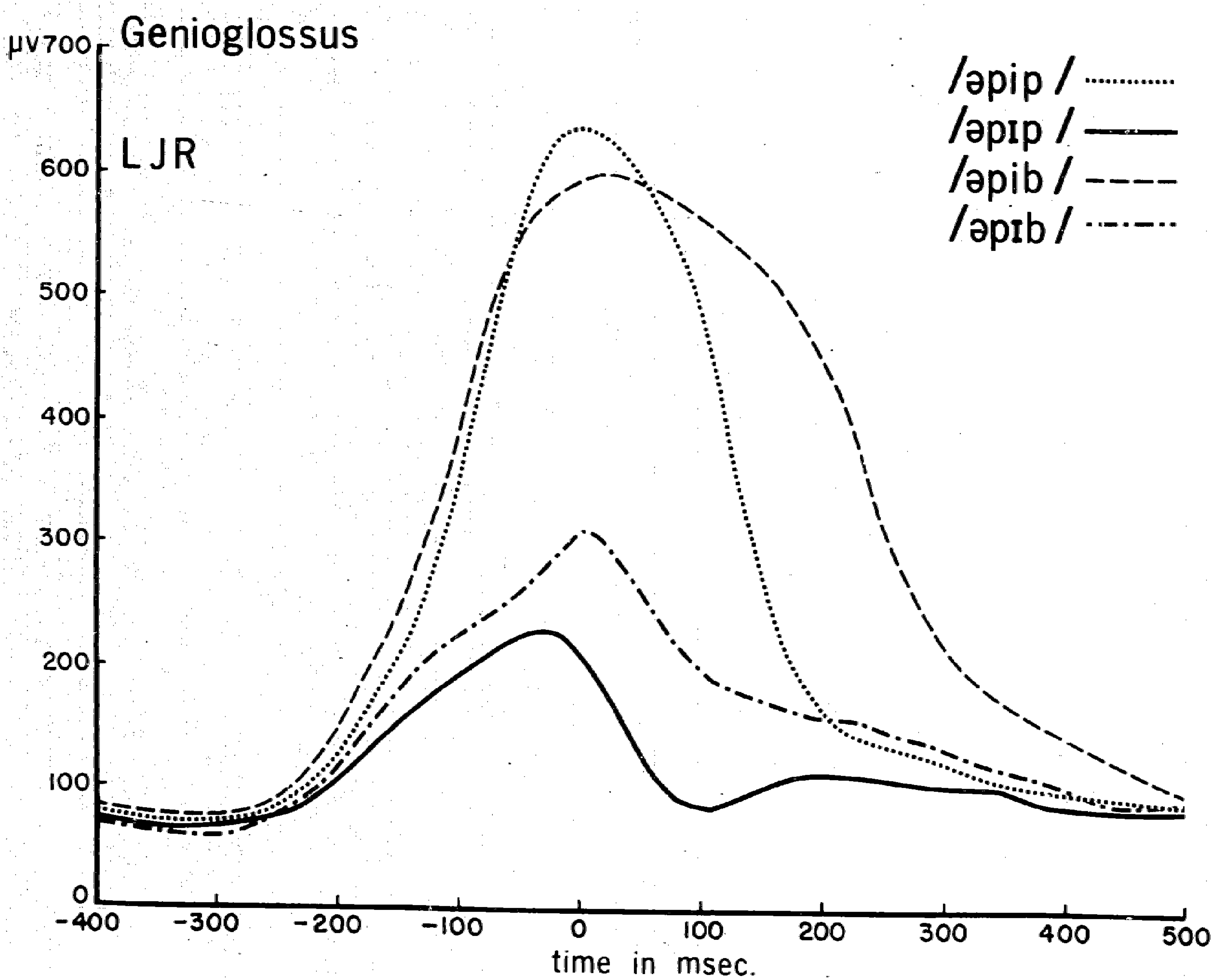

Genioglossus
LJR
µv700
600
500
400
300
200
100
0
-400
-300
-200
-100
0
100
200
300
400
500
time in msec.
/əpip /
/əpɪp /
/əpib /
/əpɪb /

Fig. 5

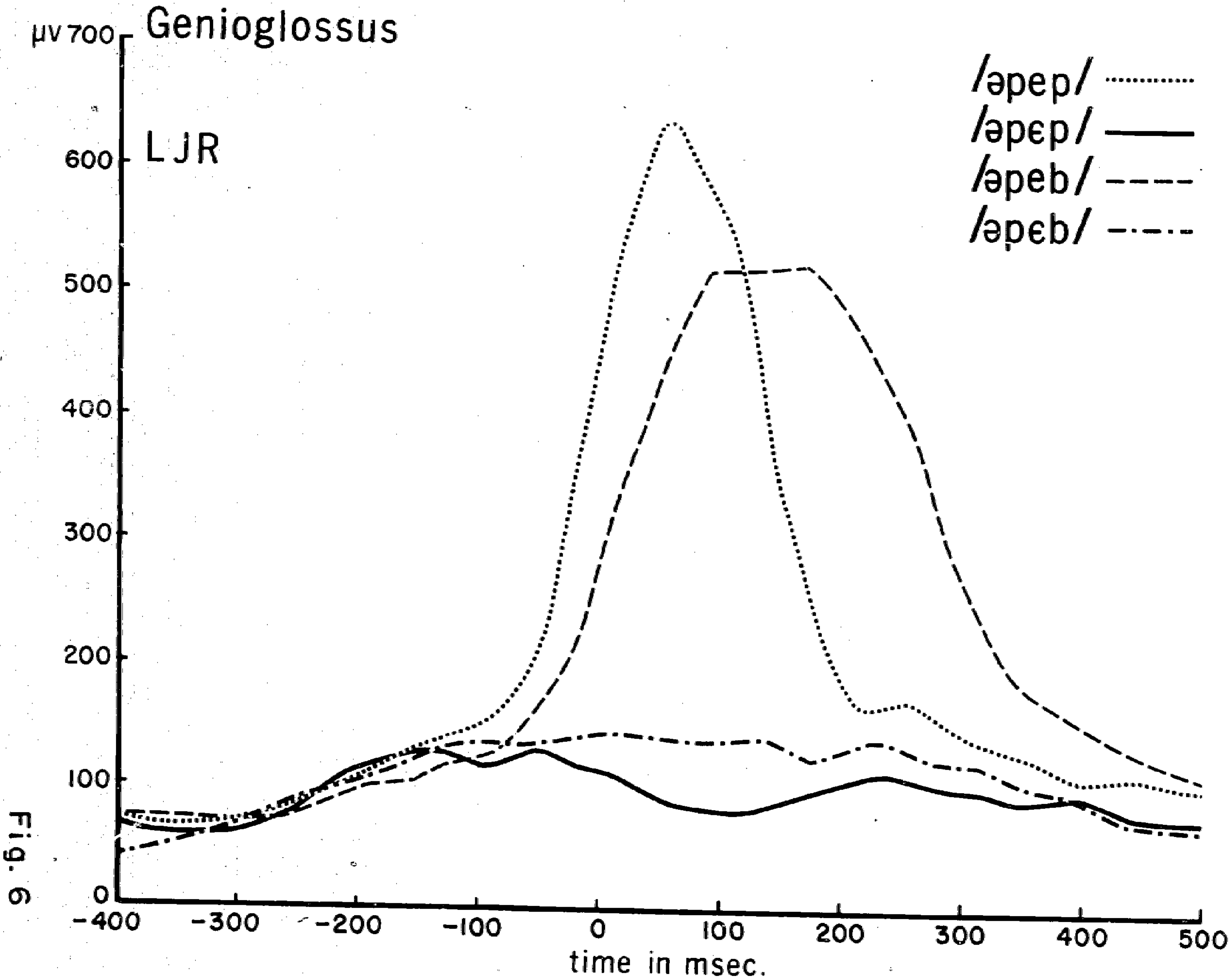

μv
700
600
500
400
300
200
100
0
Genioglossus
LJR
/əpep/
/əpɛp/
/əpeb/
/əpɛb/
-400
-300
-200
-100
0
100
200
300
400
500
time in msec.

Fig. 6

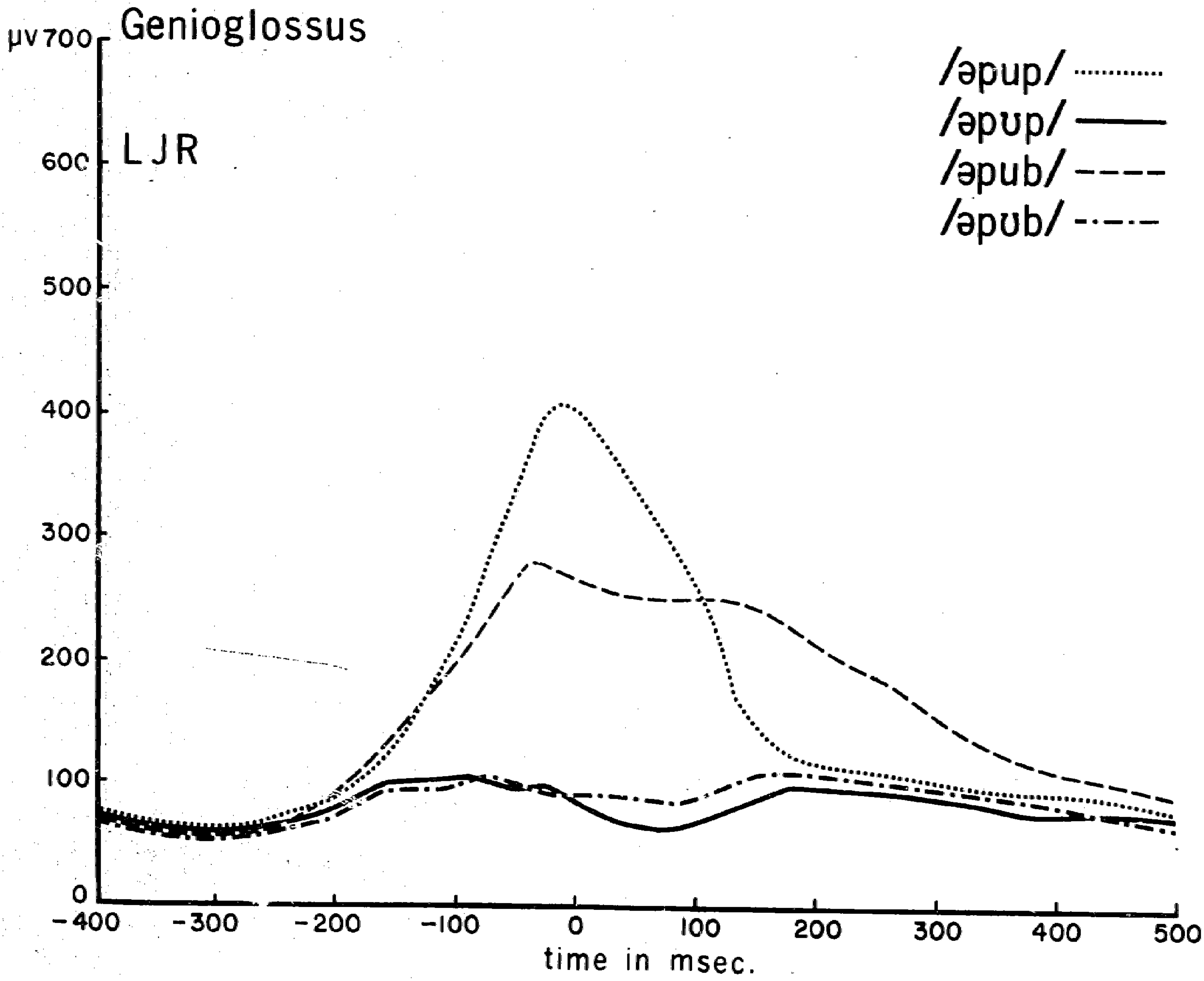

Genioglossus
LJR
μv 700
600
500
400
300
200
100
0
-400
-300
-200
-100
0
100
200
300
400
500
time in msec.
/əpup/
/əpʊp/
/əpub/
/əpʊb/

Fig. 7

as pictured in the vowel triangle is accounted for in the first subject by the action of the genioglossus. The second subject, however, presents a different picture. Figure 8 shows a typical set of curves for the front series of vowels for subject two. Here, and in general, the data for this subject present a picture in which the vowels are arranged in decreasing order of both peak activity and total activity in the order [i,e,ɪ,ɛ], though the differences between [i] and [e] and the differences between [ɪ] and [ɛ] are often small and occasionally in a direction opposite to that shown in Figure 8. In any event, [ɪ] and [e] are clearly transposed from their vowel triangle positions.

A number of possible alternative strategies might explain this apparent discrepancy. One is that for this subject (number two) the superior longitudinal, an intrinsic muscle of the tongue, may be more active in the bunching and raising of the tongue for [ɪ] and [ɛ] than is the genioglossus (MacNeilage and Sholes, 1964; MacNeilage and deClerk, 1969). We have not, as yet, found our way into the intrinsic tongue muscles with hooked-wire electrodes to verify this hypothesis.

Another possible strategy is based on the notion that tongue height may not be solely a function of the genioglossus or of tongue muscles in general, but rather a function of these muscles in conjunction with jaw opening (Lindblom and Sundberg, 1969). Thus, a particular tongue height, measured from the palate to the high point of the tongue, might be effected in more than one way: for example, wide jaw opening with maximum tongue bunching or narrow jaw opening with minimal tongue bunching. If [ɪ] is a high vowel to be paired with [i], and if it is higher than [e], we might expect to find less jaw opening for [ɪ] than for [e] to compensate for the greater tongue bunching for the latter vowel in the second subject. Perhaps the best muscle to tap as an indicator of jaw opening would be the anterior belly of the digastric. Unfortunately this muscle was not investigated for subject two. Another muscle, however, the sternohyoid, whose activity has been described as accompanying jaw opening (Ohala and Hirose, 1970) was tapped. Although the sternohyoid data, in terms of peak height, do show less jaw opening for [ɪ] than for [e], it is by no means certain that the differences between the peaks are sufficient to compensate for the genioglossus activity for the vowels (Raphael, 1971).

A final possibility for the transposition of [ɪ] and [e] as shown in the genioglossus data involves the matter of tongue backing. The vowel triangle shows both [ɪ] and [ɛ] to be retracted from the more extreme front positions of [i] and [e]. Since the genioglossus displays greater activity for the more fronted tongue positions[3] (Hirano and Smith, 1967), one would naturally expect slightly lower values for the activity of this muscle for [ɪ], and of course for [ɛ], if, in fact, these vowels are less front than [i] and [e].

Internal verification of this possibility cannot be definitely provided. The data from the superior constrictor, a muscle which has been taken to be an indicator of tongue backing, frequently does reveal greater peaks for [ɪ] and [ɛ] as opposed to [i] and [e], but the results are not consistent, differences occasionally being small and/or in the unhypothesized direction.

---

[3] Especially for the anterior electrode placement used in this experiment.

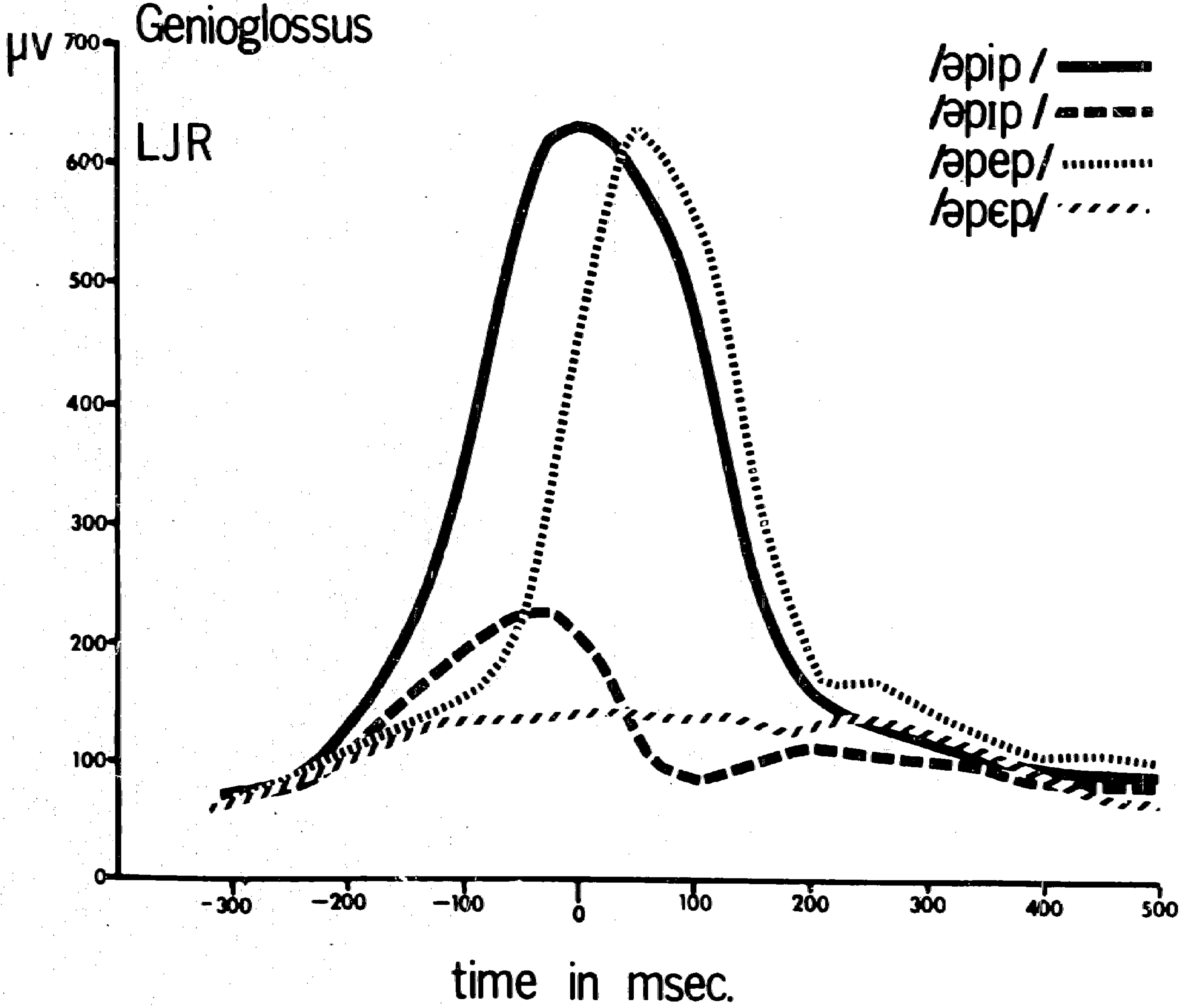

Genioglossus
LJR
μV
700
600
500
400
300
200
100
0
-300
-200
-100
0
100
200
300
400
500
time in msec.
/əpip/
/əpɪp/
/əpep/
/əpɛp/

Fig. 8

Specification of the height ordering of vowels thus awaits further experimentation, involving other muscles and perhaps an advance in technique so that the intrinsic tongue muscles may be more readily and easily investigated.

In conclusion, it appears that there is consistently greater activity on the part of the genioglossus for the vowels [i], [e], and [u] than for their counterparts on the vowel triangle [I], [ɛ], and [U]. However, both in terms of total activity and timing of peak height, this activity is inseparable from duration and quality change. Again, the investigation of other muscles, especially the intrinsic muscles of the tongue, may simplify the picture, but for the moment, although we might justify the assignment of such labels as tense and lax to the vowel categories, qualified in terms of genioglossus activity only, we would make no claims as to the primacy of the feature of tension in distinguishing the production of the vowels investigated.

REFERENCES

Hirano, Minoru and Timothy Smith. (1967) Electromyographic study of tongue function in speech: A preliminary report. UCLA Phonetics Laboratory, Working Papers in Phonetics 7, 45-56.

Lindblom, Bjorn and J. Sundberg. (1969) A quantitative model of vowel production and the distinctive features of Swedish vowels. Royal Institute of Technology, Speech Transmission Laboratory, Quarterly Progress and Status Report 1, 14-32.

MacNeilage, Peter F. and Joseph L. deClerk. (1969) On the motor control of coarticulation in CVC monosyllables. J. acoust. Soc. Amer. 45, 1217-1233.

MacNeilage, Peter F. and George N. Sholes. (1964) An electromyographic study of the tongue in vowel production. J. Speech and Hearing Assoc. 7, 209-232.

Ohala, John and Hajime Hirose. (1970) The function of the sternohyoid muscle in speech. University of Tokyo, Research Institute of Logopedics and Phoniatrics, Annual Bulletin 4, 41-44.

Raphael, Lawrence J. (1971) An electromyographic investigation of the tense-lax feature in some English vowels. Haskins Laboratories Status Report on Speech Research SR 25/26, 131-140.

Smith, Timothy. (1970) A phonetic study of the function of the extrinsic tongue muscles. Unpublished Ph.D. dissertation, University of California, Los Angeles.

