for use in the megacycle region can be simply constructed with a tan *rho-c* rubber liner. At a frequency of 3.35 megacycles a $\frac{1}{4}$-inch thick liner provides results which are close to optimum, yielding a reflection coefficient of about −30 db up to an angle of incidence of about 60°. Above 60° the liner is highly ineffective. Theoretically, one would desire that the real part of the dilatational velocity in the rubber be slightly higher than that in water. In addition, there is an optimum dilatational attenuation coefficient which is frequency dependent.

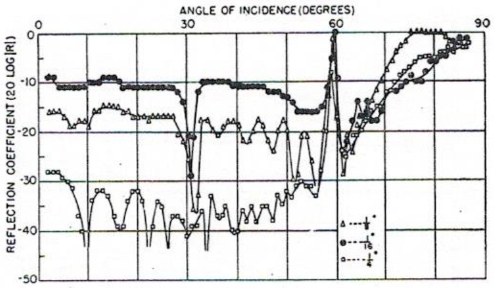

FIG. 10. Experimental reflection coefficient for an aluminum backed *rho-c* rubber sheet of indicated thickness at a frequency of 3.35 megacycles/sec and beam width of 3°.

ACKNOWLEDGMENTS

The author wishes to thank Dr. W. C. Wineland for his many helpful suggestions, and Mr. R. H. Yee for his assistance in carrying out the calculations and experiments.

THE JOURNAL OF THE ACOUSTICAL SOCIETY OF AMERICA VOLUME 25, NUMBER 1 JANUARY, 1953

The Influence of Consonant Environment upon the Secondary Acoustical Characteristics of Vowels

ARTHUR S. HOUSE AND GRANT FAIRBANKS
Speech Research Laboratory, University of Illinois, Urbana, Illinois
(Received September 24, 1952)

The consonant environments of vowels were varied by forming nonmeaningful stimulus syllables consisting of 72 combinations of six vowels and 12 consonants. The syllables were spoken by subjects, and the duration, fundamental frequency, and relative power of the vowels were measured. All three factors varied significantly in response to changes of the consonant environment. The variations were systematically related to the attributes of the consonants, the most powerful attribute being the presence or absence of vocal fold vibration, followed by manner of articulation and place of articulation, in that order.

AMONG the acoustical investigations of vowels, experiments employing wave analysis have naturally been most numerous, while studies of the secondary characteristics—duration, fundamental frequency, and intensity—have been relatively few. Especially has this been true of variation in these secondary characteristics which may be systematically related to the widely varying consonantal environments of vowels in words.

Fairbanks, House, and Stevens,[1] reporting the results of an experiment on the relative intensities of vowels, concluded that, "When the same vowel is spoken in different isolated words, its intensity sometimes varies significantly from word to word, and it seems probable that such variations are, in part at least, effects of differing consonantal environments." In that investigation the words spoken by the subjects were monosyllables in which vowels were preceded and followed by consonant elements. All consonants were voiceless and varied unsystematically among stop-plosives, fricatives, and affricates, produced in bilabial, labio-dental, lingua-dental, and velar positions. Since voicing was held constant and since only 10 words were used for each vowel, such variations were implicitly restricted. The finding that significant variation in intensity obtained even under these conditions was of unusual interest. A few previous studies have shown variation in duration also. Heffner and others[2–5] found the duration of vowels to be longer before voiced consonants than before voiceless consonants. Rositzke[6] also reported consonantal influence upon the duration of vowels, as did Hibbitt[7] for diphthongs. In addition to these results, writers such as Jones,[8] Kenyon,[9] and Thomas[10] assert that the voicing and the manner of production of a consonant following a vowel will in-

[1] Fairbanks, House, and Stevens, J. Acoust. Soc. Am. **22**, 457 (1950).

[2] R-M. S. Heffner, Language **16**, 33 (1940).
[3] R-M. S. Heffner, Am. Speech **12**, 128 (1937).
[4] W. P. Lehmann and R-M. S. Heffner, Am. Speech **15**, 377 (1940); **18**, 208 (1943).
[5] W. N. Locke and R-M. S. Heffner, Am. Speech **15**, 74 (1940).
[6] H. A. Rositzke, Language **15**, 99 (1939).
[7] G. W. Hibbitt, *Diphthongs in American Speech* (Columbia University Bookstore, New York, 1948).
[8] D. Jones, *An Outline of English Phonetics* (E. P. Dutton and Company, Inc., New York, 1948), 6th ed.
[9] J. S. Kenyon, *American Pronunciation* (George Wahr, Ann Arbor, 1940).
[10] C. K. Thomas, *An Introduction to the Phonetics of American English* (Ronald Press Company, New York, 1947).

TABLE I. Stimulus items.

