SOME ASPECTS OF THE PRODUCTION OF ORAL AND NASAL LABIAL STOPS

Katherine S. Harris, Gloria F. Lysaught and Malcolm M. Schvey
Haskins Laboratories, New York

Reprinted from
LANGUAGE AND SPEECH
Vol. 8, Part 3, July-September 1965, pp. 135 - 147

SOME ASPECTS OF THE PRODUCTION OF ORAL AND NASAL LABIAL STOPS*

KATHERINE S. HARRIS, GLORIA F. LYSAUGHT** and MALCOLM M. SCHVEY***
Haskins Laboratories, New York

Various measures were made of the electromyographic output of the muscles of the lips during the production of /p/, /b/, /m/, /mp/ and /mb/, to see if there was a difference in the amount or patterning of activity between sounds usually described as "tense" and "lax". Overall, there is some slight average tendency for "tense" sounds to be produced more forcefully than "lax", but this tendency is present only for some subjects and when large numbers of responses are averaged. It is not large enough to serve as the basis for a phonemic distinction based on muscular effort. We conclude that the essential difference between /p/ and /b/ lies elsewhere. The present results support a conclusion drawn from other studies that the real difference is in the relative timing of events at the glottis and at the place of oral occlusion.

This paper is an examination of a familiar group of sounds, and a much-debated question in phonetics, on the basis of experimental data supplied mainly by electromyographic techniques.

The study was undertaken to answer the old question as to the nature of the disinction between the two sets of English stops /p/ /t/ /k/ and /b/ /d/ /g/. It is generally recognized that in most phonetic contexts these sets differ in voicing; however, it has often been held that the *essential* difference between these sounds is a rather less well-defined opposition between "tense" and "lax".

There is a general agreement as to the physiological underpinning for "voicing"; the vocal cords vibrate during production. A satisfactory physical referent for the tense-lax opposition is somewhat more elusive. A number of authors (e.g. Fant, 1960) give a difference in the behaviour of the buccal air as the physiological referent for a tense/lax, or fortis/lenis, distinction. Recent workers (Lisker, 1963) have discussed this difference in buccal air pressure at length, concluding that "the relations among pressure, voicing and the /ptk-bdg/ set of contrasts are not as simple as stated in the previous studies. It has been found that there is no difference in pressure which invariably marks the contrasts between /p/ and /b/."

Is there another plausible referent for "tense-lax"? A possible candidate is some difference in the muscular tension in the production of tense-lax pairs. In a stop, for example, a greater muscular effort might be assumed in the production of a voiceless than a voiced stop, which might be manifested in greater tension in executing the occlusive gesture, or in the tension during occlusion, or in the force with which the

---

* *This development was supported in part by Public Health Service Grant DE- 01774, from the National Institute of Dental Research, National Institute of Health, Education and Welfare.*

** *Now at the English Language Institute, Queens College, New York.*

*** *Also, Department of Otolaryngology, Columbia Presbyterian Hospital.*

occlusion is broken. Such increased effort would presumably result in more rapid transitions into or out of the fortis sound, and perhaps a longer hold, and these would, of course, be reflected in observable changes at the acoustic level. It is difficult to find any author in the literature who commits himself in any very firm way in describing this difference in effort. However, Jakobson and Halle (1962) are about as specific as anyone when they say, "In producing lax phonemes the vocal tract exhibits the same behaviour as in generating the cognate tense phonemes but with a significant attenuation. This attenuation manifests itself by a lower air pressure in the cavity . . . by a smaller deformation of the vocal tract from its neutral, central position, and/or by a more rapid release of the constriction." This description is intended to cover the tense-lax opposition both in vowels and in consonants. One presumes that the attenuation must have muscular origins, at least in part. This study was made to investigate the difference in muscular terms.

The way we measure muscular patterns is by studying the voltage generated when articulatory muscles contract. These voltages can be picked up by suitably placed electrodes over the muscles—a technique called surface electromyography.