Word-Final Stops in Thai[*]

Arthur S. Abramson[+]
Haskins Laboratories, New Haven

As far as I can tell from the literature, agreement over the phonetic nature of word-final stop consonants in Standard Thai has not been reached. Indeed, non-Thai observers with little training and experience in auditory phonetics often have trouble in just detecting the presence of these normally unreleased stops, especially the velar stop after long /uu/. It is perhaps not surprising then that linguists have failed to be very precise in their application of vaguely defined impressionistic terms to these speech sounds.

The question must be examined against the background of the full system of Thai occlusive consonants.[1] Except possibly for my omission of the glottal stop, the phonemes[2] displayed in Table 1 will probably cause little argument. Establishing underlying forms for a generative phonology in Thai grammar is not likely to be relevant to the present phonetic analysis. Rather, it can be argued that it is necessary to have proper phonetic descriptions of utterances _before_ positing underlying forms from which to derive them by rule.

Table 1

Thai Initial Occlusives

	Labial	Dental	Alveolo-Palatal	Velar
Voiced	b	d		
Voiceless Unasp.	p	t	c	k
Asp.	pʻ	tʻ	cʻ	kʻ

---

[*] To be published in a volume on the phonetics of Thai (publication information not yet available).

[+] Also University of Connecticut, Storrs.

[1] These consonants as a set are called occlusives rather than stops only because it seems desirable to include among them the affricates /c cʻ/ which share the feature of aspiration with the simple stops.

[2] I shall restrict myself in this article to considerations of surface phonology.

The only serious disagreement in the literature with the intersecting phonetic features of Table 1 is with regard to the roles of voicing and aspiration. Richard B. Noss (1964:10-13) describes /b d/ as unaspirated lenis stops, /p t c k/ as fortis stops, and /pʻ tʻ kʻ/ as aspirated lenis stops. These appear to be his choices of paramount features, although, of course, he adds information about voicing and other aspects of production. Marvin J. Brown (1965:39), in line with his model for Thai diachronic phonology, posits glottal closure for /p t c k/. That is, he believes that there is a simultaneous glottal closure which is not released later than the oral closure.

Claims about tensity as an independently controllable parameter of speech production are tenuous (Abramson and Lisker, 1970) when one seeks experimental validation. The story on voicing and aspiration seems to be quite different. Some several years ago, Leigh Lisker and I (Lisker and Abramson, 1964) showed clearly through acoustic measurements that Thai initial stops are differentiated into three classes on the basis of voice onset time measurements. For the "voiced" stops, spectrograms show high-amplitude laryngeal pulsing during the stop occlusion; that is, the pulsing starts well before the release of the stop; for the "voiceless unaspirated" stops, voicing starts upon release of the stop or shortly thereafter; for the "voiceless aspirated" stops, voicing onset lags considerably behind stop release. The original body of data underlying these conclusions is reproduced in Table 2 (Lisker and Abramson, 1964: 396).[3] The use of negative numbers for /b d/ is simply a convention to indicate voice onset before our reference point, the release of the stop. Aspiration, then, is the acoustic consequence of exciting the vocal tract resonances by means of a noise source, turbulent air coming through the open glottis during the lag between the release of the stop and the onset of voicing. During this voicing lag, the articulator is moving away from its place of articulation, and the vocal tract is assuming a configuration for the syllabic vowel; thus, aspiration is a property of the initial portion of the vowel as well as the stop release. This turbulence is in fact also present in the short voicing lags of the voiceless unaspirated stops but has too short a duration to be very audible.[4] The voicing lead of /b d/ is typically quite audible. The perceptual relevance of voice onset time has been confirmed for Thai through experiments with synthetic speech (Abramson and Lisker, 1965; Lisker and Abramson, 1970).

Given the rather compelling efficacy of voice onset time in implementing the three-way contrast, any as yet unsubstantiated claims concerning tensity or fortisness seem gratuitous at this time. On the other hand, Brown's assertions as to glottal closure are not necessarily inconsistent with our observations on voice timing. One way to suppress phonation in speech is to swing the anterior portions of the arytenoid cartilages apart and open the

---

[3] In the 1964 cross-language study, we restricted our observations to stops; therefore, Table 2 contains no data on the affricates. Since then, however, I have seen enough additional spectrograms not only to confirm our old analysis of the Thai stops but also to validate aligning the two affricates with the voiceless unaspirated and aspirated stops, respectively.

[4] If short voicing lag is effected in part by means of a small glottal aperture (Kim, 1970), we might expect the turbulence to be low in intensity. Low intensity and short duration would combine then to yield less loud aspiration than for the aspirates.

Table 2

Thai Initial Stops: Voice Onset Time in Milliseconds
(Three Speakers)

Labials

	b	p	pʻ
Average	-97	6	64
Range	-165:-40	0:20	25:100
Number	31	32	33

Dentals

	d	t	tʻ
Average	-78	9	65
Range	-165:-40	0:25	21:125
Number	33	33	33

Velars

		k	kʻ
Average		25	100
Range		0:40	50:155
Number		32	38

glottal aperture beyond the point at which audible vocal-fold vibration can occur; another way is to close the glottis tightly. We hope soon to be able to settle such questions through the use of our flexible fiberoptic endoscope.[5] In the meantime, recent findings (Kim, 1970), for Korean at least, indicate that stops heard as voiceless inaspirates are likely to be produced not with tight closure but with a small opening of the glottis, while larger glottal apertures are required for greater amounts of aspiration.

In word-final position the phonological picture is somewhat simplified. The two affricates do not appear, and the three-way laryngeal opposition among

---

[5] See Cooper et al. (1971) and references therein.

the stops is not relevant. For each place of articulation there is just one stop phoneme: a labial, a dental, and a velar. What is the phonetic nature of this single manner category? From the point of view of the language structure, it may not matter too much; there is neutralization of the distinctive features involved. From the point of view of speech production, it does matter very much. After all, a good description of the language must include rules for pronunciation. If we suppose that these final stops are to be aligned with one of the three initial categories, let us consider the phonetic possibilities in the light of the voice timing dimension which is diagnostic for initial position. Long voicing lag or aspiration is ruled out by the fact that the final stops are not normally released, nor, for that matter, is preaspiration observed in Thai. We are left then with the mirror images of the laryngeal states of the voiced and voiceless unaspirated categories: (1) voice pulsing continues well into the stop occlusion or (2) voice pulsing ceases by the time of achievement of oral closure.

As suggested at the outset of this article, observers using purely auditory criteria have not presented very convincing pronunciation rules for the use of analysts and students of the language. One Thai writer (Rudaravanija, 1965) writes the final stops as voiceless unreleased /p t k/ in the belief that they are voiceless. Another Thai scholar (Kruatrachue, 1960:50) labels these final consonants as /p t k/ but describes them as "varying from their allophones in initial position in not being released and in being less tense or fortis." Brown (1965) writes all Modern Central Thai examples with final /p t k/, but for purposes of his historical treatment he is not necessarily matching them with initial /p t k/.

Two recent major reference works that must be taken into account in any present-day linguistic description of Thai see these final stops in a different light. For Noss (1964:10-13), the final stops share the "unaspirated lenis" feature of his initial /b d/; therefore, this necessitates positing an additional phoneme /g/, which appears only in final position. Now in his fuller phonetic specification of these consonants, Noss does say that they are fully voiced in initial position--the two that occur there--but that they are normally voiceless in final position and occasionally voiced, especially after a long high vowel. We must recall here that for Noss the primary distinction between the sets /b d g/ and /p t c k/ is based on the fortis/lenis feature rather than voicing. In the table of consonants in her dictionary Mary R. Haas (1964:xi) also posits a phoneme /g/ which occurs together with /b d/ in word-final position in her illustrative examples and in the dictionary entries. No phonetic comment is made, so one is led to believe that these stops are voiced in both positions.[6] My own experience with the Thai language has never led me to any conviction that I can hear laryngeal pulsing during the occlusions of final stops, so in my own phonemic or morphophonemic transcriptions I have always written them as /p t k/; nevertheless, in an early noninstrumental assessment of the consonants (Abramson, 1962:4), probably under the influence of Haas, I was reluctant to take a firm position and wrote in a Praguian fashion, "the view taken here is that there is a neutralization of the manner features at the end of a syllable with the archiphonemes written as /p t k/, occurring as [p t k] or [b d g]."

---

[6] This is consistent with her position in textbooks and other publications dating from 1945, too numerous to be cited here.

In the light of the foregoing, it seemed to me that it would be best at this time to approach the problem by examining the final stops acoustically in terms of the voice timing dimension that had proven so efficacious in initial position. Having on hand extensive samples of speech recorded by six educated native speakers of Central Thai, I went through all these tape recordings looking for words with final stops.[7] The speakers were university students, four men and two women, recorded between 1964 and 1971. In these recordings, made for a variety of purposes but not specifically the present one, I found a total of 140 word-final stops as displayed in Table 3. For each stop the number of tokens examined is given together with its percentage of the total. Indeed, not only is the numerical representation of types uneven as shown in Table 3, but also the array of environments in which the stops were found. That is, I simply looked for word-final stops wherever I could find them in the recordings: isolated words, citation forms of short expressions and sentences, and passages of running speech. Of course, it would have been possible to have a few informants record all the vowels of the language followed by the three consonants to form a complete paradigm.[8] My own feeling was that such an approach would achieve statistical symmetry at the price of a certain artificiality. I agree that this kind of artificiality may sometimes be necessary in linguistic investigations and even desirable, but since sufficient recordings were available to provide, as it turned out, a rather stable set of data, it seemed preferable not to call an informant's close attention to my interest in the final stops.

Table 3

Final Stops Examined

(Six Speakers: four men and two women)

	N	% of Total
/p/	18	13
/t/	45	32
/k/	77	55
Total	140	100

---

[7] I included some of these data in my review of the Haas dictionary (Abramson, 1966). This review will appear in the public domain if Volume 22 of *Word* is ever published.

[8] Naturally, as in all experimental work, anyone is free to test the generality of my results with a change in experimental design.

Wide-band sound spectrograms of all the utterances were examined for acoustic signs of laryngeal pulsing during the closures of the final stops. If there was some ambiguity as to the presence of vertical striations in the spectrograms at the fundamental frequency of the speaker's voice, especially in the samples embedded in running speech, narrow-band spectrograms were inspected as well to examine the harmonics of the voice for continuity. For the most part, the wide-band spectrograms were sufficient and preferable because of their better time resolution.

I have divided my observations of the word-final stops into the two broad classes of those occurring at the end of an utterance and those occurring embedded within an utterance. To test for voicing it should really be enough to present data on utterance-final stops since the claims in the literature seem to be intended to apply to "optimum" citation forms. I, however, wished to examine the possibility that these stops might show a definite trend toward the voiced state by progressive assimilation to a following voiced environment, while manifesting themselves as voiceless consonants in utterance-final position and before voiceless phones. Nothing in the data indicated any profit in distinguishing between utterance-final stops in citation forms from those in running speech. In running speech, any clearly marked pause or end of discourse was accepted as a sign of an utterance-final stop. The utterance-medial word-final stops appearing before voiced phones were distinguished from those appearing before voiceless phones. The results of this investigation are presented in Table 4, which shows the number of items examined for each class and the number and percentage of those for which voicing of the stop occlusions appeared in the spectrograms.

Table 4

Laryngeal Pulsing in Final Stops

	Number Examined	Voicing Present
Utterance-final	73	2 (3%)
Utterance-medial		
Before voiced phones	32	5 (16%)
Before voiceless phones	35	1 (3%)
Totals	140	8 (6%)
After long high vowels	28	2 (7%)

The data of Table 4 make it overwhelmingly clear that the only reasonable statement of a phonetic rule for word-final stops in Thai, regardless of the context, is that they be produced without voicing. Note that after the totals in the table, I have an extra entry for the stops found after long high vowels. This was done because of Noss's claim that these in particular are likely to be voiced. In fact, the two that were voiced (7 percent of 28) fall among the five that were voiced before voiced phones. It should be noted that even before voiced phones the tendency to voice the occlusions of the stops is rather weak, only 16 percent. In general, the 6 percent of the total that showed voicing is characterized by low-amplitude pulsing of the kind that we have previously called "edge vibration" (Lisker and Abramson, 1964:416-18) and would normally expect to be weak continued oscillation of the vocal folds while the glottis is opening; edge vibrations of this kind seem usually to be below auditory threshold (Lisker and Abramson, 1967:8-9). Examination of the spectrograms convinces me that this is the situation, but I do not have the precise amplitude measurements that would entitle me to make such a distinction in Table 4. In only two of these instances was the voice pulsing a convincing mirror image of the situation in word-initial voiced stops. Both of these were utterances of the dental stop in the expression /pʻûut dûaj/ in which apparently a real [d] was pronounced throughout the single sustained stop occlusion ending the first word and beginning the second.

Voice pulsing, then, clearly is not characteristic of word-final stops in Thai. The rare instances of unbroken high-amplitude laryngeal pulsing in this body of data were cases of assimilation to following homorganic stop sounds. Otherwise, the several cases observed seem to be nothing more than the weak, inaudible pulsing caused by the failure of the margins of the glottis to cease oscillating completely when the glottal aperture is not large; although normally too weak to be heard in a speech context, these pulses may have sufficient intensity to be detected by instruments. On the basis of available phonetic data, it is implausible to align word-final stops in Thai with anything but initial /p t k/.

It is unfortunate that such important reference books as the Haas dictionary and the Noss grammar can mislead students of the language as to one aspect of Thai pronunciation. Admittedly, some of the speculations of Brown as to glottal control and Noss as to the state of the supraglottal articulators should be investigated by instrumental means now at our disposal. With the knowledge of these phonomena in general and Thai phonetics in particular now available, however, I simply wish to assert that there is as yet no basis for denying the primacy of the timing of laryngeal control of voicing--and thus aspiration--for both initial and final stops of Thai.

REFERENCES

Abramson, Arthur S. (1962) The Vowels and Tones of Standard Thai: Acoustical Measurements and Experiments. (Bloomington: Indiana U. Res. Center in Anthro., Folklore and Linguistics) Pub. 20.

Abramson, Arthur S. (1966) Review of Thai-English Student's Dictionary by Mary R. Haas. Word 22 (in press).

Abramson, Arthur S. and Leigh Lisker. (1965) Voice onset time in stop consonants: Acoustic analysis and synthesis. Proc. 5th Intl. Cong. on Acoustics, D. E. Commins, ed., Vol. Ia, A51. (Liège: Imp. G. Thone).

Abramson, Arthur S. and Leigh Lisker (1970) Laryngeal behavior, the speech signal and phonological simplicity. Actes du X^e^ Congrès International des Linguistes, Vol. IV, 123-9. (Bucarest: Editions de l'Académie de la République Socialiste de Roumanie.)

Brown, Marvin J. (1965) From Ancient Thai to Modern Dialects. (Bangkok: Social Science Association Press of Thailand.)

Cooper, Franklin S., Masayuki Sawashima, Arthur S. Abramson, and Leigh Lisker. (1971) Looking at the larynx during running speech. Annals of Otology, Rhinology and Laryngology 80, 1-5.

Haas, Mary R. (1964) Thai-English Student's Dictionary. (Stanford: Stanford University Press.)

Kim, C-W. (1970) A theory of aspiration. Phonetica 21, 107-116.

Kruatrachue, Foongfuang. (1960) Thai and English: A comparative study of phonology for pedagogical applications. Ed.D. dissertation, School of Education, Indiana University.

Lisker, Leigh and Arthur S. Abramson. (1964) A cross-language study of voicing in initial stops: Acoustical measurements. Word 20, 384-422.

Lisker, Leigh and Arthur S. Abramson. (1967) Some effects of context on voice onset time in English stops. Language and Speech 10, 1-28.

Lisker, Leigh and Arthur S. Abramson. (1970) The voicing dimension: Some experiments in comparative phonetics. Proc. 6th Intl. Cong. Phon. Sci., Prague, 1967 (Prague: Academia), 563-7.

Noss, Richard B. (1964) Thai Reference Grammar. (Washington, D.C.: Foreign Service Institute, Dept. of State).

Rudaravanija, Panninee. (1965) An analysis of the elements of Thai that correspond to the basic intonation patterns of English. Ed.D. dissertation, Teachers College, Columbia University.

Audible Outputs of Reading Machines for the Blind[*]

Franklin S. Cooper, Jane H. Gaitenby, and Ignatius G. Mattingly[+]
Haskins Laboratories, New Haven

The goal of research on reading machines for the blind at Haskins Laboratories is to produce by machine methods an output of clear, audible English from an input of ordinary printed text. The core problem--generating acceptable speech from phonetic spellings--seems very near a successful solution through synthesis-by-rule methods. There is still much to be done by way of evaluating and improving the synthetic speech, but the research can now turn to some of the other problems involved in setting up a complete Reading Service Center for the blind. Thus, the present emphasis is on user tests of speech synthesized by rule, improvements in the rules (and so of the speech), and automation of the entire speech-generating process.

Evaluation by Blind Users

An article in the previous Bulletin described the completion of user trials with Compiled Speech (another kind of spoken output in which sentences are constructed from single prerecorded words). Preliminary tests were reported, also, comparing Synthetic Speech with Compiled Speech and indicating that Synthetic Speech was much preferred. The present report deals exclusively with speech that has been synthesized from phonetic spellings by various combinations of rules for synthesis; in all cases, the major part of the conversion and the generation of the tape recordings has been done by computer.

Some further testing has been done with veterans attending the Eastern Blind Rehabilitation Center, Veterans Administration Hospital, West Haven, Connecticut, with results similar to those already reported. In addition, a committee on blind students at the University of Connecticut has become interested in developing a reading center for blind students that will make use of the methods developed at Haskins Laboratories. The University has assigned a member of the faculty to help in evaluating the synthetic speech and has provided a student assistant to help in generating additional recordings for this purpose. This permits sample chapters from textbook assignments to be prepared during the coming summer. Hence, the user evaluation program is moving ahead vigorously, with present emphasis on accumulating recorded materials for student use, starting in September.

---

[*] Summary prepared for the Bulletin of Prosthetics Research BPR 10-16, Fall 1971.

[+] Also University of Connecticut, Storrs.

Automating Text Preparation

The preparation of synthetic speech recordings for user evaluation is, at present, a rather slow process since the printed text must be typed into the computer in phonetic form and sentence stress and intonation marks must be supplied by the typist. Earlier *Bulletins* have carried accounts of the phonetic input system, based on a keyboard-plus-storage oscilloscope, coupled directly to the computer.

Recent improvements in the phonetic input facility have been of two kinds. First, the editing capability has been expanded and streamlined. This allows the operator to make changes in the phonetic text both quickly and efficiently. The modified text can be synthesized immediately in order to evaluate any changes that have been made. A second major change pertains to the recording of long passages of text. Until quite recently, the task of producing an audio tape with synthesized text involved much laborious hand-editing. This process has been made automatic, i.e., audio tapes are produced under computer control in a form that is suitable for listening and evaluation with little or no further editing.

Improving the Naturalness of Synthetic Speech

Veterans and students who have listened to synthetic speech often report that it has an "accent," but that it is completely intelligible after the first few sentences. Some texts require attentive listening, and some of the listeners are not sure that the synthetic speech will be easy to listen to for long, unbroken periods of time. Many others find the speech fascinating and fun and say that even human readers cannot be listened to indefinitely. Clearly, though, improvements in naturalness can and should be made. Work along this line has been directed partly to the phonetic details of the synthesis-by-rule program itself, and partly to extensions of the rules that will mechanize the remaining stages of the speech-generating process.

Modification of the synthesis-by-rule program has concentrated on the details of the allophone tables and on the application of rules for intonation. Improvements have been made in the acoustic specification of duration and amplitude for stop consonants and for clusters of consonants in diphthongs with various stresses and in modulating the intonation over a less extreme range than before.

Rules have been developed for assigning and modifying word stress in sentences. It is a happy fact that English word stress is essentially stable, even when words appear in sentence context, though the acoustic realization of stress has to be modified to take account of word sequences, word location in breath group, and sentence intonation. Plain and sparse rules for stress modification have been applied to several thousand words of text, yielding speech that departs only rarely from expected rhythms and phrasing. A number of stress problems remain, some of them due to the multiple grammatical usage that is possible for many English words.

Rules for assigning pauses within sentences have been developed also. These depend in part on punctuation and in part on the number of words in a string and their syntactic functions. [The original rules for synthesis (fully computerized) were quite successful in realizing stresses and pauses when the input phonetic string was suitably marked by the human typist. The objective of the present rules is to develop algorithms (for later conversion to computer programs) that will automate the marking process.]

Automating a Pronouncing Dictionary

A major step in the conversion of printed English into spoken English is the derivation of the phonetic string on which the rules for synthesis will operate: the spelled form of the word must be converted to its pronounced form in phonetic symbols, and to information about its normal syntactic function(s), for use in those rules that assign stress and pause. This part of the problem is being solved by the use of a comprehensive pronouncing dictionary with syntactic annotations. A dictionary of this kind has been made available through the kind cooperation of the Speech Communications Research Laboratory; some parts of it are already in hand, and the remainder is expected by the end of the summer.