	[i]	[e]	[æ]	[ɑ]	[o]	[u]
[p]	hupeep	hupaip	hupap	hupop	hupoap	hupoop
[t]	huteet	hutait	hutat	hutot	hutoat	hutoot
[k]	hukeek	hukaik	hukak	hukok	hukoak	hukook
[f]	hufeef	hufaif	hufaf	hufof	hufoaf	hufoof
[s]	husees	husais	husas	husos	husoas	husoos
[b]	hubeeb	hubaib	hubab	hubob	huboab	huboob
[d]	hudeed	hudaid	hudad	hudod	hudoad	hudood
[g]	hugeeg	hugaig	hugag	hugog	hugoag	hugoog
[v]	huveev	huvaiv	huvav	huvov	huvoav	huvoov
[z]	huzeez	huzaiz	huzaz	huzoz	huzoaz	huzooz
[m]	humeem	humaim	humam	humom	humoam	humoom
[n]	huneen	hunain	hunan	hunon	hunoan	hunoon

fluence its duration. With respect to fundamental frequency, a search of the literature failed to disclose reports of similar variation, although such might be predicted.

The purpose of the present experiment has been to pursue the problem raised by such findings with a more extensive and systematic phonetic design. Its general plan was to place vowels in various consonant environments, to cause them to be spoken by subjects, and to make the appropriate physical measurements. Representative vowels were used, and the following articulatory characteristics of the consonants were controlled: the presence or absence of vocal-fold vibration; variations in the manner of production (fricative, stop-plosive, etc.); variations in the characteristic place of articulation (bilabial, velar, etc.).

I. PROCEDURE

Materials

After considerable study of the various alternatives, it was decided to construct stimulus materials in which only one consonant influence was present in each item. Syllables in which the vowel is both preceded and followed by the same consonant, as in the word *cease*, appeared to be suitable. Study of all such symmetrical syllables led to the conclusion that appropriate materials would be provided by restricting these items to 72, involving 12 consonants in combination with six vowels.

The following 12 consonants were selected: [p], [b], [t], [d], [k], [g], [f], [v], [s], [z], [m], [n]. It will be seen that the first 10 sounds are voiceless-voiced cognate pairs, providing direct contrast for the voicing factor. In manner of production, six stop-plosives, four fricatives, and two nasals are available, in that order. Differences in characteristic place of articulation are provided by the bilabials [p], [b], and [m]; the labiodentals [f] and [v]; the post-dentals [t], [d], [s], [z], and [n]; and the velars [k] and [g].

The six vowels chosen were [i], [e], [æ], [ɑ], [o], and [u]. These vowels span the range of tongue, mandible, and lip positions. They also vary in certain of their secondary acoustic characteristics as reported by Crandall,[11] Parmenter and Treviño,[12] Heffner and others and Black[13] for duration, by Crandall, Black, and Taylor[14] for fundamental frequency, and by Fairbanks, House, and Stevens, Black, and Sacia and Beck[15] for relative intensity.

Combinations of these consonants and vowels resulted in a mixed list of words and nonsense syllables, which appeared to be unsuitable, since control over semantic influences was regarded as important. After various attempts, it was decided to prefix each item by an unstressed syllable, creating bisyllabic nonsense items with iambic stress patterns, and to select orthographic forms that would yield least meaning. Eventually, [hə] was selected as an appropriate initial syllable, since its component sounds are easily pronounced, usually neutral, and seemed likely to have minimal effect upon adjacent sounds. Table I shows the 72 stimulus items in the form used in presentation to the subjects.

Table I may also be regarded as depicting the essentials of the statistical design. With all items produced by all subjects, the triple-classification scheme for analysis of variance was employed to test the significance of the variances attributable to vowels, consonants, subjects, and their first-order interactions, and to enable the other statistical manipulations.[16]

Subjects

Ten male students enrolled in elementary speech courses at the University of Illinois served as subjects. The mean age was 20 years, six months, and the individuals ranged from 18 years, seven months to 26 years, six months. Thirty-five potential subjects were interviewed by two judges trained in phonetics and speech pathology, who screened their speech for aberrations in pitch, loudness, duration, voice quality, articulation, and pronunciation, and questioned them concerning their hearing. The ultimate subjects were without speech disorders, had no history of hearing pathology, and spoke some form of general American dialect. In the case of the intensity data only, 10 additional subjects, selected in the same manner, were added to the basic group.

Apparatus

The procedure involved phonograph recording of the subjects' responses to the stimulus items. Equipment was arranged conventionally in a two-room laboratory, consisting of a sound-treated room and a control room. A General Radio type 759-B sound level meter, with its associated Brush 9898 crystal microphone, was set

11 J. B. Crandall, Bell System Tech. J. 4, 586 (1925).
12 C. E. Parmenter and S. N. Treviño, Am. Speech 10, 129 (1935).
13 J. W. Black, J. Speech Hearing Disorders 14, 216 (1949).
14 H. C. Taylor, J. Exptl. Psychol. 16, 565 (1933).
15 C. F. Sacia and C. J. Beck, Bell System Tech. J. 5, 393 (1926).
16 Q. McNemar, *Psychological Statistics* (John Wiley and Sons, Inc., New York, 1949).

to flat response and 60 db attenuation, and its output fed to a Presto 92-A recording amplifier, with equalizer setting out. The recordings were made at 78.26 rpm on a Presto 8D-G turntable with a Presto 1-B cutting head.