EXPERIMENTAL WORK

In this study, we will consider the results of three experiments in which we study the electromyographic production of American English /p/ and /b/ to ask if there is any difference in the behaviour of the articulators which we might legitimately describe as a difference in "force of articulation" of the /p/ and /b/ gestures.

The three experiments made use of a special-purpose electromyographic installation that was designed for research in speech. The equipment has been described in a previous article (Harris, 1964).

Preceding any formal experimentation, we explored to find electrode locations which would be diagnostic of the articulation of bilabial stops. We expected the muscles which close and open the lips to be active. We found that any location on the upper lip, close to the vermilion border, showed a large peak in activity when the lips were pursed for the "closing gesture" of /p, b/, and any location below the lower lip showed a peak for the "opening gesture". These electrode locations were used in all the experiments which follow.

The first and second experiments were performed several years ago, and have been briefly reported elsewhere (Lysaught, 1961). The third experiment was performed more recently, on an improved installation. Only the data which are relevant to the tense-lax question will be reported from the third experiment.

Experiment I

The five subjects were asked to read, in list fashion, a series of utterances which contained the labial stops /p/ and /b/ in specified contexts. The relevant stops occurred in initial and final positions, paired with a long vowel, /i/, and a short

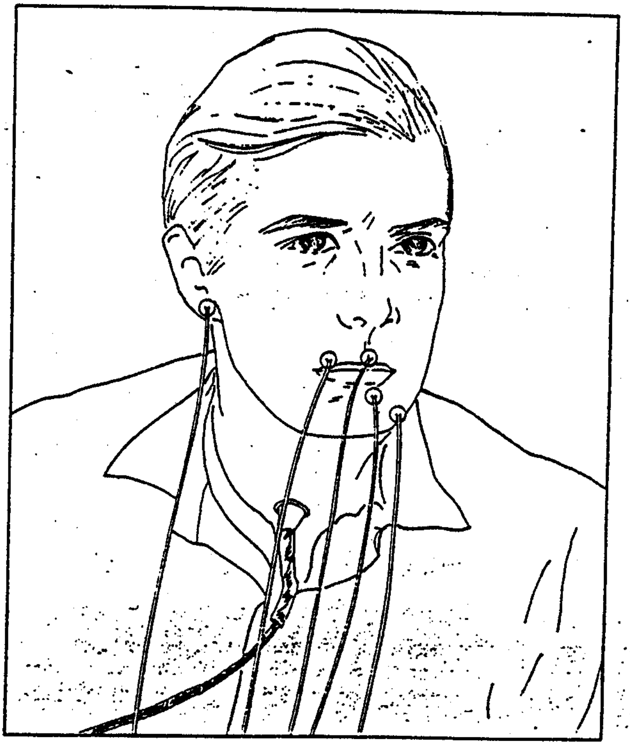

Fig. 1. Typical placement of the electrodes and the vibration transducer for subject in Experiment I.

vowel, /ɪ/. In each utterance, the labial stop was preceded or followed by an unstressed vowel, /ə/ and the stressed syllable was initiated or terminated by an unvoiced stop, /k/. The utterances to be produced were then,

əpik	əbik
əpɪk	əbɪk
kipə	kibə
kɪpə	kɪbə

These were assembled into a random order list of 400 items.