The total dictionary (as received) will contain on the order of half a million entries. It corresponds in coverage to the ordinary collegiate dictionary but has many more entries, since all inflected forms of the words are entered explicitly. Thus, in addition to such normal nouns as *cat*, there are also *cats*, *cat's*, and *cats'*. Similarly for verbs, there are such entries as *walk*, *walks*, *walked*, and *walking*. The dictionary also contains separate pronounciations for different dialects and grammatical categories. Many of these variants are not wanted in a dictionary for the projected Reading Service Center; hence, there is a substantial task involved, not only in programming for normal uses of the dictionary, but also in developing algorithms to delete the unwanted material. A substantial part of the "editing" has been done, and revised versions are being prepared for the portions of the dictionary that are in hand. It now appears that the dictionary, in final form for reading machine use, will fit comfortably onto the four discs that are a part of the Laboratories' computer installation, i.e., the entire dictionary will be available for fast random search.

Planning for a Reading Service Center

The interest and cooperation of the University of Connecticut in connection with its own program for blind students makes it feasible to plan for the establishment of a Reading Service Center at the University, probably as an extension of the University's present library services to blind students. A schedule has been set for an initial trial period during which the feasibility of such a Center will be fully assessed and equipment needs and budgets for the Center's implementation will be developed. Plans for the trial period call for substantial quantities of synthetic speech to be synthesized at Haskins Laboratories during the 1971-72 academic year. The text will be drawn from the blind students' normal reading assignments. The amount of recorded material that can be provided during the final months of 1971 will be limited by the time required to type phonetic strings into the computer. By early 1972, the automated dictionary should be operating. Any typist can then use a conventional Selectric typewriter to prepare the input text in machine-readable form; thus, the amount of material available for evaluation during the second semester should be very substantial, allowing a thorough evaluation of the utility of the projected Reading Center. During the same period--assuming that the user tests are progressing toward an encouraging conclusion--planning and engineering studies will be made to determine the type and cost of computer and optical character recognition equipment that will be needed for a full-scale Reading Service Center. The objective of the first phase is to have, by mid-1972, all the necessary data for a policy decision on whether or not to proceed with the implementation of a Center. The user trials and equipment planning for a Center for blind university students is, of course, directly applicable to decisions about a Reading Service Center for blind veterans.

The Evolution of the Human Speech Anatomy*

Philip Lieberman+
Haskins Laboratories, New Haven

Let me state at the start of this paper the two main points that I hope to cover.

First, adult Homo sapiens has a species-specific vocal tract that is necessary for producing the sounds of human speech. The sounds of human speech are necessary for human language. They are not arbitrary; they make rapid acoustic communication possible.

Second, enhanced linguistic ability was the conditioning factor in the processes of natural selection that led to the evolution of the human vocal tract. In other words, the human vocal tract evolved for the function of speech. The human vocal tract is inferior to the nonhuman vocal tract with respect to the vegetative functions of breathing, swallowing, and chewing. The only function for which the human vocal tract, i.e., the oral cavity, pharynx, larynx, and nose, is superior is generating the full range of sounds of human speech. The morphology of the base of the skull of Homo sapiens reflects the process of mutation and natural selection that resulted in the development of human speech. Human speech is as important a factor in the late stages of human evolution as chewing and upright posture are in its early stages.

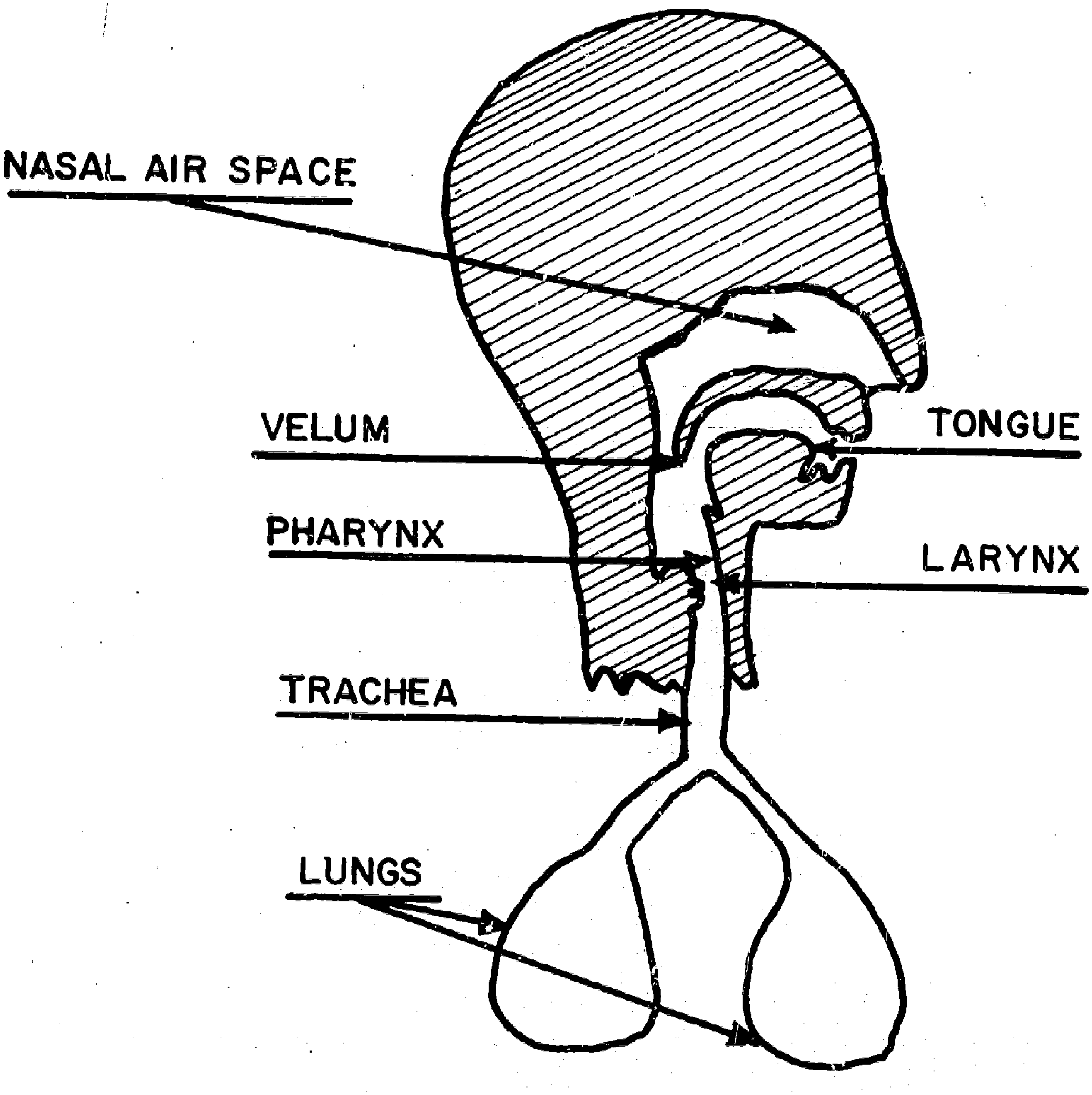
Acoustics and Anatomy of Human Speech

Let us start by briefly reviewing the acoustic and anatomic bases of human speech. Human speech acoustically is the result of a process in which a source of energy is modified by an acoustic filter. In Figure 1 a schematic view of the human vocal tract is presented. A sound like the vowel /a/ in the word father is produced by exciting the supralaryngeal vocal tract by means of puffs of air that issue from the larynx. For a typical adult male these puffs of air, which we perceive as the fundamental frequency, or pitch, of a person's voice, occur at rates of 100 to 300 puffs per second. The rate at which the larynx rapidly opens and closes can, of course, be adjusted during speech. The vowel /a/, for example, can be produced with a fundamental frequency of 120 cycles per second, or 200 cycles per second. The vowel still has the phonetic value of /a/. The phonetic characteristics of /a/ are independent of the laryngeal source. They are instead determined

---

*Invited paper presented at the Institute on Behavior and Evolution of the 70th annual meeting of the American Anthropological Association, New York, New York, 18 November 1971.

+Also University of Connecticut, Storrs.



SCHEMATIC OF RESPIRATORY SYSTEM

Figure 1: The supralaryngeal respiratory system, which determines the phonetic quality of vowels and consonants, consists of the oral and nasal cavities and the pharynx.

by the shape of the supralaryngeal airways. In Figure 2, line spectra are presented for the vowel /i/ as it was produced by the same speaker at two different fundamental frequencies. Note that energy is present at harmonics of the fundamental frequency. Note the presence of local energy maxima in the spectra at about 300 and 2200 cycles per second. These local energy maxima are determined by the formant frequencies of the supralaryngeal vocal tract. The formant frequencies are determined by the resonances of the supralaryngeal vocal tract. At these resonant frequencies the harmonics of the laryngeal source will pass through the filtering supralaryngeal vocal tract with least attenuation. Different shapes of the supralaryngeal vocal tract can result in different formant frequencies. The vowels /a/, /i/, /u/, /I/, and /ʌ/, for example, all have different formant frequencies. The inventory of phonetic elements of human languages is largely achieved by changes in the supralaryngeal vocal tract that result in different formant frequencies.

A useful musical analog to this aspect of speech production is a pipe organ. The source of acoustic excitation is similar for all the pipes. The quality of each musical note is determined by the length and shape of each pipe. We could assess some of the limitations on the music-producing capability of a particular pipe organ by examining the pipes, independent of the excitation source that the particular pipe organ actually used. We might, for example, find this partial assessment useful in reconstructing the structure of some archaic music that was written for a particular pipe organ. We would not, of course, be able to say very much about the dynamic control of the pipe organ, but we would know some of the constraints that would structure the music.

In a similar way we are now in a position to assess some of the limitations that structured the speech of extinct hominids by reconstructing and modeling their supralaryngeal vocal tracts even though we cannot say very much about their laryngeal sources or the dynamic control of their speech-producing apparatus.

The Reconstruction of the Supralaryngeal Vocal Tract

The reconstruction of the supralaryngeal vocal tract of an extinct fossil hominid at first appears to be impossible since the soft tissue of the oral cavity and pharynx is not available. We fortunately can make reasonable reconstructions by using the methods of comparative anatomy and taking advantage of skeletal similarities that exist between living primates and fossil remains. The details of these reconstructions which are largely the product of my colleague Professor Edmund S. Crelin of the Department of Anatomy of Yale University Medical School are described in our recent and forthcoming papers, (Lieberman and Crelin, 1971; Lieberman et al., in press). We have examined, reconstructed, and assessed the phonetic capabilities of a number of fossil hominids, but I shall limit the present discussion to the fossil "classic" Neanderthal man of La Chapelle-aux-Saints.

In Figure 3 lateral views of the skulls of newborn man, adult chimpanzee, the La Chapelle-aux-Saints Neanderthal man, and adult modern man are presented. The skulls have all been drawn to appear nearly equal in size. Skull features of newborn man, chimpanzee, and Neanderthal man that are similar to each other, but different from that of adult modern man, are as follows: (A) they have a generally flattened out base; (B) they lack

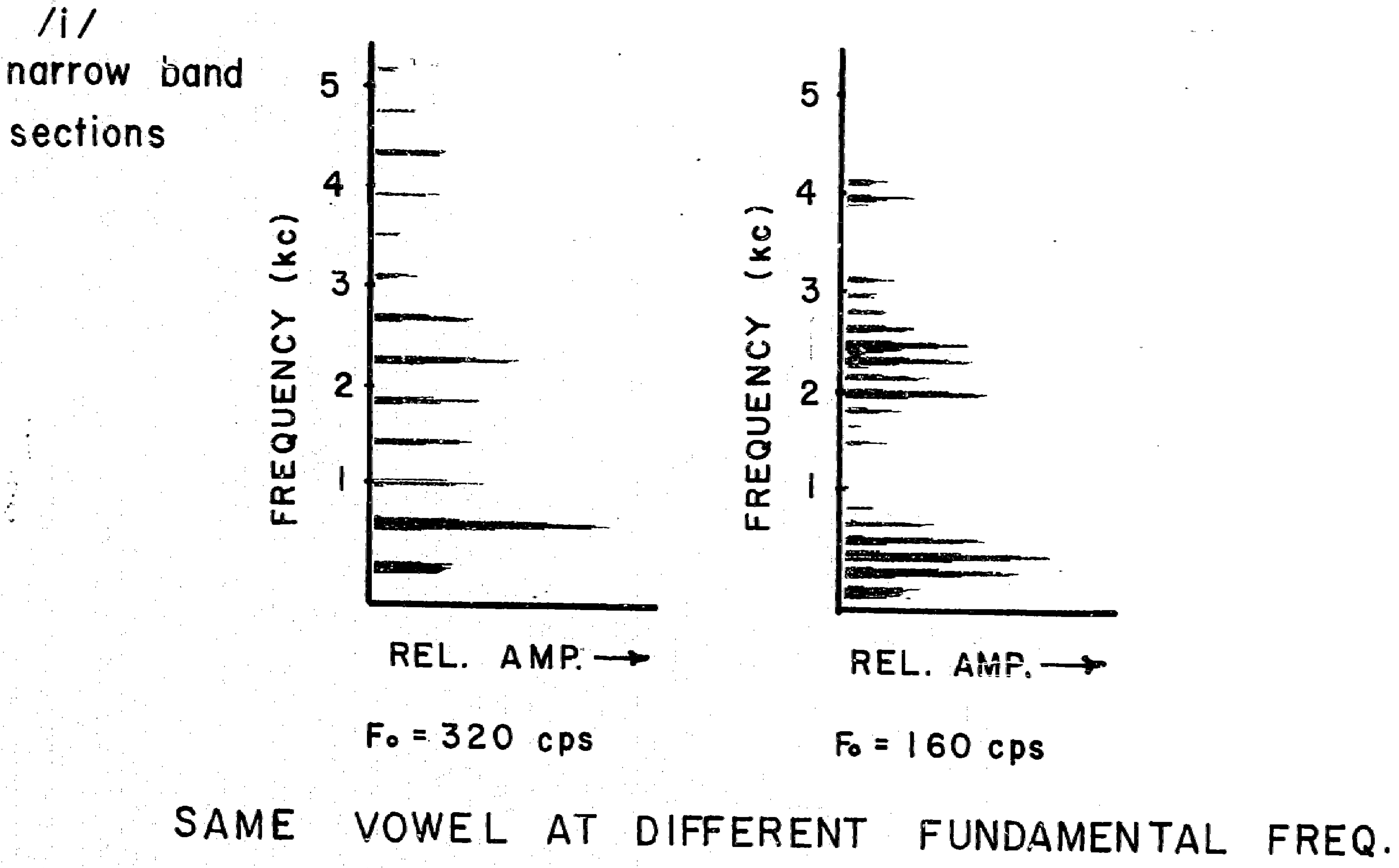


Figure 2. The local energy maxima in the vowel spectra occur in the vicinity of the formant frequencies. A speaker controls the position of the formant frequenices by varying the configuration of the supralaryngeal respiratory system.

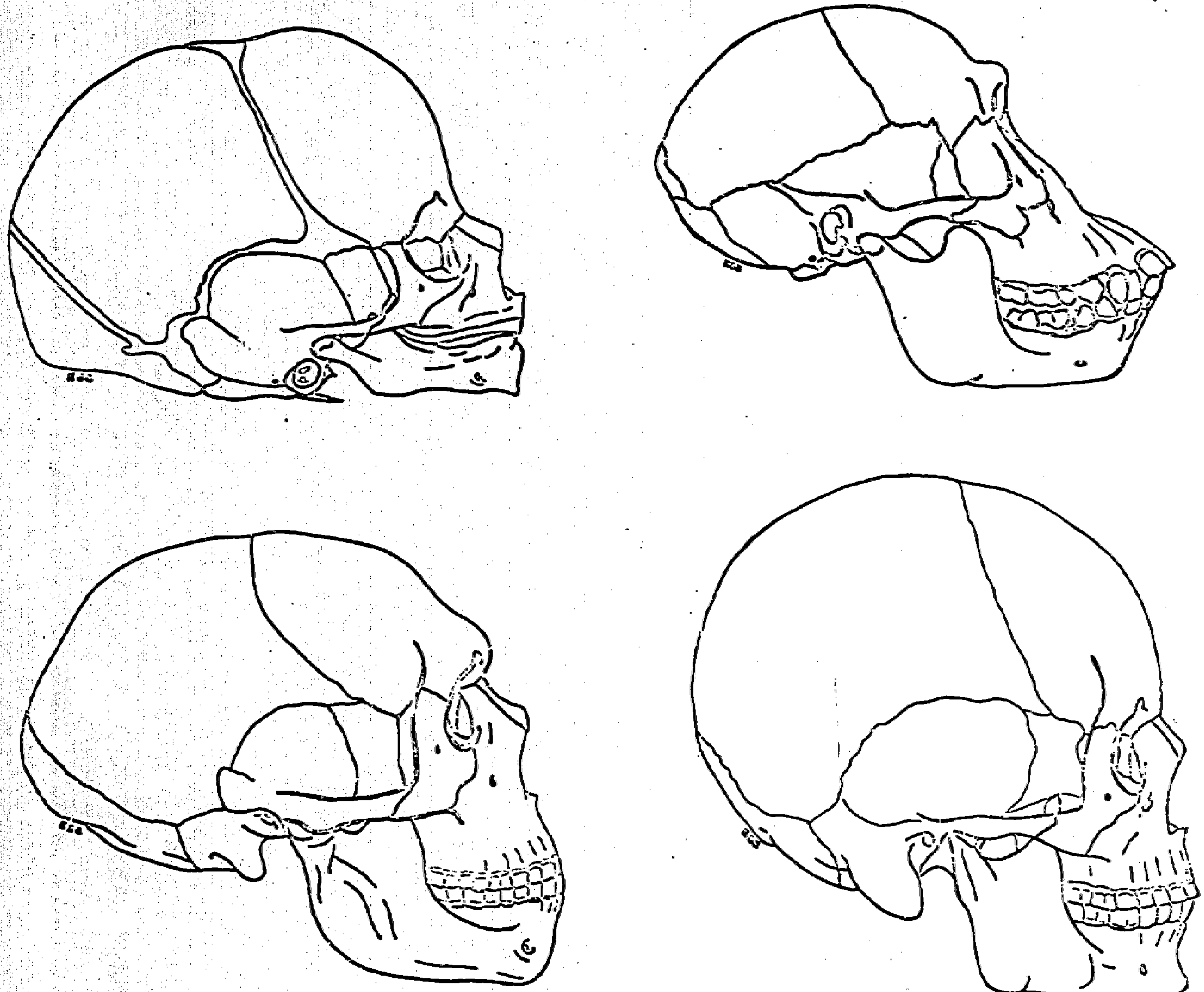


Figure 3: Skulls of human newborn, adult chimpanzee, the La Chapelle-aux-Saints fossil Neanderthal man, and adult man.

mastoid processes (very small in Neanderthal); (C) they lack a chin (occasionally present in the newborn); (D) the body of the mandible is much longer than the ramus (about 60 to 100 percent longer); (E) the posterior border of the mandibular ramus is markedly slanted away from the vertical plane; (F) they have a more horizontal inclination of the mandibular foramen leading to the mandibular canal; (G) the pterygoid process of the sphenoid bone is relatively short and its lateral lamina is more inclined away from the vertical plane; (H) the styloid process is more inclined away from the vertical plane; (I) the dental arch of the maxilla is U-shaped instead of V-shaped; (J) the basilar part of the occipital bone between the foramen magnum and the sphenoid bone is only slightly inclined away from the horizontal toward the vertical plane; (K) the roof of the nasopharynx is a relatively shallow elongated arch; (L) the vomer bone is relatively short in its vertical height, and its posterior border is inclined away from the vertical plane; (M) the vomer bone is relatively far removed from the junction of the sphenoid bone and the basilar part of the occipital bone; (N) the occipital condyles are relatively small and elongated.

In Figure 4 inferior views of the base of the skull of newborn man, Neanderthal man, and adult modern man are presented. Note that there is a relatively long space between the foramen magnum and the palate in newborn man and Neanderthal man. This long distance is reflected in the exposed portion of the sphenoid between the base of the occipital and the vomer. When the larynx is positioned with respect to the skull in Figure 5 the functional significance of the morphology of the base of the skull and mandible is apparent. The larynx is positioned high and forward in newborn and in Neanderthal. The long "flattened out" skull base, long mandible, horizontally inclined styloid processes, together with the angulation of the facets of the geniohyoid and the anterior belly of the digastric muscles at the symphysis of the mandible, are consistent with this high, fronted laryngeal position.