The recordings were played for graphic recording of relative power and for simultaneous verification of the presence of the complete set of responses for each subject. The system used included a Presto 64-A turntable with a Lear PA-200 pick-up, a Goodell PRA-1 preamplifier, a Sound Apparatus Company HPL-E high-speed level recorder with 0–50 db potentiometer, and a Goodell ATB-3 power amplifier for the monitor speaker. The level recorder was set for fast stylus speed and 10 mm/sec paper speed. Measurements of duration were made on sound spectrograms recorded on a Kay Electric Company sona-graph. The above reproducing system was employed. Oscillograms used in the measurement of fundamental frequency were obtained through the use of a locally constructed instrument similar to that described by Cowan.[17] The instrument is essentially a special type of oscillograph for the photographing of signals stored on phonograph records. A turntable is mounted on a large drum around which photographic paper is wrapped. During the playing of the record, the turntable and the drum move synchronously, and an optical lever, activated by a crystal pick-up and amplifier, is manually lowered by means of a helical screw mechanism to spiral the trace about the drum.

Experimental Procedures

The phonograph recordings were made with the subject seated comfortably in the sound-treated room and with the microphone, mounted on a boom stand, approximately 12 inches before his mouth in the horizontal plane. An experimenter was seated immediately in front of the subject and presented the stimulus items visually at approximately four-second intervals. Each item was typed in lower case on a 3×5 card. Two judges trained in phonetics independently evaluated the acceptability of each response, using reasonably liberal standards of articulation and stress pattern. When either judge did not accept a response, its stimulus card was displaced backwards in the order of presentation by at least eight items in order to present it to the subject a second time. An auxiliary list was used at the beginning of the recording session to familiarize the subject with the situation and to permit adjustment of equipment. The same auxiliary list was used to displace backwards in time stimulus items so close to the end of the list that they could not otherwise be displaced by at least eight items. The list of 72 items was randomized anew for each subject.

Before actual recordings were made each subject was given a short training period during which he read aloud all items from a standard list arranged as in Table I. He was assured that no premium would be placed on accuracy of response as such. Instructions pertaining to vocal behavior were kept to a minimum, and no attempt was made to regulate pitch, duration, or intensity. The subject was simply instructed to pronounce each word as if talking to the experimenter.

II. RESULTS

Influence upon Duration of Vowels

The identification of the beginning and end of a vowel surrounded by consonants is an arbitrary act that is both difficult and artificial. Location of these points was aided by the relative clarity with which they are shown in sound spectrograms. In this study, the spectrograms were produced with a narrow, 45-cps, band response at their base and a wide, 300-cps, response in the higher frequencies, and covered a 0–3500-cps range.[18] This procedure simplified the identification of the voice bar without weakening the vowel resonance areas. Identification of specific sounds was facilitated by information presented by Potter, Kopp, and Green[19] and Joos.[20] The preceding and following consonants were used to find the general area of each vowel, but in all instances the vowel limits were established in terms of "vowelness," based on the presence of voice and of resonance areas associated with particular vowels.

Measurements were made in millimeters and multiplied by 0.00754 to yield values in seconds.[21] For purposes of estimating reliability, a number of measurements were also made on the oscillographic films, and comparison of the two methods of measurement showed close agreement. The symmetrical structure of the test syllables greatly simplified the identification of the vowels and tended to create a situation generally favorable to reliability and validity of measurement. Hanley[22]

TABLE II. Duration of vowels in various consonant environments. Vowels pooled. All values in seconds.

Individual consonant environments		Grouped consonant environments	
p	0.159	*Voicing*	
t	0.168	Voiceless (5)	0.174
k	0.157	Voiced (7)	0.253
f	0.188	*Manner of production*	
s	0.197	Stop-plosive (6)	0.203
b	0.237	Fricative (4)	0.239
d	0.258	Nasal (2)	0.232
g	0.239	*Place of production*	
v	0.279	Bilabial (3)	0.205
z	0.291	Labio-dental (2)	0.234
m	0.219	Post-dental (5)	0.232
n	0.245	Velar (2)	0.198

[17] M. Cowan, Arch. Speech 1, Suppl., 1 (1936).

[18] The normal 0–8000-cps range of the sona-graph was modified locally with the advice of the manufacturer. In essence, the slide wire resistor was shunted by an amount determined experimentally and series resistances added to the appropriate leads to keep the total series resistance constant. These modifications resulted in a 0–3500-cps full scale recording.

[19] Potter, Kopp, and Green, *Visible Speech* (D. Van Nostrand Company, Inc., New York, 1947).

[20] M. Joos, Language 24, Suppl., 1 (1948).

[21] The instrument in question records 2.4 sec in 318.5 mm.

[22] T. D. Hanley, Speech Monog. 18, 78 (1951).