The subject produced the sounds with two pairs of suction electrodes attached to the surface of his face, in the locations described above, and a vibration transducer strapped to his throat. In Experiments I and II, the transducer was a bone-conduction receiver ; in Experiment III, it was a throat microphone of the type described by

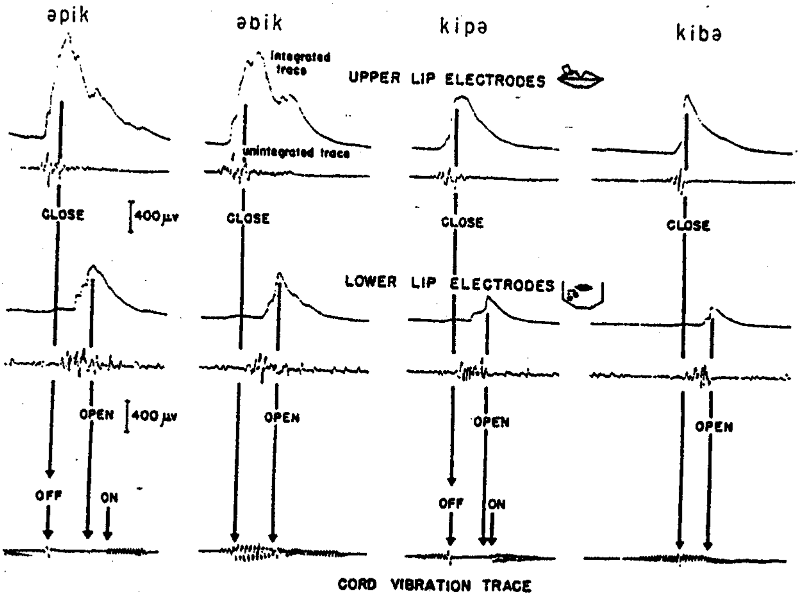

Fig. 2. Sample outputs of EMG electrodes and bone conduction receiver for four syllables.

Porter (1963). As can be seen from Fig. 1, the electrodes on the upper lip are suitably placed to make a bipolar record of the output of the orbicularis oris muscle, which runs in a band around the upper and lower lips. These fibres contract to produce a pursing of the lips as in making the occlusion for a labial stop—the " closing gesture ". The lower pair of electrodes are placed in such a way that they record the burst of activity which accompanies the parting of the lips when the labial stop occlusion is broken—the " opening gesture ".

The outputs of the throat pick-up and the electrodes were led to suitable amplifying equipment and, ultimately, to a multi-channel inkwriter.

Typical recordings from the pairs of electrodes are shown in Fig. 2. The lowest line in the figure shows the output of the bone conduction receiver. In the two utterances containing /p/, these traces show, as expected, a period of no vibration. For /b/, there is a characteristic change in the shape of the wave during the period of the occlusion. The cord vibration pattern allows identification of the time of mouth closure and re-opening, as indicated on the figure.

Two sets of myographic traces are shown. The second and fourth lines show the output without signal processing. On the first and third lines are shown the output after rectification and running integration with a 22 msec. time constant.

The integrated traces were used in all subsequent data analysis. Although the peak voltage is quite variable for repetitions of the same utterance, by the same speaker,

TABLE 1

Analysis of variance in lip opening gesture.

EXPERIMENT I

Source	*Mean Square*	*F Ratio*
/p/ *vs.* /b/	71.1111	2.0239
Position	2958.3889	84.1994*
Subjects	404.4583	11.5114*
Vowels	56.0111	1.5941
p/b x P	74.7222	2.1079
p/b x S	95.6389	2.6980
p/b x V	46.9445	1.3243
P x S	119.8472	3.3809
P x V	28.9111	<1.0
S x V	29.5806	<1.0
p/b x P x S	29.2778	<1.0
p/b x P x V	5.8667	<1.0
p/b x S x V	23.5277	<1.0
P x S x V	65.6471	1.8550
p/b x P x S x V	14.6445	<1.0

their overall shape does not fluctuate so much. Furthermore, for patterns like the ones shown, with one large peak, the height of this peak appeared to be a fairly good indicator of the area under the curve. We therefore used peak measurements as our indicators of size. (In Experiment III we show the results of a more elaborate analysis.)

The gains for each subject were adjusted to approximately the same overall height, but some differences remain. Each subject made a few errors in production, but no more than one error of any particular type occurred. We therefore eliminated one utterance at random from groups of ten in order to equalize the N in each group. Thus, the data represent 45 tokens (9 tokens × 5 subjects) of each of 8 types.