In Figure 6 the head of a young adult chimpanzee sectioned in the midsagittal plane is presented. Note the high position of the larynx. The tongue rests entirely within the oral cavity and the epiglottis can approximate with the soft palate. Note that the pharynx lies behind the oral cavity. In Figure 7 silicone rubber casts of the air passages, including the nasal cavity, of chimpanzee and newborn and adult man are shown. These casts were made by filling each side of the split air passages in sectioned heads and necks and then fusing the cases from each side of a head together. A cast of the reconstruction of the air passages of the La Chapelle-aux-Saints Neanderthal man is also shown. Note the basic similarities between the newborn human (1), the chimpanzee (2), and the Neanderthal (3) supralaryngeal air passages.

There is practically no supralaryngeal portion of the pharynx present in the direct airway out from the larynx in chimpanzee and Neanderthal and newborn man. In adult man half of the supralaryngeal vocal tract is formed by the pharyngeal cavity. This difference between chimpanzee, Neanderthal, and newborn--and adult man, is a consequence of the opening of the larynx into the pharynx, which is immediately behind the oral cavity in the chimpanzee, Neanderthal, and newborn. In adult man this opening occurs farther

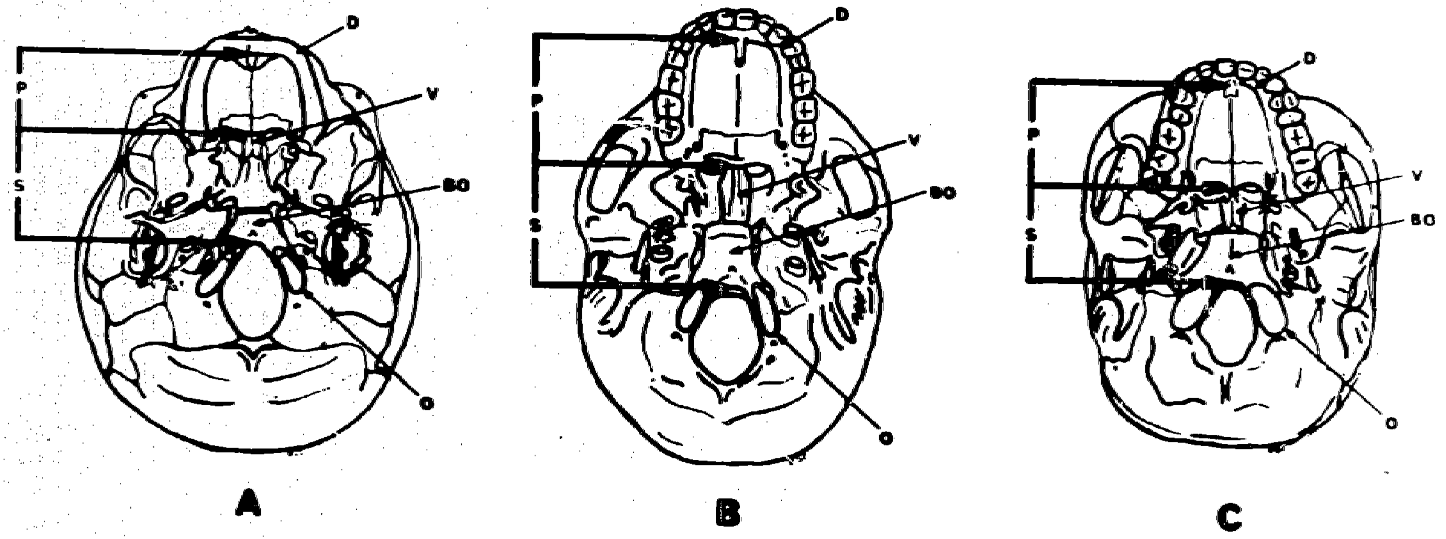


Figure 4: Inferior views of base of skull of newborn (A), Neanderthal (B), and adult man (C). D - dental arch, P - palate, S - distance between palate and foramen magnum, V - vomer bone, BO - basilar part of occipital, O - occipital condyle. (After Lieberman and Crelin, 1971.)

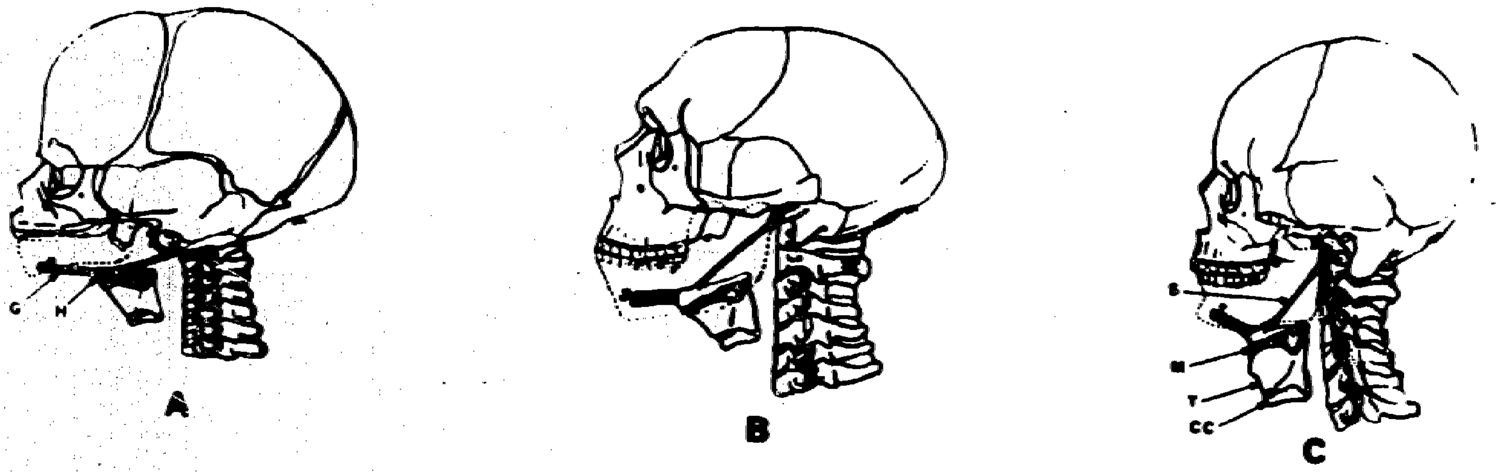


Figure 5: Skull, vertebral column, and larynx of newborn (A), reconstruction of Neanderthal (B), and adult man (C). G - geniohyoid muscle, H - hyoid bone, S - stylohyoid ligament, M - thyroid membrane, T - thyroid cartilage, CC - cricoid cartilage. (After Lieberman and Crelin, 1971.)

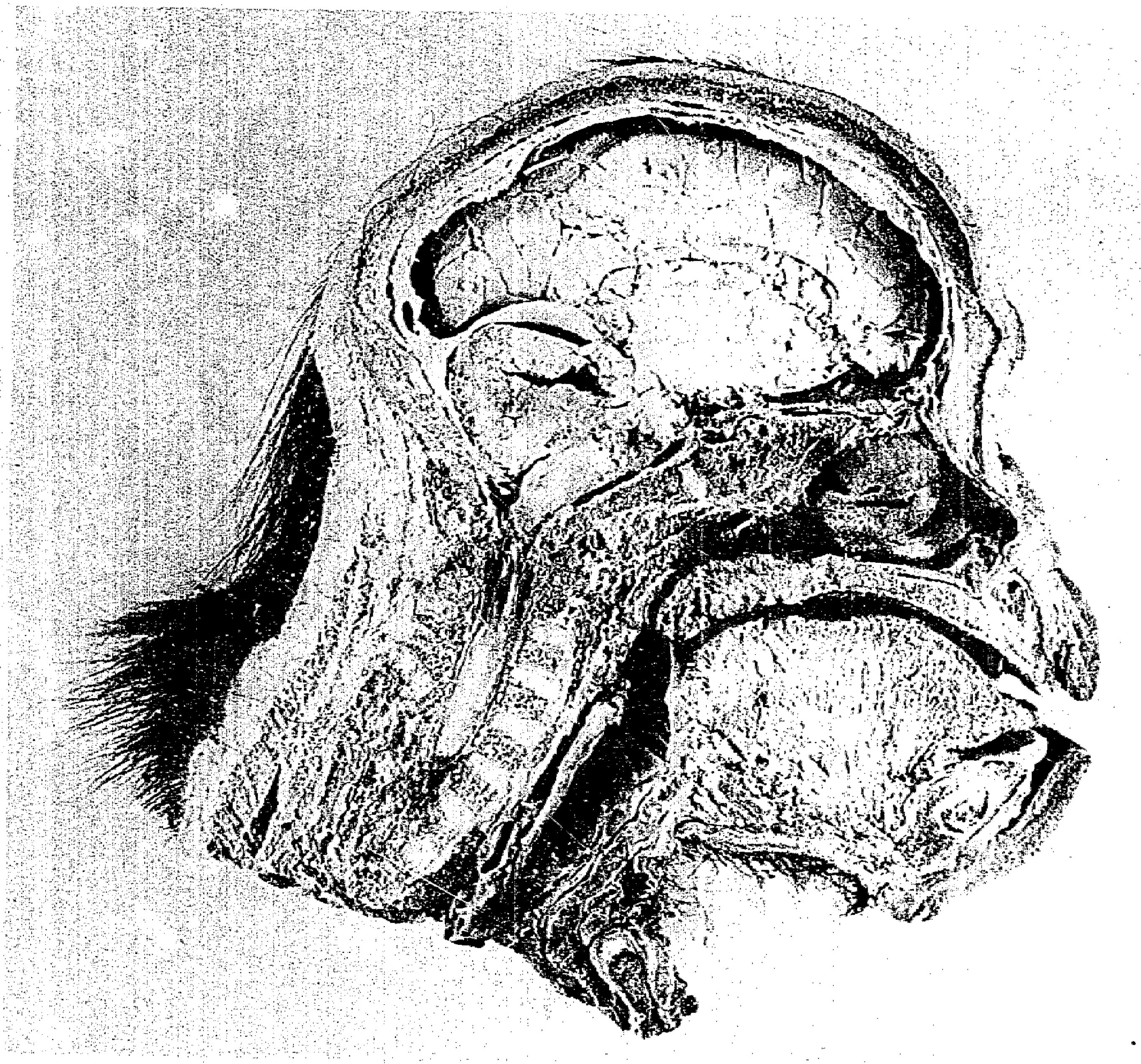
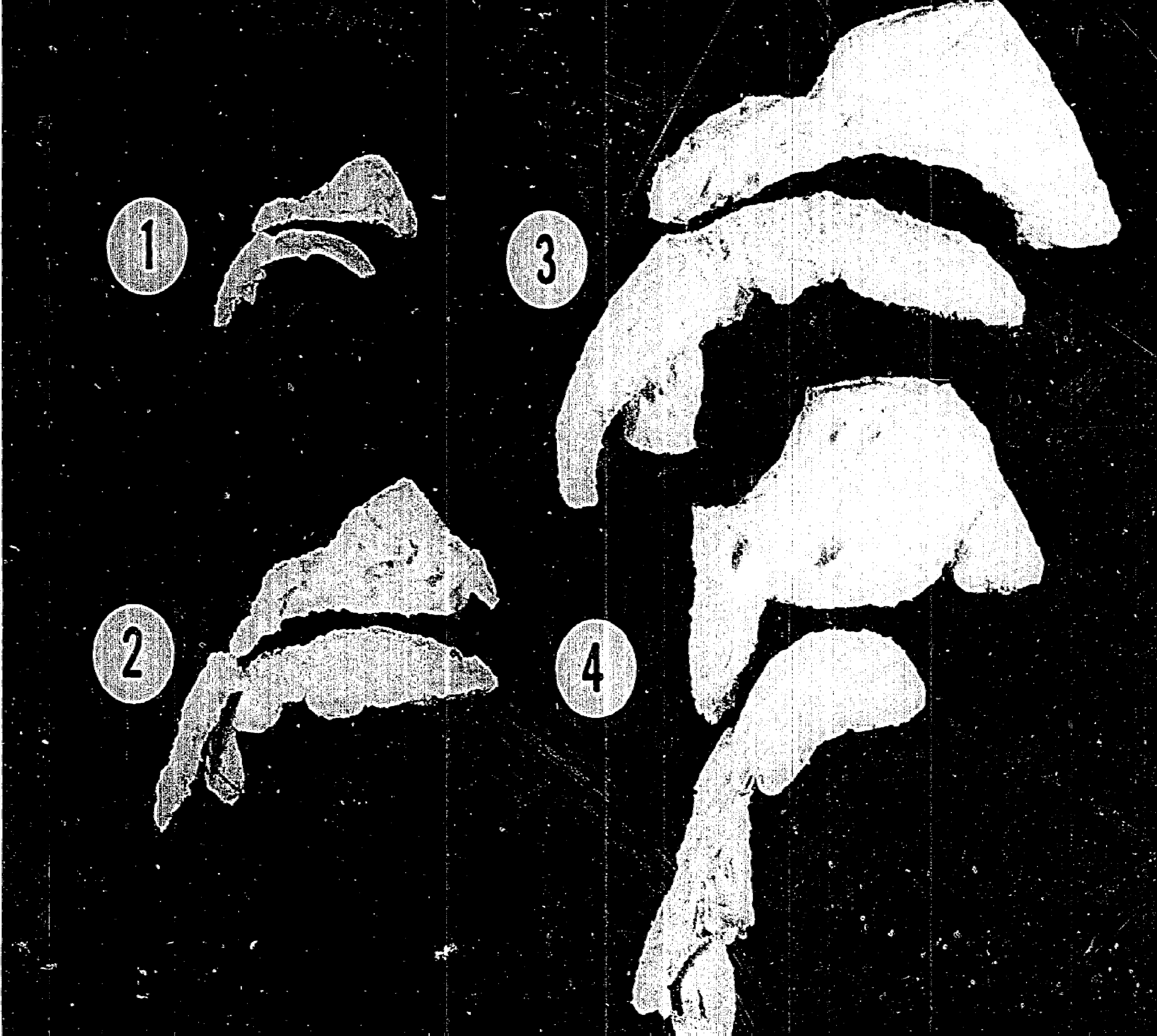


Figure 6: Left half of the head and neck of a young adult male chimpanzee sectioned in the midsaggital plane.


1
3
2
4

down in the pharynx. Note that the supralaryngeal pharynx in adult man serves both as a pathway for the ingestion of food and liquids and as an airway to the larynx. In chimpanzee, Neanderthal man, and newborn man the section of the pharynx that is behind the oral cavity is reserved for deglutition. The high epiglottis can, moreover, close the oral cavity to retain solids and liquids and allow unhampered respiration through the nose.[1]

Assessment of Limits on Phonetic Repertoire

The assessment of the limits imposed on the phonetic repertoire by the supralaryngeal vocal tract is inherently straightforward. Since the formant frequencies that determine the phonetic quality of speech sounds are specified by the shape of the supralaryngeal vocal tract (Fant, 1960), we could, if we wished, determine the range of formant frequencies by bending sheet metal into tubes that represented the limits that the vocal tract anatomy imposed on shape changes. These tubes would act as a sort of vocal tract pipe organ. We can do this in a more convenient way by modeling the range of supralaryngeal vocal tract shapes on a computer that has been programmed to act as an analog of the vocal tract. In Figure 8 three area functions that represent the most extreme deformations of a chimpanzee's supralaryngeal vocal tract in attempts to approximate the human vowels /a/, /i/, and /u/ are shown. The area function simply specifies the cross-sectional area of the supralaryngeal vocal tract along its length. It thus specifies the detailed shape of the supralaryngeal vocal tract.

In Figure 9 we have plotted the vowel-producing abilities of newborn man, chimpanzee, and Neanderthal man with respect to the vowel repertoire of adult man. Vowels can be specified by means of the first two formant frequencies The frequency of the first formant is plotted with respect to the abcissa, and that of the second formant, with respect to the ordinate. The normative data for modern man is derived from a sample of seventy-six adult men, adult women, and children (Peterson and Barney, 1952). The labeled loops enclose the data points for each vowel category. Note that none of the circles labeled "N," for Neanderthal, "1," "2," or "3" for the chimpanzee, or "X" for the newborn fall into the vowel loops for /a/, /i/, or /u/. The results of this modeling technique are consistent with acoustic measurements of living chimpanzees and newborn humans who inherently cannot produce the range of sounds necessary for human speech (Lieberman, 1968; Lieberman et al., 1969; Lieberman and Crelin, 1971, in press). The Neanderthal vocal tract also has

---

[1] The essential morphological similarities that exist between normal human newborn man and adult and juvenile "classic" Neanderthal man are discussed by Vlček (1970). Human adults never develop the specializations of adult "classic" Neanderthal man, e.g., a superorbital torus. Adult Neanderthal hominids likewise never developed the specializations of Homo sapiens, e.g., the human supralaryngeal vocal tract. Since the human supralaryngeal vocal tract is a functional anatomical specialization, it would perhaps be salutary to reserve the terms "Neanderthal" and "Neanderthaloid" to fossil forms that lack this human-like specialization. Fossil forms like Skhul V and Steinheim which appear to have had a human-like supralaryngeal vocal tract (Crelin et al., forthcoming) thus should not be classified as "Neanderthaloid" hominids.

/i/ ●—●			/a/ ■- - -■			/u/ ▲·····▲		
Formant	Freq.	Freq./1.7	Formant	Freq.	Freq./1.7	Formant	Freq.	Freq./1.7
1	610	360	1	1220	720	1	830	490
2	3400	2000	2	2550	1500	2	1800	1060
3	4420	2600	3	5070	2980	3	4080	2390

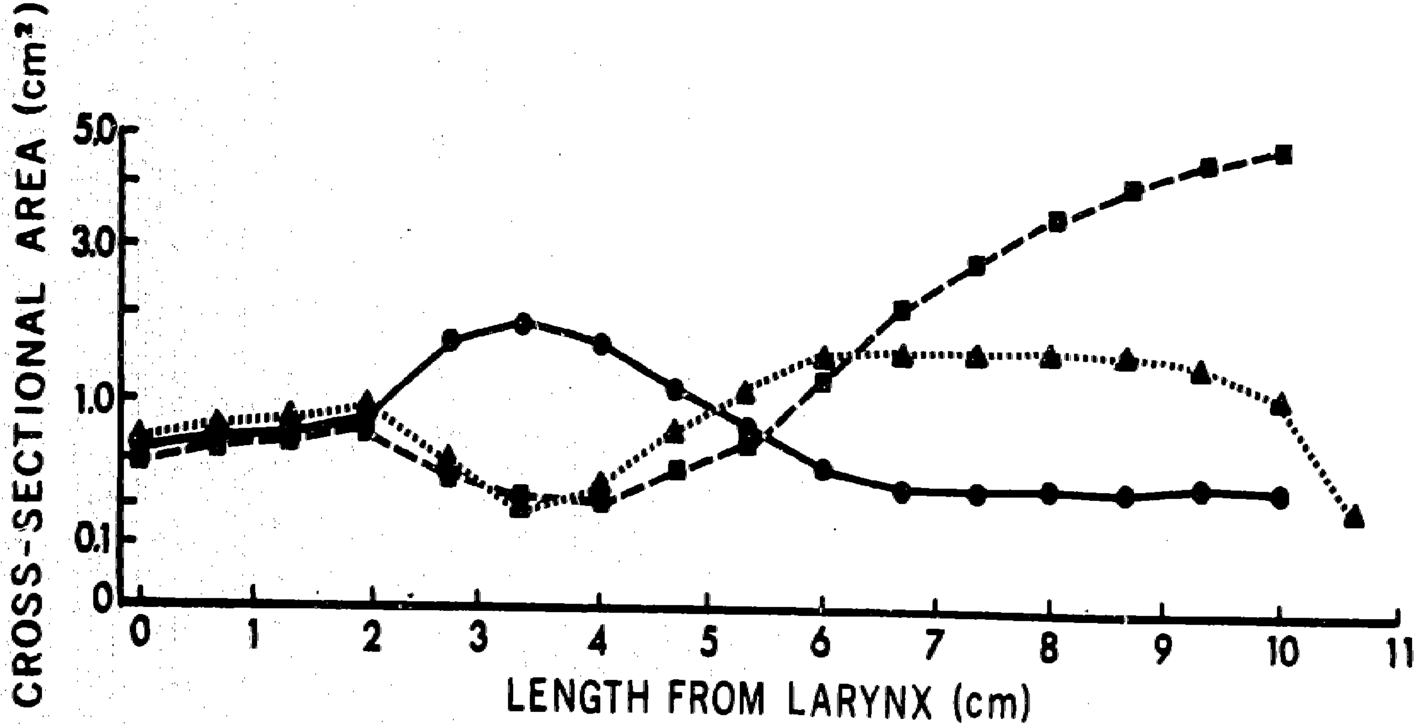


Figure 8: Chimpanzee supralaryngeal vocal tract area functions modeled on a computer. These functions were the "best" approximations that could be produced, given the anatomic limitations of the chimpanzee, to the human vowels /i/, /a/, and /u/. The formant frequencies calculated by the computer program for each vowel are tabulated and scaled to the average dimensions of the adult human vocal tract.