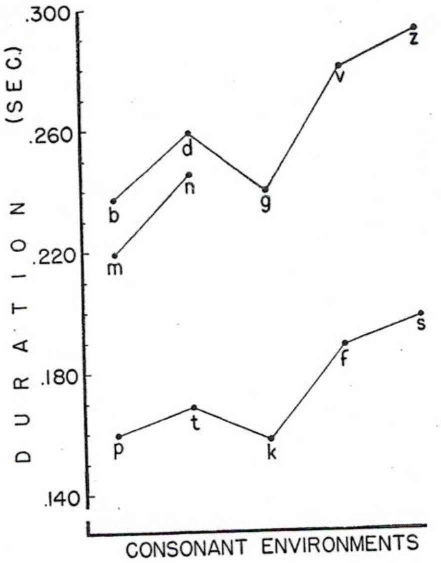

FIG. 1. Mean duration of vowels in various consonant environments. Vowels pooled.

has shown that even in connected speech durational measurements from spectrograms have high reliability.

The effect of the consonant environments on the duration of the vowels is summarized in Table II and Fig. 1. The left half of Table II shows the mean duration of all vowels spoken by the 10 subjects in each of the 12 consonant environments. These means vary over a range of 0.134 sec, and analysis of variance reveals statistically significant differences between means to exist at the 1 percent level of confidence. By use of the t statistic, differences between the various means may be evaluated with a requirement of 0.011 sec. Of the total of 66 such intercomparisons, 59 exceed the minimum.

Further study of the duration means shows that they vary systematically with certain characteristics of consonant production. A comparison of voiceless environments with their voiced cognate environments, for example, reveals larger values for the voiced environments in every case. All voiced environments, furthermore, produced vowels that differed significantly from all those produced in voiceless environments. When all responses are pooled with respect to this characteristic, as in the upper right portion of Table II, there is a statistically significant difference of 0.079 sec between the two means. The marked effect of voicing on duration is demonstrated in Fig. 1. The baseline shows the 12 consonant environments and is arranged with stop-plosive consonants to the left of the figure and fricatives to the right. Cognates and nasal correlatives are shown in the same vertical plane. Within each manner of production category sounds are arranged from left to right according to place of articulation along the antero-posterior dimension of the oral cavity (e.g., the three stop-plosives are bilabial, post-dental, and velar in order).

Attempts to interpret the effect of voicing of the consonant upon vowel duration have thus far been fruitless. It may be, for example, that the voicing of a vowel in a voiceless environment, in contrast to a voiced environment, is withheld until the physiological vowel "target" is more nearly approximated, and terminated sooner in the transition to the following consonant. The problem seems to require additional experimentation with the transition intervals, particularly of a type employing simultaneous acoustical and physiological measurements.

Table II also shows that the values for vowels surrounded by stop-plosive, fricative, and nasal classes, according to manner of production, vary over a 0.036-sec range, and demonstrate means that differ significantly. The close similarity between the fricative and nasal means is of interest, since nasal sounds are related to stop-plosives physiologically and to fricatives dynamically. The values suggest a stop-continuant dichotomy, but as the nasal class is composed only of voiced sounds, which have been shown to increase the duration of contiguous vowels, this problem cannot be analyzed definitively with these data.

Reinspection of the 12 individual environment means in Table II indicates that, voicing constant, consonants that differ in manner of production produce vowel durations that usually differ significantly. Figure 1 also shows, in both voiced and voiceless lines, the trend for fricative sounds to prolong vowels more than do stop-plosives. Apparently, the gradual, controlled movements of continuant consonants favor longer vowel durations more than do the abrupt, ballistic movements of the stop-plosives.

The remaining part of Table II shows that the duration of vowels also varies when the responses are sorted according to place of consonant articulation. The differences reach significance, but this result should be interpreted with caution in view of the findings on voicing and manner of production. Both velar consonants are stop-plosives, and both labio-dental are fricatives, while the bilabials and post-dentals are weighted with voiced consonants. The two curves for voiceless and

TABLE III. Fundamental frequency of vowels in various consonant environments. Vowels pooled. All values in cycles per second.

Individual consonant environments		Grouped consonant environments	
		Voiced	
p	127.86	Voiceless (5)	126.46
t	127.07	Voiced (7)	121.99
k	127.17	*Manner of production*	
f	124.32	Stop-plosive (6)	124.36
s	126.06	Fricative (4)	123.80
b	120.86	Nasal (2)	122.47
d	120.58	*Place of production*	
g	122.77	Bilabial (3)	123.90
v	122.20	Labio-dental (2)	123.26
z	122.62	Post-dental (5)	123.62
m	123.16	Velar (2)	124.97
n	121.77		

voiced cognates in Fig. 1, however, are remarkably similar in shape, seemingly illustrating characteristic differences between the effects of consonants, voicing constant, for place as well as for manner of articulation.

The three interactions between consonant environment, vowels and subjects were significant at the 1 percent level.

Influence upon Fundamental Frequency of Vowels

The oscillograms described above were synchronized with the spectrograms used in the measurement of duration, and analogous points in the vowel wave forms nearest to the limits of the duration interval were identified. Measurement of this distance was made in centimeters rounded to the nearest quarter of a millimeter and divided by the integral number of cycles which it subtended, yielding mean period in cm. This value, divided into the film speed, 249.3 cm/sec,[23] gave the mean fundamental frequency of the vowel.