An analysis of variance was performed on the measurements of peak heights, separately for the upper and lower lip data. The results of this analysis are shown in Tables 1 and 2.

The results of the analysis for the lip opening gesture may be discussed first. Since they do not show any very interesting effects, the analysis was not repeated in later experiments. In this experiment, as in subsequent experiments, a difference in the "force of articulation" would presumably be reflected in a *significant* difference in the peak heights for /p/ and /b/ diagnostic gestures, with or without significant higher order interaction terms. The difference between /p/ and /b/ is not significant for the lip-opening gesture. There are two significant main effects—that for "initial

TABLE 2

Analysis of variance in lip closing gesture.

EXPERIMENT I

Source	*Mean Square*	*F Ratio*
/p/ *vs.* /b/	880.4694	4.6571
Position	1,318.6694	2.00
Subjects	2,250.1222	5.15
Vowels	6.6694	<1.0
p/b x P	374.1362	14.00*
p/b x S	142.7889	5.35*
p/b x V	0.2251	<1.0
P x S	731.0445	27.34*
P x V	46.2250	1.73
S x V	13.2389	<1.0
p/b x P x S	59.1500	2.2222
p/b x P x V	4.2249	<1.0
p/b x S x V	34.5445	1.2949
P x S x V	27.2944	1.0231
p/b x P x S x V	26.6778	2.4704

vs. terminal" and that for "subjects". The significant value for subjects is of no interest, since it represents only a failure to adjust amplifier gains on the equipment to the same peak height for all subjects. The significant value for position may mean either of two things—that the labial stop opening is produced more vigorously before a stressed than before an unstressed vowel, or possibly, more vigorously before vowels of the /i-ɪ/ class than before the neutral vowel. The first alternative seems more likely, though not certain.

The results of the analysis of variance for the lip closing gesture are shown in Table 2. The lack of a significant difference between subjects represents simply a more successful gain adjustment. There are *no* significant main effects. The difference between /p/ and /b/ is not simply represented in a difference in the magnitude of the closing gesture. There are three significant second-order interaction terms in the table. Those involving position, as we stated above, should probably be investigated in a more thorough manner.

However, the most interesting, perhaps, is the significant *F* ratio for Subjects x p/b[1].

---

[1] *In this, and in all subsequent analyses of variance, we have used the techniques for pooling variance and constructing significance tests outlined in a standard statistical text (Bennett and Franklin, 1954). We have not discussed the analyses in great detail, since the central arguments in this paper do not depend on them.*

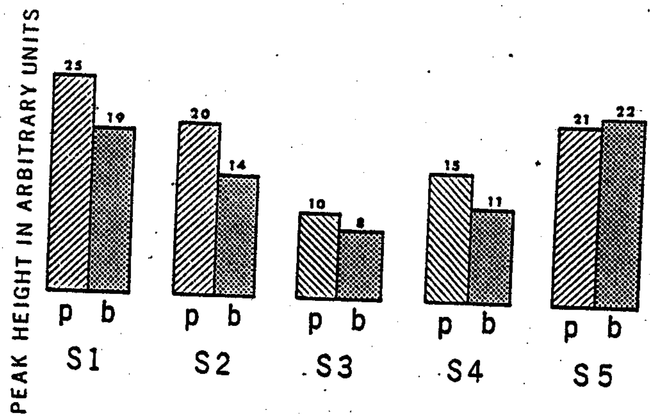

Fig. 3. Histogram of average peak response magnitudes for /p/ and /b/ closure for five subjects, Experiment I.

The mean magnitude for /p/ and /b/ separately for each subject is plotted in Fig. 3. It becomes immediately obvious why the interaction term was significant. Four of the five subjects show a small but consistent difference between /p/ and /b/ in the expected direction. The magnitudes for /p/ and /b/ are about equal, though reversed, for the fifth. Obviously, then, we have no evidence in this experiment for the kind of difference in "force" which would lead to a consistent difference of production of the two sounds.