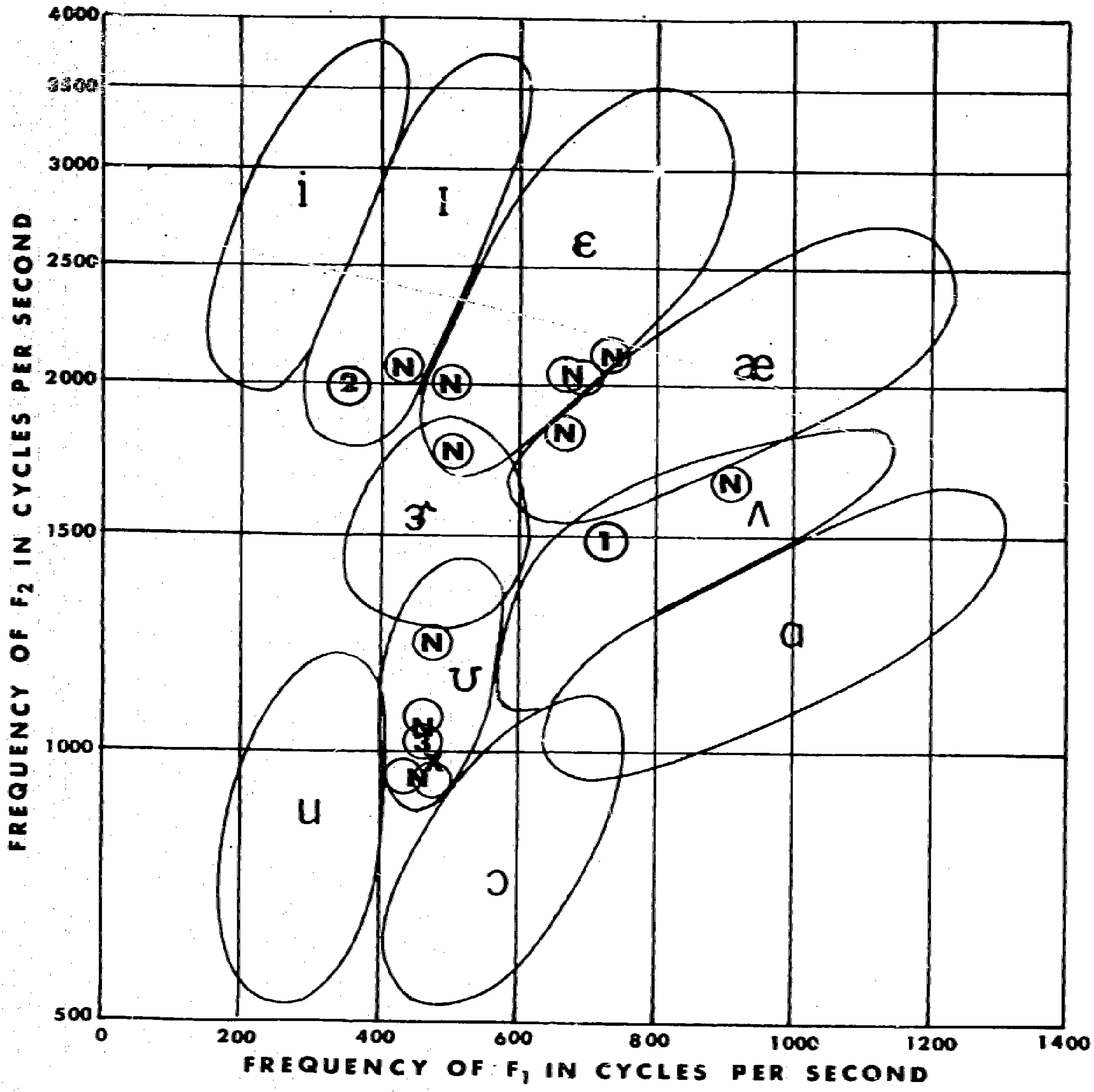


Figure 9: Vowel-producing abilities of newborn man (X), chimpanzee (1,2,3), and Neanderthal man (N) with respect to the vowel repertoire of adult man. The labeled loops enclose the data points for each vowel category.

this same deficiency. We can thus conclude that "classic" Neanderthal man inherently could not have produced the range of sounds necessary for human speech.

The uniqueness of the adult human supralaryngeal vocal tract rests in the fact that the pharynx and oral cavities are almost equal and are at right angles. No other animal has this "bent" supralaryngeal vocal tract in which the cross-sectional areas of the oral and pharyngeal cavities can be independently manipulated (Negus, 1949). In Figure 10 we have diagrammed the "bent" human vocal tract in the production of the "extreme" vowels /i/, /a/, and /u/. Note that the midpoint area functions are both extreme and abrupt. In Figure 11 the nonhuman "straight" vocal tract which is typical of Neanderthal man is diagrammed as it approximates these vowels. All area function adjustments have to take place in the oral cavity in the straight nonhuman vocal tract. Although midpoint constrictions like those needed for vowels like /a/, /i/, and /u/ can obviously be formed in the midpoint of the nonhuman vocal tract, they cannot be both extreme and abrupt. The elastic properties of the tongue prevent it from forming discontinuities that are both abrupt and extreme.

Functional Significance of Phonetic Limitations

The absence of sounds like the vowels /a/, /i/, and /u/ from the Neanderthal phonetic repertoire might at first seem interesting but trivial. After all, plenty of other sound possibilities still exist for establishing communication by means of sound. The vowels /a/, /i/, and /u/, however, have certain significant acoustic properties that relate to one of the points that I cited at the start of this talk. Human speech through a process of encoding and decoding allows communication at a rate that is about ten times faster than any other signaling system (Liberman et al., 1967; Liberman, 1970). Phonetic segments are transmitted at a rate of twenty to thirty elements per second by collapsing the acoustic cues for consonant-vowel sequences into syllable-sized units. A human listener in perceiving speech, decodes, that is, unscrambles, the acoustic cues in terms of the articulatory maneuvers and the vocal apparatus that underlie the speech signal. In order for this decoding process to function the listener needs to know the approximate size of the vocal tract that produced the speech signal (Rand, 1971). The "extreme" vowels /a/, /i/, and /u/ optimally serve this vocal tract size-calibrating function in human speech. The absence of these vowels in the phonetic repertoire of a fossil population like "classic" Neanderthal man or other examples of <u>Homo erectus</u> is, therefore, consistent with unencoded, slow, verbal communication. At worst, Neanderthal man may have completely lacked rapid, encoded human speech. At best, Neanderthal man lacked the range of phonetic possibilities of modern speech. In any event, he was not as well equipped for language as modern man.

In other ways, Neanderthal man was better equipped for life. The Neanderthal vocal tract is more efficient for breathing since airflow is not impeded by a right-angle bend (Kirchner, 1970). The Neanderthal respiratory system also cannot be blocked by food lodged in the pharynx. The Neanderthal mandible with its long body also is more efficient for chewing. Chewing efficiency in man is a function of tooth area (Manly and Braley, 1950; Manly and Shiere, 1950; Manly and Vinton, 1951). The tooth area of <u>Homo erectus</u> is substantially greater than that of <u>Homo sapiens</u>. Modern man's vocal tract

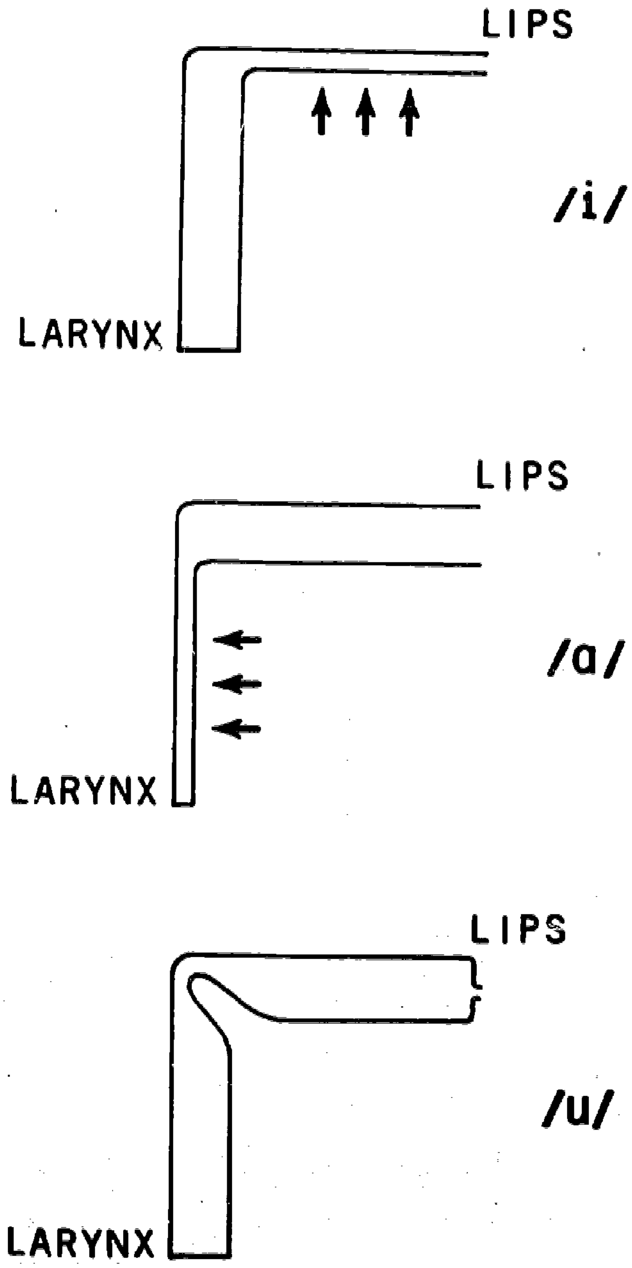


Figure 10: Schematic diagram of the "bent" human supralaryngeal vocal tract. Note that abrupt and extreme discontinuities in cross-sectional area can occur at the midpoint.

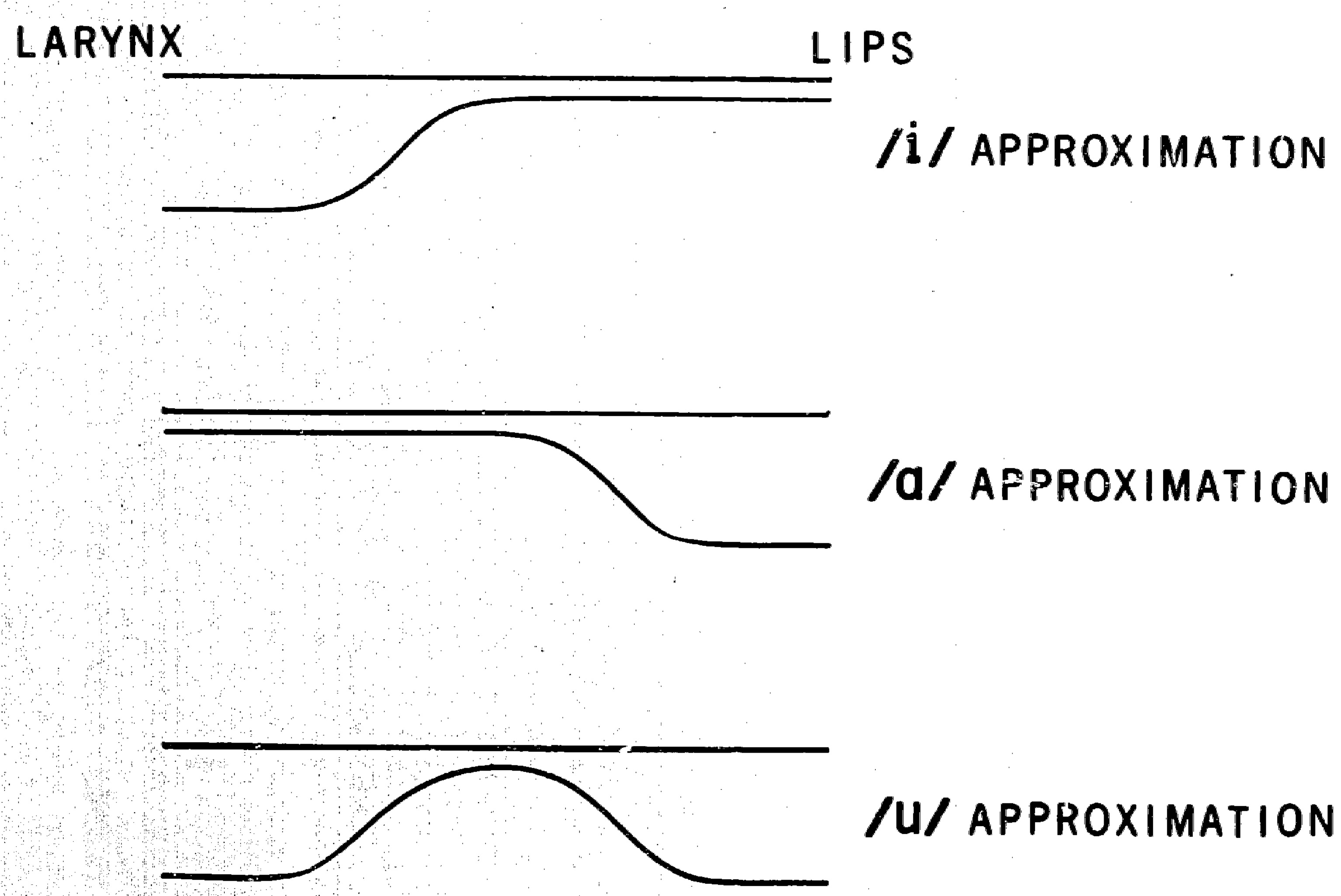


Figure 11: Schematic diagram of the straight, "single tube" nonhuman vocal tract. Note that abrupt midpoint constrictions cannot be formed.

is better suited for speech and language. He is otherwise less equipped for life. We can conclude that natural selection for enhanced speech has played as important a role in the evolution of *Homo sapiens* as upright posture and chewing played in earlier stages. Communication by means of speech may have started with the beginning of hunting, since gestural communication is limited to the line of sight. Speech communication furthermore completely frees the hands for the use of tools and weapons. Rapid, encoded communication by speech appears to be more recent. The skull of *Homo sapiens* is as functionally specialized as animals like the gorilla. The function is, however, unique with respect to all living animals insofar as it involves rapid communication by means of speech. In conclusion, we can note that the eighteenth-century philosopher La Mettrie was perhaps correct when he stated that if an ape could talk, "he would be a perfect little gentleman" (1747).

REFERENCES

Crelin, E.S., P. Lieberman, and D.H. Klatt. (forthcoming) Anatomy and related phonetic ability of the Skhul V, Steinheim, and Rhodesian fossils and the Pleisianthropus reconstruction.

Fant, G. (1960) *Acoustic Theory of Speech Production* (The Hague: Mouton).

Kirchner, J.A. (1970) *Pressman and Kelemen's Physiology of the Larynx.* Revised edition. (Rochester, Minn.: Amer. Acad. Opthal. and Otolaryn.).

La Mettrie, J.O. (1747) *De l'Homme Machine.* Critical edition, A. Vartanian, ed. (Princeton: Princeton Univ. Press, 1960).

Liberman, A.M. (1970) The grammars of speech and language. Cog. Psychol. *1*, 301-323.

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

Lieberman, P. (1968) Primate vocalizations and human linguistic ability. J. acoust. Soc. Amer. *44*, 1574-1584.

Lieberman, P. (1970) Towards a unified phonetic theory. Linguistic Inquiry *1*, 307-322.

Lieberman, P., D.H. Klatt, and W.A. Wilson. (1969) Vocal tract limitations of the vocal repertoires of Rhesus monkey and other non-human primates. Science *164*, 1185-1187.

Lieberman, P. and E.S. Crelin. (1971) On the speech of Neanderthal man. Linguistic Inquiry *2*, 203-222.

Lieberman, P., E.S. Crelin, and D.H. Klatt. (in press) Phonetic ability and related anatomy of newborn and adult human, Neanderthal man and chimpanzee. American Anthropologist *74*/3. (Also in Haskins Laboratories Status Report on Speech Research *SR-24* (1970), 57-90.)

Manly, R.S. and L.C. Braley. (1950) Masticatory performance and efficiency. J. Dent. Res. *29*, 448-462.

Manly, R.S. and F.R. Shiere. (1950) The effect of dental efficiency on mastication and food preference. Oral Surg., Oral Med., and Oral Path. *3*, 674-685.

Manly, R.S. and P. Vinton. (1951) A survey of the chewing ability of denture wearers. J. Dent. Res. *30*, 314-321.

Negus, V.E. (1949) *The Comparative Anatomy and Physiology of the Larynx.* (New York: Hafner).

Peterson, G.E. and H.L. Barney. (1952) Control methods used in a study of the vowels. J. acoust. Soc. Amer. *24*, 175-184.

Rand, T.C. (1971) Vocal tract size normalization in the perception of stop consonants. Haskins Laboratories Status Report on Speech Research, SR-25/26, 141-146.

Vlček, E. (1970) Etude comparitive onto-phylogénétique de l'enfant du Pech-de-l'Azé par rapport à d'autres enfants néanderthaliens. In L'Enfant du Pech-de-l'Azé, by D. Ferembach et al. (Paris: Masson et C^{ie}).

2/223

PUBLICATIONS AND REPORTS[*]

Publications and Manuscripts

Dichotic Backward Masking of Complex Sounds. Christopher J. Darwin. *Quarterly Journal of Experimental Psychology* (November 1971), *23*, 386-392.

Distinctive Features and Laryngeal Control. Leigh Lisker and Arthur S. Abramson. *Language* (December 1971), *47*, 767-785.

Subsequent Recognition of Items Subject to Proactive Interference in Short-Term Memory. Michael T. Turvey, D.L. Mosher, and L. Katz. *Psychonomic Science* (December 1971), *25*, 365-367.

Electrode Insertion Technique for Laryngeal Electromyography. Hajime Hirose, Thomas Gay, and Marshall Strome. *Journal of the Acoustical Society of America* (December 1971) *50*, 1449-1450.

Audible Outputs of Reading Machines for the Blind. Franklin S. Cooper, Jane H. Gaitenby, and Ignatius G. Mattingly. *Bulletin of Prosthetics Research* (Fall 1971), *BPR 10-16*.

Electromyography of the Intrinsic Laryngeal Muscles During Phonation. Thomas Gay, Hajime Hirose, Marshall Strome, and Masayuki Sawashima. To appear in *Ann. Oto. Rhinol. Laryng.*

Word-Final Stops in Thai. Arthur S. Abramson. To appear in a volume on the phonetics of Thai, to be published in Bangkok.

On Peripheral and Central Processes in Vision: Inferences from an Information-Processing Analysis of Masking with Patterned Stimuli. Michael T. Turvey.

Laryngeal Adjustments for Vowel Devoicing in Japanese: An Electromyographic Study. Hajime Hirose.

The Development of Auditory Feedback Monitoring: IV. Delayed Auditory Feedback Studies on the Vocalizations of Children Between Six and Nineteen Months. N. Fargo Belmore, Diane K. Port, Richard L. Mobley, and Violet E. Goodman.

Reports and Oral Papers

Some Possible Parallels Between Speech Perception and Verbal Learning. James J. Jenkins. Eastern Verbal Investigators League, Northampton, Mass., 10 October 1971.

An Electromyographic Investigation of the Feature of Tension in Some American English Vowels. Lawrence Raphael. Presented at 82nd meeting of the Acoustical Society of America, Denver, Colo., 19 October 1971.

---

[*] The contents of this report, SR-28, are included in this listing.

Separate Speech and Nonspeech Processing in Dichotic Listening? Ruth S. Day and James C. Bartlett. Presented at the 82nd meeting of the Acoustical Society of America, Denver, Colo., October 1971.

Listening to Speech with One, Two, or Three Ears. Michael Studdert-Kennedy. Presented at 11th Annual Meeting of the New England Psychological Association, New Haven, Conn., 12-13 November 1971.

Viewing Spoken Language as a Motor Performance by the Speaker. Katherine S. Harris. Presented at 11th Annual Meeting of the New England Psychological Association, New Haven, Conn., 12-13 November 1971.

Basic Relationships Between Talking and Reading. Michael T. Turvey. Presented at 11th Annual Meeting of the New England Psychological Association, New Haven, Conn., 12-13 November 1971.

A Brief Review of Research Programs and Their Relevance to Sensory Aids for the Blind and Deaf. Franklin S. Cooper. Presented at 11th Annual Meeting of the New England Psychological Association, New Haven, Conn., 12-13 November 1971.

The Activity of Intrinsic Laryngeal Muscles in Voicing Control: An Electromyographic Study. Hajime Hirose and Thomas Gay. Presented at the 1971 Convention of the American Speech and Hearing Association, Chicago, Ill., November 1971. (Expanded text appears in this Status Report.)