The means for the different consonant environments shown in the left column of Table III vary significantly over a range of 7.28 cps. Comparison of these mean values to requirements of 2.40 cps and 1.82 cps at the 1-percent and 5-percent levels, respectively, shows that 41 of the 66 possible differences exceed the minimum at the 5-percent level, and that 35 of these exceed the 1-percent requirement.

Inspection of these means reveals that the fundamental frequencies of vowels in voiceless environments are invariably higher than those in voiced environments. With the exception of [f] compared to [m] and [g], all of these differences are significant at the 5-percent level or better. The means are graphed in Fig. 2, the arrangement being the same as for Fig. 1. Voiceless and voiced groups were formed, with means as shown in the upper right of Table III, and the difference was tested with the expected result. The conclusion is reached that the presence or absence of vocal-fold vibration during consonants has a real effect upon the fundamental frequency of adjacent vowels in the direction mentioned.

In an attempt to explain this effect, frequency curves of a number of complete responses were plotted from the oscillograms, using 0.05-sec sampling intervals. From these curves, and also by ear in the responses generally, it was observed that the pitch inflection of the first, and unstressed, syllable [hə] was usually downward and reached a frequency considerably lower than the dominant level of the vowel in the following stressed syllable. Phonation continued from this low frequency with a rising inflection into the second syllable when the initial consonant of the second syllable was voiced. When that consonant was voiceless, and the characteristic interruption of phonation separated the two syllables, the voicing of the stressed syllable usually started at a higher frequency. It was suspected, further that the fundamental frequency of voiced consonant environments might influence that of the vowel. Crandall[11] reports lower fundamental frequencies for voiced consonants than for vowels, but his conditions were dissimilar. As it was deemed impractical at this time to make such a comparison with all 720 responses in the present study, one response involving voiced continuant consonants was selected at random from each of the 10 subjects. By chance such selection yielded syllables including four vowels and all four consonants, each of the latter at least twice. For each response the frequencies of the consonants immediately contiguous to the vowel were measured on the oscillogram. A sample of 0.10 sec was used, unless the duration of the consonant was shorter than that interval. For the preceding and following consonants, respectively, the means were three and 10 cps lower than that of the vowels. In other words, the calculations indicated a consonant-vowel-consonant inflection that was circumflex. No definitive explanation can be advanced on the basis of these data, but if the natural fundamental frequency of voiced consonants is lower than that of vowels, it is not implausible to suggest that a vowel surrounded by voiced consonants might have a lower mean fundamental frequency than when these influences are absent. Further study of this problem is in progress.

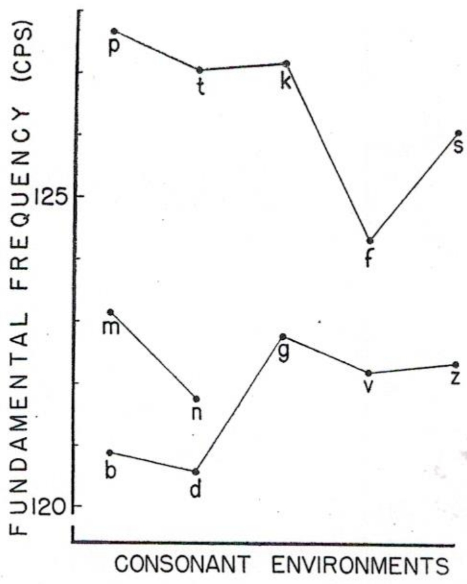

FIG. 2. Mean fundamental frequency of vowels in various consonant environments. Vowels pooled.

Comparison of the fundamental frequencies of vowels with respect to the manner of production of adjacent consonants, stop-plosive, fricative, or nasal, is somewhat complex. It may be carried out most readily by studying in conjunction the left column of means in Table II, especially as graphed in Fig. 2, and the right central portion of that same table. The latter measures differ significantly, but the absolute size of the maxi-

[23] Recording paper length × recording speed/sec per min $=191.2\times78.26/60=249.3$.

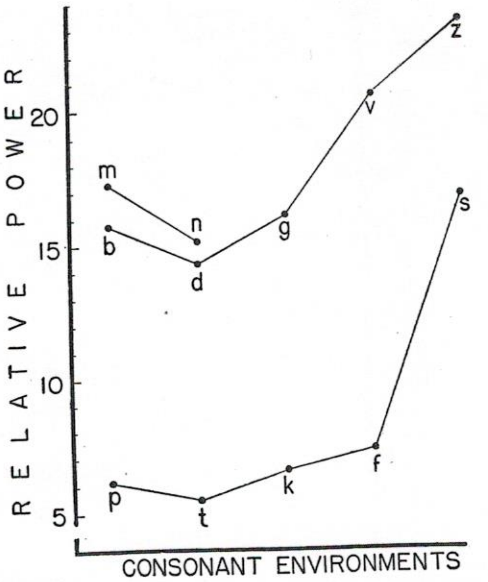

FIG. 3. Mean relative power of vowels in various consonant environments. Vowels pooled.

mum difference is only 1.89 cps. Study of Fig. 2 and evaluation of the statistical significance of the differences there shown discloses that although individual consonant environments may differ in their effects from class to class, the differences are small, often not significant, and variable in direction, when voicing, the predominant factor, is held constant.