Still another way of looking at the data is seen in the histogram of peak heights for /p/ and /b/ productions for a single subject, shown in Fig. 4.

In this plot of the data, all productions of /p/ and /b/ have been pooled in a single histogram, regardless of phonetic environment. As one can easily see, the largest size productions are /p/, the smallest /b/, but there is a great deal of overlap. If the experimenter were presented with a single production of /p/ or /b/, he could not decide with any confidence which class it should be assigned to on the basis of height.

Experiment II

There were five subjects in Experiment II, as in Experiment I. Again, subjects read lists of "words" containing bilabial stops in intervocalic position. The utterances themselves were different. They were of the form :

$[C_1]$ æ $[C_2]$ l̩

where C_1 could either be zero, or the consonant /d/, /k/, /s/ or /r/. The intervocalic

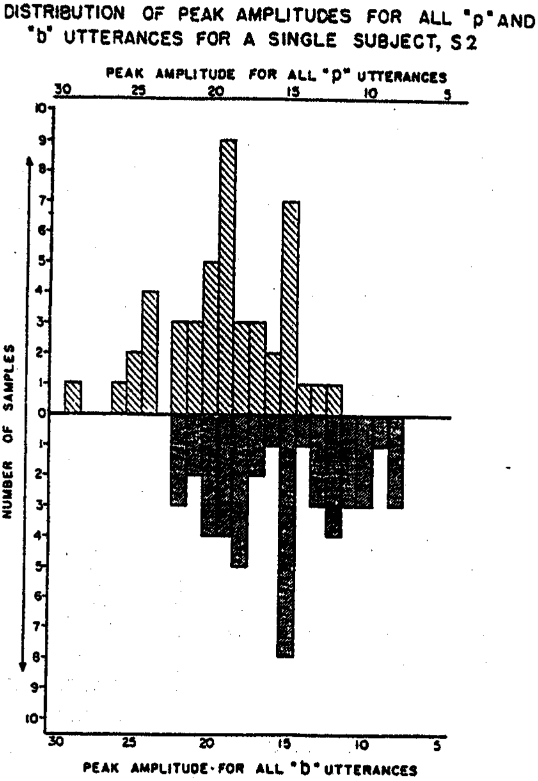

Fig. 4. Histogram of frequency distribution of individual peak response magnitudes for /p/ and /b/ lip closure for a single subject, Experiment I.

consonant, C_2, could be /p/, /b/, /m/, /mp/ or /mb/. There were eight repetitions of each utterance by each subject. All permutations lead to a list of 200 utterances.

The methods of recording and analysing the data were exactly the same as in Experiment I, except for two minor changes. Records were not made from the "lip-opening" positions, and a single electrode (monopolar recording) was used to register "lip closing". The electrode was placed on the upper lip with a distant reference electrode located on an ear lobe. This change does not seem to have affected the shape of the response. Again, peak heights of the integrated traces were measured.

The results of the analysis of variance are shown in Table 3. If there was a significant size difference between "tense" and "lax" consonants, we would expect /p/ would be larger than /b/ and /m/, and perhaps even, that /mp/ would be

TABLE 3

Analysis of variance in lip closing gesture.

EXPERIMENT II

Source	*Mean Square*	*F Ratio*
Medials	280.3575	2.3944
Initials	91.9575	<1.0
Subjects	1,469.2625	12.2222*
I x M	103.2188	3.1269*
M x S	120.1994	3.6413*
I x S	48.2306	1.4611
I x M x S	22.3453	<1.0

larger than /mb/. This would appear in the analysis of variance as a significant *F* ratio for the variable "medials". Again, we do not find a significant value for medials, but there is a significant interaction between subjects and medials. This result exactly conforms to the result of the previous experiment, where the value for "p/b x S" was significant; the other main effect that shows significance, that for subjects, represents again a failure to equalize gains perfectly.