Velopharyngeal Function in Oral/Nasal Articulation and Voicing Gestures. Fredericka B. Berti and Hajime Hirose. Presented at the 1971 Convention of the American Speech and Hearing Association, Chicago, Ill., November 1971.

The Organization of Articulation Schema. Katherine S. Harris. Presented at the 1971 Convention of the American Speech and Hearing Association, Chicago, Ill., November 1971. (Revised text, "Vowel Stress and Articulatory Reorganization," appears in this Status Report.)

Some Effects of Oral Anesthesia on Speech: A Perceptual and Electromyographic Analysis. Gloria Borden. Presented at the 1971 Convention of the American Speech and Hearing Association, Chicago, Ill., November 1971.

Brain Mechanisms for Speech. Donald P. Shankweiler. Presented at the Institute on Behavior and Evolution of the American Anthropological Association, New York, N.Y., 18 November 1971.

The Evolution of Human Speech Anatomy. Philip Lieberman. Presented at the Institute on Behavior and Evolution of the American Anthropological Association, New York, N.Y., 18 November 1971.

Vocal Production Program Chairman: Lawrence Raphael. Annual Meeting of the Speech Communication Association, San Francisco, Calif., 29 December 1971.

Dichotic Fusion Along an Acoustic Continuum. James E. Cutting and Ruth S. Day. To be presented at the 83rd meeting of the Acoustical Society of America, Buffalo, N.Y., April 1972.

STATUS REPORT VOLUME NUMBERS

1965: SR-1, SR-2, SR-3, SR-4

1966: SR-5/6, SR-7/8

1967: SR-9, SR-10, SR-11, SR-12

1968: SR-13/14, SR-15/16

1969: SR-17/18, SR-19/20

1970: SR-21/22, SR-23, SR-24

1971: SR-25/26, SR-27, SR-28

AUTHOR INDEX

Haskins Laboratories Status Report on Speech Research
SR-1 (1965) - SR-28 (1971)

Abramson, Arthur S. (See Lisker, SR-1, 3.1)

__________ (See Hadding-Koch, SR-2, 2.1)

__________ and L. Lisker. Voice Onset Time in Stop Consonants: Acoustic Analysis and Synthesis. SR-3 (1965), 1.1.

__________. Book Review: Mary R. Haas, Thai-English Student's Dictionary. SR-3 (1965), 4.1.

__________, L. Lisker, and F.S. Cooper. Laryngeal Activity in Stop Consonants. SR-4 (1965), 6.1.

__________ (See Lisker, SR-5/6, 1.1)

__________ (See Lisker, SR-5/6, 4.1)

__________. On Matching the Speech Signal to Mechanisms of Production. SR-10 (1967), 2.1.

__________ (See Lisker, SR-11, 9)

__________ and L. Lisker. Discriminability Along the Voicing Continuum: Cross-Language Tests. SR-11 (1967), 17.

__________ and L. Lisker. Laryngeal Behavior, the Speech Signal, and Phonological Simplicity. SR-11 (1967), 23.

__________ and L. Lisker. Voice Timing: Cross-Language Experiments in Identification and Discrimination. SR-13/14 (1968), 49.

__________ (See Lisker, SR-15/16, 81)

__________. Tonal Experiments in Whispered Thai. SR-19/20 (1969), 37.

__________ (See Sawashima, SR-19/20, 201)

__________ (See Lisker, SR-21/22, 201)

__________. Phonetics: An Overview. SR-23 (1970), 3.

__________ (See Lisker, SR-27, 133)

__________ (See Lisker, SR-27, 175)

__________ and L. Lisker. Voice Timing in Korean Stops. SR-27 (1971), 179.

__________ Word-Final Stops in Thai. SR-28 (1971), 193.

Bailin, Howard M. (See Borst, SR-3, 2.1)

Bartlett, James C. (See Day, SR-28, 93).

Benson, Peter (See Preston, SR-13/14, 179)

Berti, Fredericka B. (See Liberman, I.Y., SR-24, 17)

__________. The Velopharyngeal Mechanism: An Electromyographic Study. A Preliminary Report. SR-25/26 (1971), 117.

__________ and H. Hirose. Velopharyngeal Function in Oral/Nasal Articulation and Voicing Gestures. SR-28 (1971), 143.

Borst, John M. and H.M. Bailin. Input Equipment for Speech Research Facility. SR-3 (1965), 2.1.

Chase, R.A. (See MacNeilage, SR-5/6, 5.1)

Cooper, Franklin S. (See Liberman, A.M., SR-1, 2.1)

__________ (See Music, SR-1, 5.1)

__________ (See Liberman, A.M., SR-3, 3.1)

__________ (See Liberman, A.M., SR-4, 1.1)

__________. Computers in Speech Research--Part II. SR-4 (1965), 4.1.

__________ (See Abramson, SR-4, 6.1)

__________. Describing the Speech Process in Motor Command Terms. SR-5/6 (1966), 2.1.

__________ (See Lisker, SR-5/6, 4.1)

__________ (See Liberman, A.M., SR-9, 1.1)

__________ (See Liberman, A.M., SR-10, 1.1)

__________, J.H. Gaitenby, and I.G. Mattingly. High-Performance Reading Machines for the Blind. SR-15/16 (1968), 111.

__________ and I.G. Mattingly. Computer-Controlled PCM System for Investigation of Dichotic Speech Perception. SR-17/18 (1969), 17.

__________, J.H. Gaitenby, I.G. Mattingly, and N. Umeda. Reading Aids for the Blind: A Special Case of Machine-to-Man Communication. SR-17/18 (1969), 117.

__________ (See Liberman, A.M., SR-19/20, 9)

__________ (See Studdert-Kennedy, SR-19/20, 139)

__________ (See Sawashima, SR-19/20, 201)

__________ (See Lisker, SR-21/22, 201)

__________, T.C. Rand, R.S. Music, and I.G. Mattingly. A Voice for the Laboratory Computer. SR-24 (1970), 91.

__________, J.H. Gaitenby, and I.G. Mattingly. Audible Outputs of Reading Machines for the Blind. SR-24 (1970), 97.

__________. How is Language Conveyed by Speech? SR-27 (1971), 3.

__________, J.H. Gaitenby, and I.G. Mattingly. Audible Outputs of Reading Machines for the Blind. SR-28 (1971), 201.

Copeland, Paul M. (See Day, SR-27, 193)

Crelin, Edmund S. (See Lieberman, SR-21/22, 157)

__________ (See Lieberman, SR-24, 57)

Crowder, Robert G. (See Darwin, SR-25/26, 13)

Cullen, John K., Jr. (See Yeni-Komshian, SR-11, 47)

__________ (See Preston, SR-11, 55)

Cutting, James E. (See Day, SR-24, 35)

__________ (See Day, SR-27, 193)

__________ and R.S. Day. Dichotic Fusion Along an Acoustic Continuum. SR-28 (1971), 103.

Darwin, Christopher J. Stimuli for Dichotic Listening Experiments. SR-11 (1967), 61.

__________. Ear Differences in the Recall of Fricatives and Vowels. SR-21/22 (1970), 39.

__________, M.T. Turvey, and R.G. Crowder. An Auditory Analogue of the Sperling Partial Report Procedure: Evidence for Brief Auditory Storage. SR-25/26 (1971), 13.

__________. Stimulus Versus Response Competition and Ear Asymmetry in Dichotic Listening. SR-25/26 (1971), 29.

__________. Dichotic Backward Masking of Complex Sounds. SR-27 (1971), 199.

Day, Ruth S. Temporal Order Judgments in Speech: Are Individuals Language-Bound or Stimulus-Bound? SR-21/22 (1970), 71.

__________ and J.E. Cutting. Perceptual Competition Between Speech and Nonspeech. SR-24 (1970), 35.

__________. Temporal Order Perception of a Reversible Phoneme Cluster. SR-24 (1970), 47.

__________ (See Wood, SR-25/26, 41)

__________ and C.C. Wood. Interactions Between Linguistic and Nonlinguistic Processing. SR-27 (1971), 185.

__________, J.E. Cutting, and P.M. Copeland. Perception of Linguistic and Nonlinguistic Dimensions of Dichotic Stimuli. SR-27 (1971), 193.

__________ and J.C. Bartlett. Separate Speech and Nonspeech Processing in Dichotic Listening? SR-28 (1971), 93.

__________ (See Cutting, SR-28, 103)

DeClerk, Joseph L. (See MacNeilage, SR-12, 9)

Dooley, Patricia M. (See Shelton, SR-12, 79)

Elbert, L. (See Music, SR-1, 5.1)

Epstein, Robert. A Transistorized Formant-Type Synthesizer. SR-1 (1965), 7.1.

Gaitenby, Jane H. The Elastic Word. SR-2 (1965), 3.1.

__________. Problems in Machine Conversion of Print to "Speech." SR-5/6 (1966), 7.1.

__________. Rules for Word Stress Analysis for Conversion of Print to Synthetic Speech. SR-12 (1967), 131.

__________ (See Cooper, SR-15/16, 111)

__________ (See Cooper, SR-17/18, 117)

__________ (See Cooper, SR-24, 97)

__________ (See Cooper, SR-28, 201)

Garcia, Erica. The Identification and Discrimination of Synthetic Nasals. SR-7/8 (1966), 3.1.

__________. Labelling of Synthetic Nasals (II). SR-9 (1967), 4.1.

__________. Identification and Discrimination of Synthetic Nasals. SR-10 (1967), 8.1.

__________. Discrimination of Three-Formant Nasal-Vowel Syllables (III). SR-12 (1967), 143.

__________. The Perception of Lateral Versus Nasal Syllabic Resonants. SR-13/14 (1968), 117.

__________. The Perception of Liquids, I: /r/ vs. /w/. SR-17/18 (1969), 41.

Gay, Thomas J. A Perceptual Study of English Diphthongs. (Ph.D. dissertation) Supplement to the Status Report, June 1967.

__________ (See Harris, SR-13/14, 137)

__________. Effect of Speaking Rate on Diphthong Formant Movements. SR-15/16 (1968), 97.

__________. Current Status of Electromyography in Speech Research. SR-15/16 (1968), 125.

__________ (See Liberman, SR-17/18, 83)

__________ (See Harris, SR-19/20, 127)

__________ (See Sawashima, SR-19/20, 211)

__________. Effects of Filtering and Vowel Environment on Consonant Perception. SR-21/22 (1970), 133.

__________, H. Hirose, M. Strome, and M. Sawashima. Electromyography of the Intrinsic Laryngeal Muscles During Phonation. SR-25/26 (1971), 97.

__________ (See Hirose, SR-28, 115)

Goff, William R. (See Wood, SR-25/26, 41)

Hadding-Koch, Kerstin and M. Studdert-Kennedy. A Study of Semantic and Psychophysical Test Responses to Controlled Variations in Fundamental Frequency. SR-1 (1965), 4.1.

__________ and A.S. Abramson. Duration Versus Spectrum in Swedish Vowels: Some Perceptual Experiments. SR-2 (1965), 2.1.

__________ and M. Studdert-Kennedy. Further Experimental Studies of Fundamental Frequency Contours. SR-4 (1965), 5.1.

Hadding, Kerstin. (See Studdert-Kennedy, SR-27, 153)

Haggard, Mark P. A Simple Program for Synthesizing Speech by Rules. SR-9 (1967), 5.1.

__________. Figural After-Effect Phenomena for Speech and Non-Speech Stimuli. SR-9 (1967), 6.1.

__________ and I.G. Mattingly. A Simple Program for Synthesizing British English. SR-12 (1967), 117.

__________. Perception of Semi-Vowels and Laterals. SR-17/18 (1969), 143.

Halwes, Terry G. Some Attempts to Make Vowels Behave Like Speech. SR-11 (1967), 57.

__________ (See Mattingly, SR-15/16, 67)

__________. Effects of Dichotic Fusion on the Perception of Speech. (Ph.D. dissertation) Supplement to the Status Report, September 1969.

__________ (See Mattingly, SR-21/22, 99)

Harris, Katherine S. (See Liberman, A.M., SR-1, 2.1)

__________, G.F. Lysaught, and M.H. Schvey. Some Aspects of the Production of Oral and Nasal Labial Stops. SR-2 (1965), 1.1.

__________ (See Liberman, A.M., SR-3, 3.1)

__________ (See Huntington, SR-3, 5.1)

__________ (See Liberman, A.M., SR-4, 1.1)

__________ (See Shankweiler, SR-4, 2.1)

__________, D.A. Huntington, and G.N. Sholes. Coarticulation of Some Disyllabic Utterances Measured by Electromyographic Techniques. SR-5/6 (1966), 3.1.

__________ (See Huntington, SR-7/8, 2.1)

__________ (See Shankweiler, SR-9, 2.1)

__________ (See Huntington, SR-10, 6.1)

__________ (See Shelton, SR-12, 79)

__________, T.J. Gay, G.N. Sholes, and P. Lieberman. Some Stress Effects on Electromyographic Measures of Consonant Articulations. SR-13/14 (1968), 137.

__________ (See Lieberman, SR-17/18, 23)

__________ (See Lieberman, SR-17/18, 83)

__________, T.J. Gay, P. Lieberman, and G.N. Sholes. The Function of Muscles in Control of Stress and Intonation. SR-19/20 (1969), 127.

__________ (See Studdert-Kennedy, SR-19/20, 139)

__________ (See Sawashima, SR-19/20, 211)

__________. Physiological Aspects of Articulatory Behavior. SR-23 (1970), 49.

__________ (See Liberman, I.Y., SR-24, 17)

__________. Action of the Extrinsic Musculature in the Control of Tongue Position: Preliminary Report. SR-25/26 (1971), 87.

__________. Vowel Stress and Articulatory Reorganization. SR-28 (1971), 167.

Hirose, Hajime. Electromyography of the Articulatory Muscles: Current Instrumentation and Technique. SR-25/26 (1971), 73.

__________ (See Gay, SR-25/26, 97)

__________. An Electromyographic Study of Laryngeal Adjustments During Speech Articulation: A Preliminary Report. SR-25/26 (1971), 107.

__________ and T.J. Gay. The Activity of the Intrinsic Laryngeal Muscles in Voicing Control: An Electromyographic Study. SR-28 (1971), 115.

__________ (See Berti, SR-28, 143)

__________. Laryngeal Adjustments for Vowel Devoicing in Japanese: An Electromyographic Study. SR-28 (1971), 157.

Howie, John M. Experiments on the Perception of Mandarin Vowels and Tones, I. SR-7/8 (1966), 7.1.

__________. Experiments on the Perception of Mandarin Vowels and Tones, II. SR-13/14 (1968), 189.

Huntington, Dorothy A. and K.S. Harris. Some Electromyographic Observations of the Speech of Persons with Impaired Hearing. SR-3 (1965), 5.1.

__________ (See Harris, SR-5/6, 3.1)

__________, K.S. Harris, and G.N. Sholes. An Electromyographic Study of Articulatory Dynamics in Hearing-Impaired and Normal Speakers. SR-7/8 (1966), 2.1.

__________, K.S. Harris, D.P. Shankweiler, and G.N. Sholes. Some Observations on Monosyllable Production by Deaf and Dysarthric Speakers. SR-10 (1967), 6.1.

Kirstein, Emily F. Perception of Second-Formant Transitions in Non-Speech Patterns. SR-7/8 (1966), 9.1.

__________ and D.P. Shankweiler. Selective Listening for Dichotically Presented Stop Consonants and Vowels. SR-17/18 (1969), 133.

__________. Selective Listening for Temporally Staggered Dichotic CV Syllables. SR-21/22 (1970), 63.

__________. Perception of Dichotically Presented Steady-State Vowels as a Function of Interaural Delay. SR-24 (1970), 31.

__________. Temporal Factors in Perception of Dichotically Presented Stop Consonants and Vowels. (Ph.D. dissertation) SR-24 (1970), Part II, 1.

Klatt, Dennis H. (See Lieberman, SR-15/16, 157)

__________ (See Lieberman, SR-24, 57)

Knudson, Ronald (See Lieberman, SR-13/14, 153)

Liberman, Alvin M., F.S. Cooper, K.S. Harris, P.F. MacNeilage, and M. Studdert-Kennedy. Some Observations on a Model for Speech Perception. SR-1 (1965), 2.1.

__________, M. Studdert-Kennedy, K.S. Harris, and F.S. Cooper. A Reply to "Identification and Discrimination Functions for a Visual Continuum and Their Relation to the Motor Theory of Speech Perception," by Cross, Lane, and Sheppard. SR-3 (1965), 3.1.

__________, F.S. Cooper, M. Studdert-Kennedy, K.S. Harris, and D.P. Shankweiler. Some Observations on the Efficiency of Speech Sounds. SR-4 (1965), 1.1.

__________, F.S. Cooper, D.P. Shankweiler, and M. Studdert-Kennedy. Perception of the Speech Code. SR-9 (1967), 1.1.

__________, F.S. Cooper, and M. Studdert-Kennedy. Why are Spectrograms Hard to Read? SR-10 (1967) 1.1.

__________ (See Stevens, SR-13/14, 47)

__________ (See Mattingly, SR-15/16, 9)

__________. Some Characteristics of Perception in the Speech Mode. SR-15/16 (1968), 39.

__________ (See Mattingly, SR-15/16, 67)

__________ and F.S. Cooper. In Search of the Acoustic Cues. SR-19/20 (1969), 9.

__________. The Grammars of Speech and Language. SR-19/20 (1969), 79.

__________ (See Studdert-Kennedy, SR-19/20, 139)

__________ (See Mattingly, SR-21/22, 99)

__________, I.G. Mattingly, and M.T. Turvey. Language Codes and Memory Codes. SR-27 (1971), 59.

233

Liberman, Isabelle Y., D.P. Shankweiler, C. Orlando, K.S. Harris, and F.B. Berti. Letter Confusions and Reversals of Sequence in the Beginning Reader: Implications for Orton's Theory of Developmental Dyslexia. SR-24 (1970), 17. (See also Erratum, SR-27, 211)

__________. Basic Research in Speech and Lateralization of Language: Some Implications for Reading Disability. SR-25/26 (1971), 51.

__________ (See Shankweiler, SR-27, 35)

Lieberman, Philip. Primate Vocalizations and Human Linguistic Ability. SR-13/14 (1968), 65.

__________ (See Harris, SR-13/14, 137)

__________, R. Knudson, and J. Mead. Determination of the Rate of Change of Fundamental Frequency with Respect to Subglottal Air Pressure During Sustained Phonation. SR-13/14 (1968), 153.

__________. On the Physical Bases of Phonologic Features. SR-15/16 (1968), 137.

__________, D.H. Klatt, and W.A. Wilson. Vocal Tract Limitations on the Vowel Repertoires of Rhesus Monkey and Other Nonhuman Primates. SR-15/16 (1968), 157.

__________, K.S. Harris, P. Wolff, and L.R. Russell. Newborn Infant Cry and Nonhuman Primate Vocalizations. SR-17/18 (1969), 23.

__________, M. Sawashima, K.S. Harris, and T.J. Gay. An Articulatory Implementation of the Breath-Group and Prominence: Crico-Thyroid Muscular Activity in Intonation. SR-17/18 (1969), 83.

__________. Towards a Unified Phonetic Theory. SR-17/18 (1969), 157.

__________ (See Harris, SR-19/20, 127)

__________ and E.S. Crelin. On the Speech of Neanderthal Man. SR-21/22 (1970), 157.

__________. A Study of Prosodic Features. SR-23 (1970), 179.

__________, E.S. Crelin, and D.H. Klatt. Phonetic Ability and Related Anatomy of the Newborn and Adult Human, Neanderthal Man, and the Chimpanzee. SR-24 (1970), 57.

__________. On the Evolution of Human Language. SR-27 (1971), 113.

__________. The Evolution of Human Speech Anatomy. SR-28 (1971), 205.

Lindblom, Björn E.F. and M. Studdert-Kennedy. On the Role of Formant Transitions in Vowel Recognition. SR-9 (1967), 7.1.

Lisker, Leigh, and A.S. Abramson. Voice Onset Time in the Production and Perception of English Stops. SR-1 (1965), 3.1.

__________ (See Abramson, SR-3, 1.1)

__________. Supraglottal Air Pressure in the Production of English Stops. SR-4 (1965), 3.1.

__________ (See Abramson, SR-4, 6.1)

__________ and A.S. Abramson. Some Effects of Context on Voice Onset Time in English Stops. SR-5/6 (1966), 1.1.

__________, A.S. Abramson, F.S. Cooper, and M.H. Schvey. Transillumination of the Larynx in Running Speech. SR-5/6 (1966), 4.1.