The effects of varying the characteristic place of articulation are generally similar. A test of significance of the lower right means in Table III allows the rejection of the hypothesis of no difference at the 5-percent level of confidence, but a close inspection of the variables suggests that most of the differences probably may be attributed to chance.

The interactions involving subjects reached statistical significance, while the vowel-by-consonant interaction failed to reach the 5-percent level. The relative variation of the individual vowels in the differing consonant environments was strikingly similar.

General study of the data presented in this section indicates that the effects of consonants upon fundamental frequency, although significant, are probably less than the variations in fundamental frequency natural to the vowels themselves when consonant environments are constant (see Table V).

Influence upon Relative Power of Vowels

The intensity curves produced by the high speed level recorder showed typical bimodal forms with the second and greater mode corresponding to the stressed syllable. For each such syllable the maximum level was measured. The phonetic structure of the material and the characteristics of the records allow the assumptions that this point was reached during the production of the vowel, that it occurred within the time interval measured for duration and fundamental frequency, and that it furnished valid data concerning the intensity of the vowel. The measurements for each subject were expressed in db above the lowest value for that subject and in turn converted to relative power to facilitate arithmetic treatment.[24] These manipulations tended to minimize variation in intensity from subject to subject. It will be recalled that this vocal characteristic was not controlled in the original procedure, where each subject was permitted to establish his own general level. The existence of individual differences in vocal output is a well-known phenomenon and was not of interest in this study. It should be noted further that results of analysis of variance indicate that subject variation was not completely obliterated.

Study of the consonant environment means in the left column of Table IV shows them to range from 5.43 to 23.28, and a test of this variation allows the rejection of the null hypothesis at the 1-percent level. If the 66 possible differences between these means are evaluated against a requirement of 4.04, 42 are seen to be significant at the 1-percent level. Further study of these means shows that voicing of the consonant environment was almost uniformly productive of greater mean power. When voiceless and voiced categories are formed, as in the right of the table, this difference is seen to be large on the average and is significant. Statistical evaluation of the differences between the individual consonant means shows that all voiced environments produced significantly greater power than all voiceless environments with the single exception of [s]. The nature of these differences may be visualized in the graphs of these means in Fig. 3. This finding would be an expected one in that the continuation of phonation throughout the consonants as well as the vowel would be likely to favor greater maximum intensity.

The data on the power of vowels in stop-plosive, fricative and nasal environments are presented in Table IV and represent statistically significant differences. These variations may be observed in Fig. 3. In view of the marked effect of voicing, mentioned above, the

TABLE IV. Relative power of vowels in various consonant environments. Vowels pooled. All values in relative power.

Individual consonant environments		Grouped consonant environments	
		Voicing	
p	6.11	Voiceless (5)	8.40
t	5.43	Voiced (7)	17.46
k	6.48	*Manner of production*	
f	7.21	Stop-plosive (6)	10.68
s	16.78	Fricative (4)	16.94
b	15.72	Nasal (2)	16.19
d	14.28	*Place of production*	
g	16.08	Bilabial (3)	13.03
v	20.52	Labio-dental (2)	13.87
z	23.28	Post-dental (5)	14.98
m	17.25	Velar (2)	11.28
n	15.13		

[24] Relative power was taken as equal to antilog$_{10}$ $N/10$, where N was expressed in db.

comparison is probably valid only for the stop-plosive and fricative classes. Voicing constant, differences between the individual fricative and stop-plosive environment means of Table IV generally are significant. An exception is the fricative [f], which did not differ from the voiceless stop-plosives.

The lower right section of Table IV shows that the effect of the place of production of adjacent consonants is small. Although this effect is statistically significant, since the low intensity velar environments are all voiceless, and the post-dental environments at the other extreme are weighted in favor of voicing, the evidence for differences caused by variation in place of articulation is regarded as inconclusive. Nevertheless, the semi-parallel curves of Fig. 3 indicate similarities of consonant effects within a given voicing class.

When relative power is considered, the interaction between vowels and consonants reaches significance barely at the 5-percent level. The interactions involving subjects are significant.

Differences between Vowels

Table V presents the data for the six specific vowels listed in order of the conventional physiological vowel diagram. The general behavior of the three secondary characteristics may be observed in the top row of each section of Table V where the various environments have been pooled, and these data are also graphed in Fig. 4. Analyses of the variances attributable to vowels showed them to be significant beyond the 1-percent level for all three variables.

It will be seen that the duration of vowels is directly related to size of mouth opening and inversely related to tongue height. The conformity of [e] and [o] to the progression is interesting, since they are commonly diphthongized and longer durations would not have been surprising. These trends are in general agreement with the data reported by Black,[13] by Heffner and others,[2–5] and by Parmenter and Treviño,[12] while Crandall's data[11] show an inversion of the tendency. The results shed considerable doubt upon the classification of [i], [ɑ], and [u] as "long" vowels and of [e], [æ], and [o] as "short" vowels by Jones,[8] and upon the assertion by Thomas[10] that "tense" vowels are generally longer in duration than "lax" vowels. That durational variation should progress as shown is plausible, and probably may be explained on grounds of varying extent of articulatory movement with correspondingly varying time.