We can look more closely at the result for M x S by plotting average peak heights for each subject. These plots are shown as Fig. 5. The height for /p/ is larger than for /b/ or /m/ in four out of five cases with one reversal, as before. Interestingly enough, /mp/ is larger than /mb/ for the same proportion of subjects, but the reversal occurs for a different individual. Incidentally, different subjects served in the two experiments. We do not know how to interpret the significant I x M term without further work.

Experiment III

Experiment III is much like Experiment II, except that the equipment was much improved in a variety of ways, most of which are not relevant here. One change is that calibration procedures were improved, permitting a transformation of height measurements into microvolts. In the experiment, a single subject read utterances like those of Experiment II, except that the initial consonant was zero, /t/ or /s/. There were twelve repetitions of each utterance.

The measurement procedure was somewhat changed. Instead of measuring peak amplitude, each trace was measured at 20 msec. intervals, using the onset of voicing as "zero" time. The ordinates were then averaged for each measurement interval, summing like ordinates for all three initial consonants to obtain a measure for each medial consonant so that each ordinate represents the average of 36 measurements.

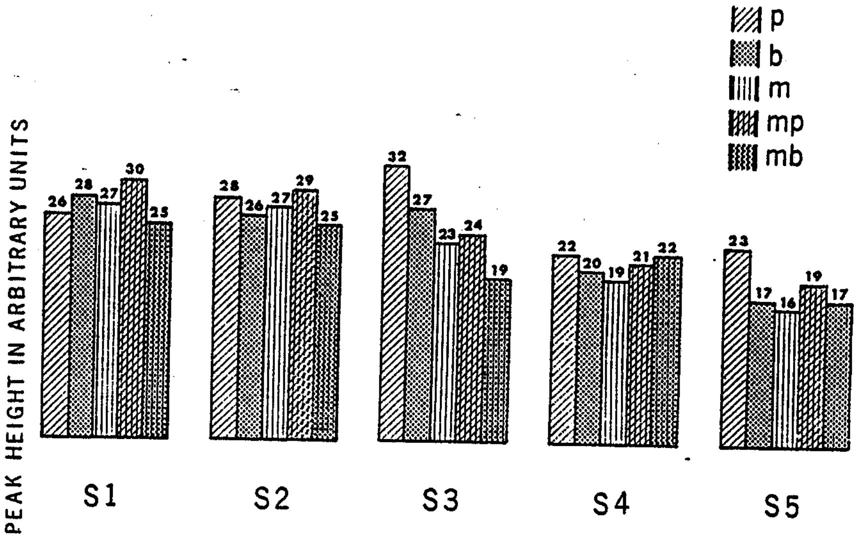

Fig. 5. Histogram of average peak response magnitudes for /p/, /b/, /m/, /mp/ and /mb/ lip closure for five subjects, Experiment II.

The resulting data are shown in Fig. 6 for /p/, /b/ and /m/, and in Fig. 7 for /mp/ and /mb/.

The shape of the response is very similar for /p/, /b/ and /m/, as can be seen from Fig. 6. About halfway through the first syllable, /sæ/, /tæ/ or /æ/, there is a sharp rise in activity, which peaks at about the time the lips are closed for the labial stop.[2] There is very little difference in the average peak height between /p/ and /b-m/. Indeed, the shapes of the two classes are also closely similar, suggesting that the difference in peak height was a reasonable measure to choose of the overall difference in response in the previous experiments. The /p/ response has a slight tendency to have sharper sides than those for /b/ or /m/, although this tendency would not be

---

[2] *There appears to be a secondary peak associated with lip opening in some traces. The second peak is not observed for all subjects. We presume that it represents the activity of a group of muscles other than the orbicularis oris, but further work will be necessary to be specific on this point.*