__________. Measuring Stop Closure Duration from Intraoral Pressure Records. SR-7/8 (1966), 5.1.

__________. Pattern to "Holes" in the Pattern? SR-10 (1967), 5.1.

__________ and A.S. Abramson. The Voicing Dimension: Some Experiments in Comparative Phonetics. SR-11 (1967), 9.

__________ (See Abramson, SR-11, 17)

__________ (See Abramson, SR-11, 23)

__________ (See Abramson, SR-13/14, 49)

__________ and A.S. Abramson. Distinctive Features and Laryngeal Control. SR-15/16 (1968), 81.

__________. Stop Duration and Voicing in English. SR-19/20 (1969), 27.

__________ (See Sawashima, SR-19/20, 201)

__________, M. Sawashima, A.S. Abramson, and F.S. Cooper. Cinegraphic Observations of the Larynx During Voiced and Voiceless Stops. SR-21/22 (1970), 201.

__________. On Time and Timing in Speech. SR-23 (1970), 151.

__________. On Learning a New Contrast. SR-24 (1970), 1.

__________ and A.S. Abramson. Distinctive Features and Laryngeal Control. SR-27 (1971), 133.

__________ and A.S. Abramson. Glottal Modes in Consonant Distinctions. SR-27 (1971), 175.

__________ (See Abramson, SR-27, 179)

Lorge, Beatrice. A Study of the Relationship Between Production and Perception of Initial and Intervocalic /t/ and /d/ in Individual English-Speaking Adults. SR-9 (1967), 3.1.

Lysaught, Gloria F. (See Harris, SR-2, 1.1)

MacNeilage, Peter F. (See Liberman, A.M. SR-1, 2.1)

__________ (See Rootes, SR-1, 8.1)

__________, T.P. Rootes, and R.A. Chase. Speech Production and Perception in a Patient with Severe Impairment of Somesthetic Perception and Motor Control. SR-5/6 (1966), 5.1.

__________ (See Winterkorn, SR-11, 41)

__________ and Joseph L. DeClerk. On the Motor Control of Coarticulation in CVC Monosyllables. SR-12 (1967), 9.

Mattingly, Ignatius G. Synthesis by Rule of General American English. SR-7/8 (1966), 4.1.

__________. Speech Synthesis by Rule as a Research Technique. SR-10 (1967), 3.1.

__________. Experimental Methods for Speech Synthesis by Rule. SR-12 (1967), 103.

__________ (See Haggard, SR-12, 117)

__________. Synthesis by Rule of General American English. (Ph.D. dissertation) Supplement to the Status Report, April 1968.

__________. Some Rules for Synthesis of General American English. SR-13/14 (1968), 37.

__________ and A.M. Liberman. The Speech Code and the Physiology of Language. SR-15/16 (1968), 9.

__________, A.M. Liberman, A.K. Syrdal, and T. Halwes. Discrimination of F2 Transitions in Speech Context and in Isolation. SR-15/16 (1968), 67.

__________ (See Cooper, SR-15/16, 111)

__________. Programming for the Glace-Holmes Synthesizer. SR-17/18 (1969), 9.

__________ (See Cooper, SR-17/18, 17)

__________ (See Cooper, SR-17/18, 117)

__________. Synthesis by Rule as a Tool for Phonological Research. SR-19/20 (1969), 61.

__________, A.M. Liberman, A.K. Syrdal, and T. Halwes. Discrimination in Speech and Nonspeech Modes. SR-21/22 (1970), 99.

__________. Speech Synthesis for Phonetic and Phonological Models. SR-23 (1970), 117.

__________ (See Cooper, SR-24, 91)

__________ (See Cooper, SR-24, 97)

__________. Reading, the Linguistic Process, and Linguistic Awareness. SR-27 (1971), 23.

__________ (See Liberman, A.M., SR-27, 59)

__________. Speech Cues and Sign Stimuli. SR-27 (1971), 89.

__________ (See Cooper, SR-28, 201)

Mead, Jere (See Liberman, SR-13/14, 153)

Music, Richard S., L. Elbert, N. Reilly, and F.S. Cooper. Automating the Collection and Reduction of Electromyographic Data on Speech Production, I and II. SR-1 (1965), 5.1.

__________ (See Cooper, SR-24, 91)

Ohman, Sven E.G. (See Stevens, SR-13/14, 47)

Orlando, Charles (See Liberman, I.Y., SR-24, 17)

Phillips, Juliet (See Preston, SR-13/14, 185)

Pisoni, David B. (See Studdert-Kennedy, SR-25/26, 1)

__________. On the Nature of Categorical Perception of Speech Sounds. (Ph.D. dissertation) Abstract: SR-27 (1971), 209. Text: Supplement to Status Report, November 1971.

Port, Diane K. (See Preston, SR-13/14, 181)

__________ (See Preston, SR-19/20, 189)

__________. The EMG Data System. SR-25/26 (1971), 67.

Porter, Robert J. The Effect of Temporal Overlap on the Perception of Dichotically and Monotically Presented CV Syllables. SR-25/26 (1971), 35.

Preston, Malcolm S. A Study of Voicing in the Stops Found in the Pre-Linguistic Babblings of Infants. SR-7/8 (1966), 6.1.

__________, G. Yeni-Komshian, and R.E. Stark. A Study of Voicing in Initial Stops Found in the Pre-Linguistic Vocalizations of Infants from Different Language Environments. SR-10 (1967), 7.1.

__________ (See Winterkorn, SR-11, 41)

__________ (See Yeni-Komshian, SR-11, 47)

__________ and G. Yeni-Komshian. Studies on the Development of Stop-Consonants in Children. SR-11 (1967), 49.

__________ and J.K. Cullen. Computer Synthesis of Infant Vocalizations. SR-11 (1967), 55.

__________, G. Yeni-Komshian, and P. Benson. A Dichotic ABX Procedure for Use in Laterality Studies. SR-13/14 (1968), 179.

__________, G. Yeni-Komshian, R.E. Stark, and D.K. Port. Developmental Studies of Voicing in Stops. SR-13/14 (1968), 181.

__________, J. Phillips, and G. Yeni-Komshian. Procedures for Examining the Perceptual Capabilities for Processing Speech Sounds in Children Under Two Years of Age. SR-13/14 (1968), 185.

__________ and D.K. Port. Further Results on the Development of Voicing in Stop-Consonants in Young Children. SR-19/20 (1969), 189.

Rand, Timothy C. (See Cooper, SR-24, 91)

__________. Vocal Tract Size Normalization in the Perception of Stop Consonants. SR-25/26 (1971), 141.

Raphael, Lawrence J. An Electromyographic Investigation of the Tense-Lax Feature in Some English Vowels. SR-25/26 (1971), 131.

__________. Vowel Duration as a Cue to the Perceptual Separation of Cognate Sounds in American English. (Ph.D. dissertation) Abstract: SR-25/26 (1971), 147. Text: Supplement to the Status Report, August 1971.

__________. An Electromyographic Investigation of the Feature of Tension in Some American English Vowels. SR-28 (1971), 179.

Reilly, Norman. (See Music, SR-1, 5.1)

Rootes, Thomas P. and P.F. MacNeilage. Coordinated Studies of a Syndrome of Impairment in Somesthetic Perception and Motor Function. SR-1 (1965), 8.1.

__________ (See MacNeilage, SR-5/6, 5.1)

Russell, Lorraine H. (See Lieberman, SR-17/18, 23)

Sawashima, Masayuki (See Lieberman, SR-17/18, 83)

__________, A.S. Abramson, F.S. Cooper, and L. Lisker. Observing Laryngeal Adjustments During Running Speech by Use of a Fiberoptics System. SR-19/20 (1969), 201.

__________, T.J. Gay, and K.S. Harris. Laryngeal Muscle Activity During Vocal Pitch and Intensity Changes. SR-19/20 (1969), 211.

__________. Glottal Adjustments for English Obstruents. SR-21/22 (1970), 187.

__________ (See Lisker, SR-21/22, 201)

__________. Laryngeal Research in Experimental Phonetics. SR-23 (1970), 69.

__________ (See Gay, SR-25/26, 97)

Schvey, Malcolm H. (See Harris, SR-2, 1.1)

__________ (See Lisker, SR-5/6, 4.1)

Schulman, Susan. (See Studdert-Kennedy, SR-21/22, 89)

Shankweiler, Donald P. Laterality Effects in Perception of Speech and Other Sounds. SR-3 (1965), 6.1.

__________ (See Liberman, A.M., SR-4, 1.1)

__________ and K.S. Harris. An Experimental Approach to the Problem of Articulation in Aphasia. SR-4 (1965), 2.1.

__________ and M. Studdert-Kennedy. Identification of Consonants and Vowels Presented to Left and Right Ears. SR-5/6 (1966), 6.1.

__________. Hemispheric Specialization and the Perception of Speech. SR-7/8 (1966), 1.1.

__________ (See Liberman, A.M., SR-9, 1.1)

__________, K.S. Harris, and M.L. Taylor. Electromyographic Studies of Articulation in Aphasia. SR-9 (1967), 2.1.

__________ and M. Studdert-Kennedy. An Analysis of Perceptual Confusions in Identification of Dichotically Presented CVC Syllables. SR-10 (1967), 4.1.

__________ (See Huntington, SR-10, 6.1)

__________. An Analysis of Laterality Effects in Speech Perception. SR-13/14 (1968), 9.

__________ (See Kirstein, SR-17/18, 133)

__________ (See Studdert-Kennedy, SR-21/22, 1)

__________ (See Studdert-Kennedy, SR-21/22, [illegible]

__________ (See Liberman, I.Y., SR-24, 17)

__________ (See Studdert-Kennedy, SR-25/26, 1)

__________ and I.Y. Liberman. Misreading: A Search for Causes. SR-27 (1971), 35.

Shelton, Ralph L., K.S. Harris, G.N. Sholes, and P.M. Dooley. Study of Non-Speech Voluntary Palate Movements by Scaling and Electromyographic Techniques. SR-12 (1967), 79.

Sholes, George N. Automating the Collection and Reduction of Electromyographic Data on Speech Production, III: Measuring and Averaging the Data. SR-1 (1965), 6.1.

__________ (See Harris, SR-5/6, 3.1)

__________ (See Huntington, SR-7/8, 2.1)

__________ (See Huntington, SR-10, 6.1)

__________ (See Shelton, SR-12, 79)

__________ (See Harris, SR-13/14, 137)

__________ (See Harris, SR-19/20, 127)

Singer, Andrew J. The Design and Implication of a Narrative Data Reduction System. Supplement to the Status Report, October 1966.

Stark, Rachel E. (See Preston, SR-10, 7.1)

__________ (See Preston, SR-13/14, 181)

Stevens, Kenneth N., A.M. Liberman, M. Studdert-Kennedy, and S.E.G. Ohman. Cross-Language Study of Vowel Perception (A). SR-13/14 (1968), 47.

Strome, Marshall (See Gay, SR-25/26, 97)

Studdert-Kennedy, Michael (See Liberman, A.M., SR-1, 2.1)

__________ (See Hadding-Koch, SR-1, 4.1)

__________ (See Liberman, A.M., SR-3, 3.1)

__________ (See Liberman, A.M., SR-4, 1.1)

__ ________ (See Hadding-Koch, SR-4, 5.1)

__________ (See Shankweiler, SR-5/6, 6.1)

__________ (See Liberman, A.M., SR-9, 1.1)

__________ (See Liberman, A.M., SR-10, 1.1)

__________ (See Shankweiler, SR-10, 4.1)

__________ (See Stevens, SR-13/14, 47)

__________, A.M. Liberman, K.S. Harris, and F.S. Cooper. A Motor Theory of Speech Perception: A Reply to Lane's Critical Review. SR-19/20 (1969), 139.

__________ and D.P. Shankweiler. Hemispheric Specialization for Speech Perception. SR-21/22 (1970), 1.

__________, D.P. Shankweiler, and S. Schulman. Opposed Effects of a Delayed Channel on Perception of Dichotically and Monotically Presented CV Syllables. SR-21/22 (1970), 89.

__________. The Perception of Speech. SR-23 (1970), 15.

__________, D.P. Shankweiler, and D.B. Pisoni. Auditory and Phonetic Processes in Speech Perception: Evidence from a Dichotic Study. SR-25/26 (1971), 1.

__________ and K. Hadding. Auditory and Linguistic Processes in the Perception of Intonation Contours. SR-27 (1971), 153.

Syrdal, Ann K. (See Mattingly, SR-15/16, 67)

__________ (See Mattingly, SR-21/22, 99)

Taylor, Martha L. (See Shankweiler, SR-9, 2.1)

Turvey, Michael T. (See Darwin, SR-25/26, 13)

__________ (See Liberman, A.M., SR-27, 59)

__________. On Peripheral and Central Processes in Vision: Inferences from an Information-Processing Analysis of Masking with Patterned Stimuli. SR-28 (1971), 1.

Umeda, Noriko (See Cooper, SR-17/18, 117)

Vinegrad, Michael D. The Scaling of Speech Sounds. SR-7/8 (1966), 8.1.

__________. The Scaling of Synthetic Speech. SR-11 (1967), 37.

__________. A Direct Magnitude Scaling Method to Investigate Categorical Versus Continuous Modes of Speech Perception. SR-21/22 (1970), 147.

Wilson, William A. (See Lieberman, SR-15/16, 157)

Winterkorn, J.M.S., P.F. MacNeilage, and M.S. Preston. Perception of Voiced and Voiceless Stop Consonants in Three-Year-Old Children. SR-11 (1967), 41.

Wolff, Peter. (See Lieberman, SR-17/18, 23)

Wood, Charles C., W.R. Goff, and R.S. Day. Auditory Evoked Potentials During Speech Perception. SR-25/26 (1971), 41.

__________ (See Day, SR-27, 185)

Yeni-Komshian, Grace (See Preston, SR-10, 7.1)

__________, M.S. Preston, and J.K. Cullen. A Brief Report on a Study of Imitation of Synthetic Speech Stimuli by Children. SR-11 (1967), 47.

__________ (See Preston, SR-11, 49)

__________ (See Preston, SR-13/14, 179)

__________ (See Preston, SR-13/14, 181)

__________ (See Preston, SR-13/14, 185)

TITLE INDEX

Haskins Laboratories Status Report on Speech Research

SR-1 (1965) - SR-28 (1971)

Action of the Extrinsic Musculature in the Control of Tongue Position: Preliminary Report. K.S. Harris. SR-25/26 (1971), 87.

The Activity of the Intrinsic Laryngeal Muscles in Voicing Control: An Electromyographic Study. H. Hirose and T.J. Gay. SR-28 (1971), 115.

An Analysis of Laterality Effects in Speech Perception. D.P. Shankweiler. SR-13/14 (1968), 9.

An Analysis of Perceptual Confusions in Identification of Dichotically Presented CVC Syllables. D.P. Shankweiler. SR-10 (1967), 4.1.

An Articulatory Implementation of the *Breath-Group* and *Prominence*: Crico-Thyroid Muscular Activity in Intonation. P. Lieberman, M. Sawashima, K.S. Harris, and T.J. Gay. SR-17/18 (1968), 83.

Some Aspects of the Production of Oral and Nasal Labial Stops. K.S. Harris, G.F. Lysaught, and M.H. Schvey. SR-2 (1965), 1.1.

Some Attempts to Make Vowels Behave Like Speech. T.G. Halwes. SR-11 (1967), 57.

Audible Outputs of Reading Machines for the Blind. F.S. Cooper, J.H. Gaitenby, and I.G. Mattingly. SR-24 (1970), 97.

Audible Outputs of Reading Machines for the Blind. F.S. Cooper, J.H. Gaitenby, and I.G. Mattingly. SR-28 (1971), 201.

An Auditory Analogue of the Sperling Partial Report Procedure: Evidence for Brief Auditory Storage. C.J. Darwin, M.T. Turvey, and R.G. Crowder. SR-25/26 (1971), 13.

Auditory Evoked Potentials During Speech Perception. C.C. Wood, W.R. Goff, and R.S. Day. SR-25/26 (1971), 41.

Auditory and Linguistic Processes in the Perception of Intonation Contours. M. Studdert-Kennedy and K. Hadding. SR-27 (1971), 153.

Auditory and Phonetic Processes in Speech Perception: Evidence from a Dichotic Study. M. Studdert-Kennedy, D.P. Shankweiler, and D.B. Pisoni. SR-25/26 (1971), 1.

Automating the Collection and Reduction of Electromyographic Data on Speech Production, I and II. R.S. Music, L. Elbert, N. Reilly, and F.S. Cooper. SR-1 (1965), 5.1.

Automating the Collection and Reduction of Electromyographic Data on Speech Production, III: Measuring and Averaging the Data. G.N. Sholes. SR-1 (1965), 6.1.

Basic Research in Speech and Lateralization of Language: Some Implications for Reading Disability. I.Y. Liberman. SR-25/26 (1971), 51.

Book Review: Mary R. Haas, Thai-English Student's Dictionary. A.S. Abramson. SR-3 (1965), 4.1.

A Brief Report on a Study of Imitation of Synthetic Speech Stimuli by Children. G. Yeni-Komshian, M.S. Preston, and J.K. Cullen. SR-11 (1967), 47.

Some Characteristics of Perception in the Speech Mode. A.M. Liberman. SR-15/16 (1968), 39.

Cinegraphic Observations of the Larynx During Voiced and Voiceless Stops. L. Lisker, M. Sawashima, A.S. Abramson, and F.S. Cooper. SR-21/22 (1970), 201.

Coarticulation of Some Disyllabic Utterances Measured by Electromyographic Techniques. K.S. Harris, D.A. Huntington, and G.N. Sholes. SR-5/6 (1966), 3.1.

Computer-Controlled PCM System for Investigation of Dichotic Speech Perception. F.S. Cooper and I.G. Mattingly. SR-17/18 (1969), 17.

Computers in Speech Research -- Part II. F.S. Cooper. SR-4 (1965), 4.1.

Computer Synthesis of Infant Vocalizations. M.S. Preston and J.K. Cullen. SR-11 (1967), 55.

Coordinated Studies of a Syndrome of Impairment of Somesthetic Perception and Motor Function. T.P. Rootes and P.F. MacNeilage. SR-1 (1965), 8.1.

Cross-Language Study of Vowel Perception (A). K.N. Stevens, A.M. Liberman, M. Studdert-Kennedy, and S.E.G. Ohman. SR-13/14 (1968), 47.

Current Status of Electromyography in Speech Research. T.J. Gay. SR-15/16 (1968), 125.

Describing the Speech Process in Motor Command Terms. F.S. Cooper. SR-5/6 (1966), 2.1.

The Design and Implementation of a Narrative Data Reduction System. A.J. Singer. Supplement to the Status Report, October 1966.

Determination of the Rate of Change of Fundamental Frequency with Respect to Subglottal Air Pressure During Sustained Phonation. P. Lieberman. SR-13/14 (1968), 153.

Developmental Studies of Voicing in Stops. M.S. Preston, G. Yeni-Komshian, R.E. Stark, and D.K. Port. SR-13/14 (1968), 181.

A Dichotic ABX Procedure for Use in Laterality Studies. M.S. Preston, G. Yeni-Komshian, and P. Benson. SR-13/14 (1968), 179.

Dichotic Backward Masking of Complex Sounds. C.J. Darwin. SR-27 (1971), 199.

Dichotic Fusion Along an Acoustic Continuum. J.E. Cutting and R.S. Day. SR-28 (1971), 103.

A Direct Magnitude Scaling Method to Investigate Categorical Versus Continuous Modes of Speech Perception. M.D. Vinegrad. SR-21/22 (1970), 147.

Discriminability Along the Voicing Continuum: Cross-Language Tests. A.S. Abramson and L. Lisker. SR-11 (1967), 17.

Discrimination of F2 Transitions in Speech Context and in Isolation. I.G. Mattingly, A.M. Liberman, A.K. Syrdal, and T. Halwes. SR-15/16 (1968), 67.

Discrimination in Speech and Nonspeech Modes. I.G. Mattingly, A.M. Liberman, A.K. Syrdal, and T. Halwes. SR-21/22 (1970), 99.

Discrimination of Three-Formant Nasal-Vowel Syllables (III). E. Garcia. SR-12 (1967), 143.

Distinctive Features and Laryngeal Control. L. Lisker and A.S. Abramson. SR-15/16 (1968), 81.

Distinctive Features and Laryngeal Control. L. Lisker and A.S. Abramson. SR-27 (1971), 133.

Duration Versus Spectrum in Swedish Vowels: Some Perceptual Experiments. K. Hadding-Koch and A.S. Abramson. SR-2 (1965), 2.1.

Ear Differences in the Recall of Fricatives and Vowels. C.J. Darwin. SR-21/22 (1970), 39.

Some Effects of Context on Voice Onset Time in English Stops. L. Lisker and A.S. Abramson. SR-5/6 (1966), 1.1.

Effects of Dichotic Fusion on the Perception of Speech. T. Halwes. (Ph.D. dissertation). Supplement to the Status Report, September 1969.

Effects of Filtering and Vowel Environment on Consonant Perception. T.J. Gay. SR-21/22 (1970), 133.

Effect of Speaking Rate on Diphthong Formant Movements. T.J. Gay. SR-15/16 (1969), 97.