In the middle graph of Fig. 4 it will be observed that fundamental frequency varies systematically and directly with the usual vertical location of the high point of the tongue. This finding of a "vowel-pitch triangle" is in general agreement with data reported by Crandall,[11] Black,[13] Peterson and Barney,[25] and Taylor.[14] The concomitant variation of fundamental frequency and tongue position has been explained by the latter author as dynamogenetic radiation from the tongue musculature to the laryngeal muscles controlling the tension of the vocal folds. Thus, in comparison to a "low" vowel, the increase in tongue height of a "high" vowel is accompanied by increased tension of the tongue musculature. Such variations in degree of tension are irradiated to the laryngeal musculature, producing corre-

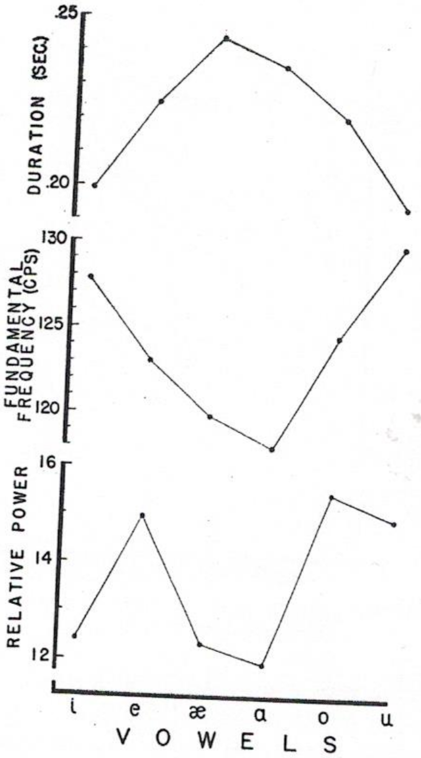

FIG. 4. Mean duration, fundamental frequency, and relative power of vowels. Consonant environments pooled.

TABLE V. Acoustic characteristics of specific vowels in various types of consonant environments. Top line of each section shows vowel means when all environments are pooled; remaining lines show vowel means for five mutually exclusive classes of environment. See also Figs. 4 and 5.

	[i]	[e]	[æ]	[ɑ]	[o]	[u]
			Duration (sec)			
All environments (12)	0.199	0.225	0.244	0.236	0.221	0.195
Voiceless stops (3)	0.138	0.171	0.184	0.180	0.157	0.138
Voiceless fricatives (2)	0.177	0.199	0.215	0.218	0.187	0.161
Nasals (2)	0.209	0.238	0.253	0.235	0.241	0.217
Voiced stops (3)	0.215	0.251	0.276	0.267	0.244	0.215
Voiced fricatives (2)	0.277	0.283	0.304	0.295	0.293	0.261
			Fundamental frequency (cps)			
All environments (12)	127.86	123.03	119.78	118.00	124.64	129.83
Voiceless stops (3)	132.07	126.44	122.00	120.97	128.57	133.79
Voiceless fricatives (2)	129.59	124.05	119.80	119.59	125.86	132.86
Nasals (2)	125.89	119.93	118.75	117.78	122.64	129.83
Voiced stops (3)	125.34	121.02	117.87	115.80	123.13	125.26
Voiced fricatives (2)	125.58	123.00	119.69	116.25	121.22	127.69
			Relative power (see text)			
All environments (12)	12.43	14.94	12.35	11.97	15.49	14.94
Voiceless stops (3)	4.16	8.30	5.50	5.78	7.35	4.91
Voiceless fricatives (2)	12.43	12.70	11.39	11.17	12.83	11.48
Nasals (2)	17.05	15.96	17.18	12.14	16.84	17.97
Voiced stops (3)	13.80	16.65	13.66	14.94	18.29	14.78
Voiced fricatives (2)	18.19	23.58	16.74	17.42	24.82	30.67

[25] G. E. Peterson and H. L. Barney, J. Acoust Soc. Am. 24, 175 (1952).

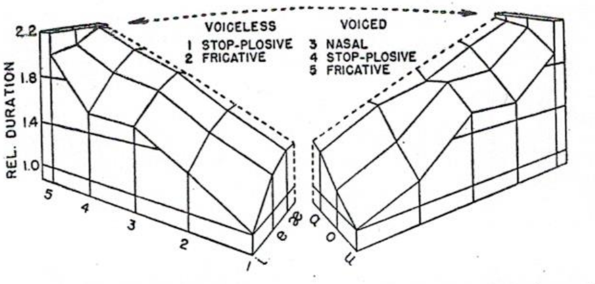

Fig. 5. Three-dimensional graph showing mean duration of specific vowels in various classes of consonant environment. Figure is divided along the median plane and the two halves rotated to show both sides.

sponding variations in vocal-fold tension and in the fundamental frequency of the output.