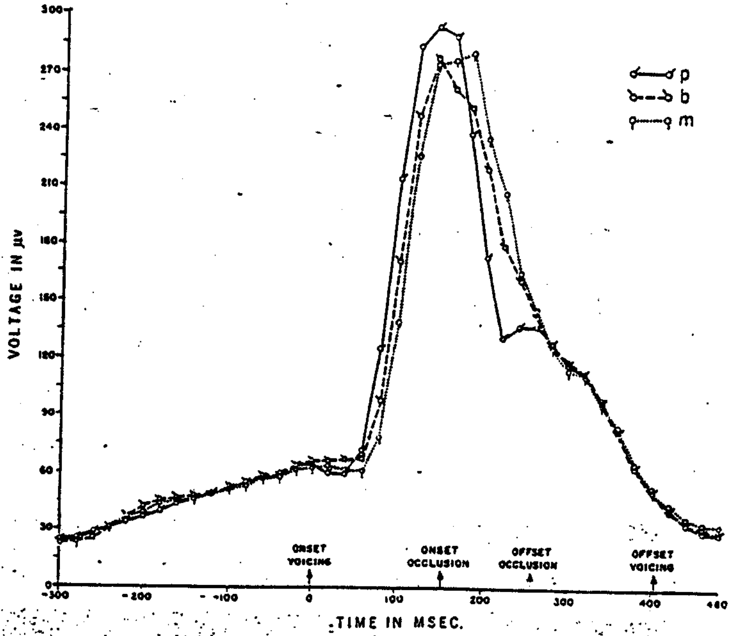

Fig. 6. Average response curve for /p/, /b/, and /m/, Experiment III. Each plotted point represents the average value for 36 utterances, measured at the indicated time separation from the onset of voicing.

visible for individual response-class members.[3] This slight difference in average values is also observed for the /mp/ and /mb/ responses, but again does not appear large or stable enough to serve as a basis for the phonemic distinction.

---

[3] *There are a few fine points which are worth mentioning. First, the averaging technique we used will tend to favour a sharper peak for /p/. Traces are lined up for averaging at the onset of voicing. As is well known, vowels preceding a voiced consonant are longer than before a voiceless consonant. Hence the /p/ peaks will be closer to their line-up, which will tend to produce better superposition. Of course we did not have the complicating factor in Experiments I and II, where we simply measured peak height regardless of time position.*

The second, related point is that there appears to be a second peak for /p/ in this trace, which is not visible for /b/ or /m/. As we said, the peak appears to be related to lip opening; probably it would show in the /b/ and /m/ records if registration were better. It is perfectly obvious in both the /mp/ and /mb/ records, where the first and second peaks are further apart.

This averaging problem is soluble by any of several techniques, the most obvious of which is to use multiple rather than single alignment points. We intend to do further work on the alignment problem.

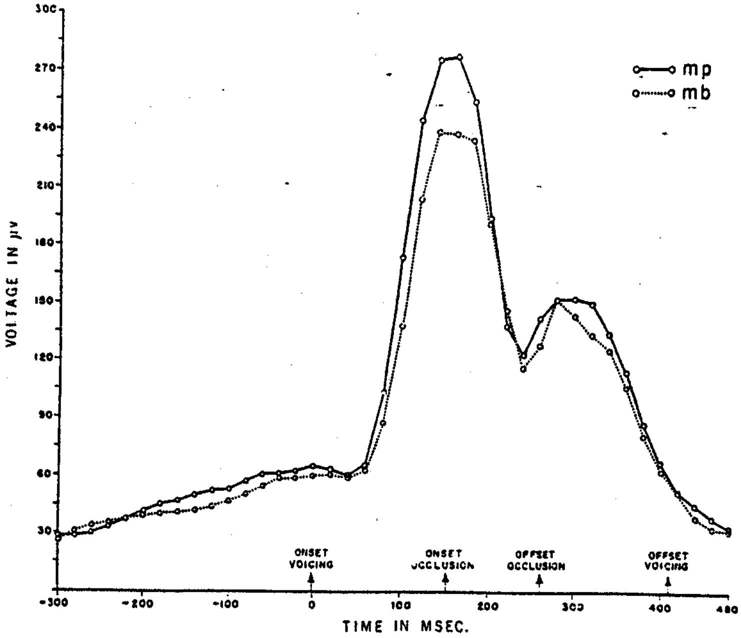

Fig. 7. Average response curve for /mp/ and /mb/, Experiment III. Plotted as in Fig. 6.