The Effect of Temporal Overlap on the Perception of Dichotically and Monotically Presented CV Syllables. R.J. Porter. SR-25/26 (1971), 35.

The Elastic Word. J.H. Gaitenby. SR-2 (1965), 3.1.

An Electromyographic Investigation of the Feature of Tension in Some American English Vowels. L.J. Raphael. SR-28 (1971), 179.

An Electromyographic Investigation of the Tense-Lax Feature in Some English Vowels. L.J. Raphael. SR-25/26 (1971), 131.

Some Electromyographic Observations of the Speech of Persons with Impaired Hearing. D.A. Huntington and K.S. Harris. SR-3 (1965), 5.1.

Electromyographic Studies of Articulation in Aphasia. D. Shankweiler, K.S. Harris, and M.L. Taylor. SR-9 (1967), 2.1.

An Electromyographic Study of Articulatory Dynamics in Hearing-Impaired and Normal Speakers. D.A. Huntington, K.S. Harris, and G.N. Sholes. SR-7/8 (1966), 2.1.

An Electromyographic Study of Laryngeal Adjustments During Speech Articulation: A Preliminary Report. H. Hirose. SR-25/26 (1971), 107.

Electromyography of the Articulatory Muscles: Current Instrumentation and Technique. H. Hirose. SR-25/26 (1971), 73.

Electromyography of the Intrinsic Laryngeal Muscles During Phonation. T.J. Gay, H. Hirose, M. Strome, and M. Sawashima. SR-25/26 (1971), 97.

The EMG Data System. D.K. Port. SR-25/26 (1971), 67.

On the Evolution of Human Language. P. Lieberman. SR-27 (1971), 113.

The Evolution of Human Speech Anatomy. P. Lieberman. SR-28 (1971), 205.

The Experimental Approach to the Problem of Articulation in Aphasia. D.P. Shankweiler and K.S. Harris. SR-4 (1965), 2.1.

Experimental Methods for Speech Synthesis by Rule. I.G. Mattingly. SR-12 (1967), 103.

Experiments on the Perception of Mandarin Vowels and Tones, I. J.M. Howie. SR-7/8 (1966), 7.1.

Experiments on the Perception of Mandarin Vowels and Tones, II. J.M. Howie. SR-13/14 (1968), 189.

Figural After-Effect Phenomena for Speech and Non-Speech Stimuli. M.P. Haggard. SR-9 (1967), 6.1.

The Function of Muscles in Control of Stress and Intonation. K.S. Harris, T.J. Gay, P. Lieberman, and G.N. Sholes. SR-19/20 (1966), 127.

Further Experimental Studies of Fundamental Frequency Contours. K. Hadding-Koch and M. Studdert-Kennedy. SR-4 (1965), 5.1.

Further Results on the Development of Voicing in Stop Consonants in Young Children. M.S. Preston and D.K. Port. SR-19/20 (1969), 189.

Glottal Adjustments for English Obstruents. M. Sawashima. SR-21/22 (1970), 187.

Glottal Modes in Consonant Distinctions. L. Lisker and A.S. Abramson. SR-27 (1971), 175.

The Grammars of Speech and Language. A.M. Liberman. SR-19/20 (1969), 79.

Hemispheric Specialization and the Perception of Speech. D.P. Shankweiler. SR-7/8 (1966), 1.1.

Hemispheric Specialization for Speech Perception. M. Studdert-Kennedy and D.P. Shankweiler. SR-21/22 (1970), 1.

High-Performance Reading Machines for the Blind. F.S. Cooper, J.H. Gaitenby, and I.G. Mattingly. SR-15/16 (1968), 111.

How is Language Conveyed by Speech? F.S. Cooper. SR-27 (1971), 3.

Identification of Consonants and Vowels Presented to Left and Right Ears. D.P. Shankweiler and M. Studdert-Kennedy. SR-5/6 (1966), 6.1.

The Identification and Discrimination of Synthetic Nasals. E. Garcia. SR-7/8 (1966), 3.1.

Identification and Discrimination of Synthetic Nasals. E. Garcia. SR-10 (1967), 8.1.

Input Equipment for Speech Research Facility. J. Borst and H.M. Bailin. SR-3 (1965), 2.1.

Interactions Between Linguistic and Nonlinguistic Processing. R.S. Day and C.C. Wood. SR-27 (1971), 85.

Labelling of Synthetic Nasals (II). E. Garcia SR-9 (1967), 4.1.

Language Codes and Memory Codes. A.M. Liberman. SR-27 (1971), 59.

Laryngeal Activity in Stop Consonants. A.S. Abramson, L. Lisker, and F.S. Cooper. SR-4 (1965), 6.1.

Laryngeal Adjustments for Vowel Devoicing in Japanese: An Electromyographic Study. H. Hirose. SR-28 (1971), 157.

Laryngeal Behavior, the Speech Signal, and Phonological Simplicity. A.S. Abramson and L. Lisker. SR-11 (1967), 23.

Laryngeal Muscle Activity During Vocal Pitch and Intensity Changes. M. Sawashima, T.J. Gay, and K.S. Harris. SR-19/20 (1969), 211.

Laryngeal Research in Experimental Phonetics. M. Sawashima. SR-23 (1970), 69.

Laterality Effects in Perception of Speech and Other Sounds. D.P. Shankweiler. SR-3 (1965), 6.1.

On Learning a New Contrast. L. Lisker. SR-24 (1970), 1.

Letter Confusions and Reversals of Sequence in the Beginning Reader: Implications for Orton's Theory of Developmental Dyslexia. I.Y. Liberman, D.P. Shankweiler, C. Orlando, K.S. Harris, and F.B. Berti. SR-24 (1970), 17. (See also Erratum, SR-27, 211)

On Matching the Speech Signal to Mechanisms of Production. A.S. Abramson. SR-10 (1967), 2.1.

Measuring Stop Closure Duration from Intraoral Pressure Records. L. Lisker. SR-7/8 (1966), 5.1.

Misreading: A Search for Causes. D.P. Shankweiler and I.Y. Liberman. SR-27 (1971), 35.

On the Motor Control of Coarticulation in CVC Monosyllables. P.F. MacNeilage and J.L. DeClerk. SR-12 (1967), 9.

A Motor Theory of Speech Perception: A Reply to Lane's Critical Review. M. Studdert-Kennedy, A.M. Liberman, K.S. Harris, and F.S. Cooper. SR-19/20 (1969), 139.

On the Nature of Categorical Perception of Speech Sounds. D.B. Pisoni. (Ph.D. dissertation) Abstract: SR-27 (1971), 209. Text: Supplement to the Status Report, November 1971.

Newborn Infant Cry and Nonhuman Primate Vocalizations. P. Lieberman, K.S. Harris, P. Wolff, and L.R. Russell. SR-17/18 (1969), 23.

Some Observations on the Efficiency of Speech Sounds. A.M. Liberman, F.S. Cooper, M. Studdert-Kennedy, K.S. Harris, and D.P. Shankweiler. SR-4 (1965), 1.1.

Some Observations on Monosyllable Production by Deaf and Dysarthric Speakers. D.A. Huntington, K.S. Harris, D.P. Shankweiler, and G.N. Sholes. SR-10 (1967), 6.1.

Some Observations on a Model for Speech Perception. A.M. Liberman, F.S. Cooper, K.S. Harris, P.F. MacNeilage, and M. Studdert-Kennedy. SR-1 (1965), 2.1.

Observing Laryngeal Adjustments During Running Speech by Use of a Fiberoptics System. M. Sawashima, A.S. Abramson, F.S. Cooper, and L. Lisker. SR-19/20 (1969), 201.

Opposed Effects of a Delayed Channel on Perception of Dichotically and Monotically Presented CV Syllables. M. Studdert-Kennedy. D.P. Shankweiler, and S. Schulman. SR-21/22 (1970), 89.

Pattern to "Holes" in the Pattern? L. Lisker. SR-10 (1967), 5.1.

Perception of Dichotically Presented Steady-State Vowels as a Function of Interaural Delay. E.F. Kirstein. SR-24 (1970), 31.

The Perception of Lateral Versus Nasal Syllabic Resonants. E. Garcia. SR-13/14 (1968), 117.

Perception of Linguistic and Nonlinguistic Dimensions of Dichotic Stimuli. R.S. Day, J.E. Cutting, and P.M. Copeland. SR-27 (1971), 193.

The Perception of Liquids, I: /r/ vs. /w/. E. Garcia. SR-17/18 (1969), 41.

Perception of Second-Formant Transitions in Non-Speech Patterns. E.F. Kirstein. SR-7/8 (1966), 9.1.

Perception of Semi-Vowels and Laterals. M.P. Haggard. SR-17/18 (1969), 143.

The Perception of Speech. M. Studdert-Kennedy. SR-23 (1970), 15.

Perception of the Speech Code. A.M. Liberman, F.S. Cooper, D.P. Shankweiler, and M. Studdert-Kennedy. SR-9 (1967), 1.1.

Perception of Voiced and Voiceless Stop Consonants in Three-Year-Old Children. J.M.S. Winterkorn, P.F. MacNeilage, and M.S. Preston. SR-11 (1967), 41.

Perceptual Competition Between Speech and Nonspeech. R.S. Day and J.E. Cutting. SR-24 (1970), 35.

A Perceptual Study of English Diphthongs. T.J. Gay. (Ph.D. dissertation) Supplement to the Status Report, June 1967.

On Peripheral and Central Processes in Vision: Inferences from an Information-Processing Analysis of Masking with Patterned Stimuli. M.T. Turvey. SR-28 (1971), 1.

Phonetic Ability and Related Anatomy of the Newborn and Adult Human, Neanderthal Man, and the Chimpanzee. P. Lieberman, E.S. Crelin, and D.H. Klatt. SR-24 (1970), 57.

Phonetics: An Overview. A.S. Abramson. SR-23 (1970), 3.

On the Physical Bases of Phonetic Features. P. Lieberman. SR-15/16 (1968), 137.

Physiological Aspects of Articulatory Behavior. K.S. Harris. SR-23 (1970), 49.

Primate Vocalizations and Human Linguistic Ability. P. Lieberman. SR-13/14 (1968), 65.

Problems in Machine Conversion of Print to "Speech." J.H. Gaitenby. SR-5/6 (1966), 7.1.

Procedures for Examining the Perceptual Capabilities for Processing Speech Sounds in Children Under Two Years of Age. M.S. Preston, J. Phillips, and G. Yeni-Komshian. SR-13/14 (1968), 185.

Programming for the Glace-Holmes Synthesizer. I.G. Mattingly. SR-17/18 (1969), 9.

Reading Aids for the Blind: A Special Case of Machine-to-Man Communication. F.S. Cooper, J.H. Gaitenby, I.G. Mattingly and N. Umeda. SR-17/18 (1969), 117.

Reading, the Linguistic Process, and Linguistic Awareness. I.G. Mattingly. SR-27 (1971), 23.

A Reply to "Identification and Discrimination Functions for a Visual Continuum and Their Relation to the Motor Theory of Speech Perception," by Cross, Lane, and Sheppard. A.M. Liberman, M. Studdert-Kennedy, K.S. Harris, and F.S. Cooper. SR-3 (1965), 3.1.

On the Role of Formant Transitions in Vowel Recognition. B.E.F. Lindblom and M. Studdert-Kennedy. SR-9 (1967), 7.1.

Some Rules for Synthesis of General American English. I.G. Mattingly. SR-13/14 (1968), 37.

Rules for Word Stress Analysis for Conversion of Print to Synthetic Speech. J.H. Gaitenby. SR-12 (1967), 131.

The Scaling of Speech Sounds. M.D. Vinegrad. SR-7/8 (1966), 8.1.

The Scaling of Synthetic Speech. M.D. Vinegrad. SR-11 (1967), 37.

In Search of the Acoustic Cues. A.M. Liberman and F.S. Cooper. SR-19/20 (1969), 9.

Selective Listening for Dichotically Presented Stop Consonants and Vowels. E.F. Kirstein and D.P. Shankweiler. SR-17/18 (1969), 133.

Selective Listening for Temporally Staggered Dichotic CV Syllables. E.F. Kirstein. SR-21/22 (1970), 63.

Separate Speech and Nonspeech Processing in Dichotic Listening. R.S. Day and J.C. Bartlett. SR-28 (1971), 93.

A Simple Program for Synthesizing British English. M.P. Haggard and I.G. Mattingly. SR-12 (1967), 117.

A Simple Program for Synthesizing Speech by Rules. M.P. Haggard. SR-9 (1967), 5.1.

The Speech Code and the Physiology of Language. I.G. Mattingly and A.M. Liberman. SR-15/16 (1968), 9.

Speech Cues and Sign Stimuli. I.G. Mattingly. SR-27 (1971), 89.

On the Speech of Neanderthal Man. P. Lieberman and E.S. Crelin. SR-21/22 (1970), 157.

Speech Production and Perception in a Patient with Severe Impairment of Somesthetic Perception and Motor Control. P.F. MacNeilage, T.P. Rootes, and R.A. Chase. SR-5/6 (1966), 5.1.

Speech Synthesis for Phonetic and Phonological Models. I.G. Mattingly. SR-23 (1970), 117.

Speech Synthesis by Rule as a Research Technique. I.G. Mattingly. SR-10 (1967), 3.1.

Stimuli for Dichotic Listening Experiments. C.J. Darwin. SR-11 (1967), 61.

Stimulus Versus Response Competition and Ear Asymmetry in Dichotic Listening. C.J. Darwin. SR-25/26 (1971), 29.

Stop Duration and Voicing in English. L. Lisker. SR-19/20 (1969), 27.

Some Stress Effects on Electromyographic Measures of Consonant Articulations. K.S. Harris, T.J. Gay, G.N. Sholes, and P. Lieberman. SR-13/14 (1968), 137.

Studies on the Development of Stop Consonants in Children. M.S. Preston and G. Yeni-Komshian. SR-11 (1967), 49.

Study of Non-Speech Voluntary Palate Movements by Scaling and Electromyographic Techniques. R.L. Shelton, K.S. Harris, G.N. Sholes, and P.M. Dooley. SR-12 (1967), 79.

A Study of Prosodic Features. P. Lieberman. SR-23 (1970), 179.

A Study of the Relationship Between Production and Perception of Initial and Intervocalic /t/ and /d/ in Individual English-Speaking Adults. B. Lorge. SR-9 (1967), 3.1.

A Study of Semantic and Psychophysical Test Responses to Controlled Variations in Fundamental Frequency. K. Hadding-Koch and M. Studdert-Kennedy. SR-1 (1965), 4.1.

A Study of Voicing in Initial Stops Found in the Pre-Linguistic Vocalizations of Infants from Different Language Environments. M.S. Preston, G. Yeni-Komshian, and R.E. Stark. SR-10 (1967), 7.1.

A Study of Voicing in the Stops Found in the Pre-Linguistic Babblings of Infants. M.S. Preston. SR-7/8 (1966), 6.1.

Supraglottal Air Pressure in the Production of English Stops. L. Lisker. SR-4 (1965), 3.1.

Synthesis by Rule of General American English. I.G. Mattingly. SR-7/8 (1966), 4.1.

Synthesis by Rule of General American English. I.G. Mattingly. (Ph.D. dissertation) Supplement to the Status Report, April 1968.

Synthesis by Rule as a Tool for Phonological Research. I.G. Mattingly. SR-19/20 (1969), 61.

Temporal Factors in Perception of Dichotically Presented Stop Consonants and Vowels. E.F. Kirstein. (Ph.D. dissertation) SR-24 (1970), Part II, 1.

Temporal Order Judgments in Speech: Are Individuals Language-Bound or Stimulus-Bound? R.S. Day. SR-21/22 (1970), 71.

Temporal Order Perception of a Reversible Phoneme Cluster. R.S. Day. SR-24 (1970), 47.

On Time and Timing in Speech. L. Lisker. SR-23 (1970), 151.

Tonal Experiments in Whispered Thai. A.S. Abramson. SR-19/20 (1969), 37.

Towards a Unified Phonetic Theory. P. Lieberman. SR-17/18 (1969), 157.

Transillumination of the Larynx in Running Speech. L. Lisker, A.S. Abramson, F.S. Cooper, and M.H. Schvey. SR-5/6 (1966), 4.1.

A Transistorized Formant-Type Synthesizer. R. Epstein. SR-1 (1965), 7.1.

Velopharyngeal Function in Oral/Nasal Articulation and Voicing Gestures. F.B. Berti and H. Hirose. SR-28 (1971), 143.

The Velopharyngeal Mechanism: An Electromyographic Study. A Preliminary Report. F.B. Berti. SR-25/26 (1971), 117.

Vocal Tract Limitations on the Vowel Repertoires of Rhesus Monkey and Other Nonhuman Primates. P. Lieberman, D.H. Klatt, and W.A. Wilson. SR-15/16 (1968), 157.

Vocal Tract Size Normalization in the Perception of Stop Consonants. T.C. Rand. SR-25/26 (1971), 141.

A Voice for the Laboratory Computer. F.S. Cooper, T.C. Rand, R.S. Music, and I.G. Mattingly. SR-24 (1970), 91.

Voice Onset Time in the Production and Perception of English Stops. L. Lisker and A.S. Abramson. SR-1 (1965), 3.1.

Voice Onset Time in Stop Consonants: Acoustic Analysis and Synthesis. A.S. Abramson and L. Lisker. SR-3 (1965), 1.1.

Voice Timing: Cross-Language Experiments in Identification and Discrimination. A.S. Abramson and L. Lisker. SR-13/14 (1968), 49.

Voice Timing in Korean Stops. A.S. Abramson and L. Lisker. SR-27 (1971), 179.

The Voicing Dimension: Some Experiments in Comparative Phonetics. L. Lisker and A.S. Abramson. SR-11 (1967), 9.

Vowel Duration as a Cue to the Perceptual Separation of Cognate Sounds in American English. L.J. Raphael. (Ph.D. dissertation) Abstract: SR-25/26 (1971), 147. Text: Supplement to the Status Report, August 1971.

Vowel Stress and Articulatory Reorganization. K.S. Harris. SR-28 (1971), 167.

Word-Final Stops in Thai. A.S. Abramson. SR-28 (1971), 193.

Why are Spectrograms Hard to Read? A.M. Liberman, F.S. Cooper, and M. Studdert-Kennedy. SR-10 (1967), 1.1.

UNCLASSIFIED
Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author)
Haskins Laboratories, Inc.
270 Crown Street
New Haven, Conn. 06510

2a. REPORT SECURITY CLASSIFICATION
Unclassified

2b. GROUP
N/A

3. REPORT TITLE
Status Report on Speech Research, no. 28, October-December 1971

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)
Interim Scientific Report

5. AUTHOR(S) (First name, middle initial, last name)
Staff of Haskins Laboratories; Franklin S. Cooper, P.I.

6. REPORT DATE
January 1972

7a. TOTAL NO. OF PAGES
218

7b. NO. OF REFS
243

8a. CONTRACT OR GRANT NO.
ONR Contract N00014-67-A-0129-0001
b. NIDR: Grant DE-01774
NICHD: Grant HD-01994
c. NIH/DRFR: Grant RR-5596
VA/PSAS Contract V101(134)P-71
d. NICHD Contract NIH-71-2420

9a. ORIGINATOR'S REPORT NUMBER(S)
SR-28 (1971)

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
None

10. DISTRIBUTION STATEMENT
Distribution of this document is unlimited.*

11. SUPPLEMENTARY NOTES
N/A

12. SPONSORING MILITARY ACTIVITY
See No. 8

13. ABSTRACT

This report (for 1 October-31 December) is one of a regular series on the status and progress of studies on the nature of speech, instrumentation for its investigation, and practical applications. Manuscripts and extended reports cover the following topics:

- On Peripheral and Central Processes in Vision: Inferences from an Information-Processing Analysis of Masking with Patterned Stimuli
- Separate Speech and Nonspeech Processing in Dichotic Listening?
- Dichotic Fusion Along an Acoustic Continuum
- The Activity of the Intrinsic Laryngeal Muscles in Voicing Control: An Electromyographic Study
- Velopharyngeal Function in Oral/Nasal Articulation and Voicing Gestures
- Laryngeal Adjustments for Vowel Devoicing in Japanese: An Electromyographic Study
- Vowel Stress and Articulatory Reorganization
- An Electromyographic Investigation of the Feature of Tension in Some American English Vowels
- Word-Final Stops in Thai
- Audible Outputs of Reading Machines for the Blind
- The Evolution of Human Speech Anatomy

DD FORM 1 NOV 65 1473 (PAGE 1)
S/N 0101-807-6811

*This document contains no information not freely available to the general public. It is distributed primarily for library use.

UNCLASSIFIED
Security Classification

A-31408

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Perceptual and central visual mechanisms Visual masking Speech and nonspeech processing Dichotic fusion Electromyography Voicing control Laryngeal muscles Velopharyngeal function Vowel devoicing Stop consonants Thai Reading machines for the blind Synthetic speech						

DD FORM 1 NOV 65 1473 (BACK)
0101-807-6821

A-31409