Discussion

These experiments were performed for two reasons. First, we wanted to find out if electromyographic data on the production of the sounds of English could be made to yield information about the nature of those sounds which would be useful and complementary to information from other sources. The study has served its purpose as a trial experiment although subsequent work (MacNeilage and Sholes, 1964) has established the feasibility of this type of approach somewhat more comprehensively. Second, we were interested in pinning down a rather vague referent which has floated through the literature for many years, to see what it might mean. Are tense consonants pronounced with more energy than their lax counterparts ? The answer seems to be that such differences exist on the average, but they are small at best and non-existent for some subjects. The differences observed in these experiments could not serve as the basis for a workable phonemic distinction based on muscular tenseness or laxness.

Unfortunately, a negative result is never quite as satisfactory as a positive one. It might be argued that there is some sort of a difference in energy of production between tense and lax consonants, but that electromyographic methods are too insensitive to register them. The only comment we can make is that sensitivity seems entirely adequate, though it is always possible that our search for "diagnostic" electrode placements was incomplete. It may be noted, in this connection, that in a study of the vowels made subsequent to the collection of these data (MacNeilage and Sholes, 1964) it was found that electromyographic methods were at least sensitive enough to show discriminably different tracings for each class of American English vowels from a variety of locations.

A stronger argument is that the difference between the bilabial stops can be accounted for on the basis of the behaviour of the glottis. During the series recorded here, the vocal cords were always vibrating at the time the stop occlusion was released for voiced sounds, and never were vibrating at the time of release of voiceless sounds. This difference in glottal behaviour has been extensively investigated by workers at this laboratory (Lisker and Abramson, 1964), although in initial position. It is their conclusion that "It would seem that such features as voicing, aspiration and force of articulation are predictable consequences of differences in relative timing of events at the glottis and at the place of oral occlusion".

REFERENCES

BENNETT, C. A. and FRANKLIN, N. L. (1954). Statistical Analysis in Chemistry and the Chemical Industry (New York and London).

FANT, G. (1960). Acoustic Theory of Speech Production (S'Gravenhage).

HARRIS, K. S., ROSOV, R., COOPER, F. S. and LYSAUGHT, G. F. (1964). A multiple suction electrode system. *Electroenceph. Clin. Neurophysiol.*, 17, 698.

JAKOBSON, R. and HALLE, M. (1962). Tenseness and laxness. In *Selected Writings of Roman Jakobson* (S'Gravenhage).

LISKER, L. (1963). On Hultzén's "Voiceless lenis stops in prevocalic clusters." *Word*, 19, 376.

LISKER, L. and ABRAMSON, A. S. (1964). A cross-language study of voicing in initial stops: acoustical measurements. *Word*, 20, 384.

LYSAUGHT, G. F., ROSOV, R. J. and HARRIS, K. S. (1961). Electromyography as a speech research technique with an application to labial stops. *J. acoust. Soc. Amer.*, 33, 842.

MACNEILAGE, P. F. and SHOLES, G. N. (1964). An electromyographic study of the tongue during vowel production. *J. Sp. Hrg. Res.*, 7, 209.

PORTER, H. C. (1963). Extraction of pitch from the trachea. AFCRL-63-24. Research Report, Communication Sciences Laboratory, Air Force Cambridge Research Laboratories, L. G. Hanscom Field, Mass.