The group data for relative power are shown in the lowest portion of Fig. 4. With this arrangement of vowels, regularity of progression, as in the case of duration and fundamental frequency, is absent. The curve also differs in certain respects from previous results. While the vowels [æ] and [ɑ] are here seen to be lowest in mean power, Sacia and Beck,[15] Black,[13] and Fairbanks, House and Stevens[1] report reverse findings. These vowels, furthermore, are known to have the largest anterior diameters of the vocal conduit, an aspect which Fairbanks[26] has shown to be closely related to vowel intensity. The vowels fall into two groups of greater and lesser power, and within each group the range is very small. While the general analysis of variance revealed significant differences among the vowels, comparison of the individual means within each of the two groups of vowels mentioned yielded no statistically significant differences.

It would appear that this atypical vowel curve, if not resulting from chance, might be a product of the present experimental design, which differs from those of previous investigations. For one thing, consonant environments were less restricted in the present experiment, which might be important in view of the substantial variance for consonants mentioned above. Investigation of this factor by regrouping according to the three main characteristics of consonant environment, however, showed that similar curves were found under all these conditions. Another difference from previous experiments, although it seems an unlikely source, is that the stimuli in this instance were bisyllabic with the syllable studied being preceded by a common unstressed syllable. A third difference that should be mentioned was that all stimuli were nonmeaningful, although this factor would appear to operate, if at all, in the opposite direction. A more plausible factor is the phonetic symmetry of the present syllables, which required a subject to begin from and return to the same consonant position. It seems reasonable to suggest that this condition might restrict the extent of movement to the vowel position, and that the restriction might be greatest for vowels of normally large mouth opening such as [æ] and [ɑ]. An additional condition, involving the spelling of stimulus items, should be mentioned. It will be seen in Table I that the vowels [æ] and [ɑ] were represented for the subjects by spellings of one letter while the other vowels were spelled with two letters. Also, it would seem likely that the spoken vowels intended by the four two-letter spellings would, in general, be more clear to the naive subject, and during the course of the experiment it was observed that subjects conditioned more swiftly to them than to the one-letter spellings. It will be recalled that the original recordings included those responses which necessitated repetition because of misarticulation of the vowels. A count of these showed that [æ] and [ɑ] were involved approximately equally in more than two-thirds of the total, or double the chance expectancy. A confident explanation of the atypical findings for relative power cannot be advanced, but both of the latter two possibilities present problems which are themselves worthy of investigation.

It having been shown that the three dependent variables vary significantly among vowels, and that consonant environments also exert significant and systematic influences, it is of interest to determine whether the influence of one is to be felt under the various conditions of the other. The main portions of Table V show means for each of the six vowels in each of five different, mutually exclusive classes of consonant environment. These classes involved voicing and manner of articulation, the two factors demonstrated to be most powerful. Examination of the data will show the influence of both vowel and consonant environment. Thus, the values for any given consonant environment change from vowel to vowel in a manner generally similar to the change when all environments are pooled. The means in any column are observed to progress more

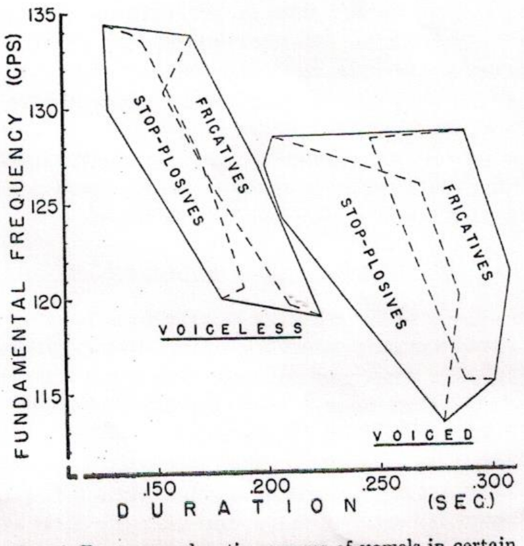

Fig. 6. Frequency-duration ranges of vowels in certain consonant environments.

[26] G. Fairbanks, Speech Monog. 17, 390 (1950).

or less systematically down the column in the direction shown to be generally appropriate for that variable when all vowels are pooled.

The nature of these interactions may best be appreciated by reference to Fig. 5 which displays the values for duration in Table V and is exemplary of the general findings. In this three-dimensional figure, the ordinate is duration while consonant classes and vowels are shown along the horizontal axes. The figure is split and spread to show both sides. The systematic variation of vowel duration in response to changes in both consonant environment and in vowel is clearly seen. It is of considerable interest that the influence of neither factor is obscured by the other, which is a finding of obvious implication for experimental design.

Interrelationships between Acoustical Characteristics

In the above discussion, duration, fundamental frequency, and power have been considered separately. In this section they are brought together for purposes of illustrating their covariation.

Figure 6 shows the effects of voicing and manner of production of consonants on vowels. The figure depicts frequency-duration areas for voiced and voiceless stop-plosives and fricatives. Each area boundary connects the most divergent coordinate values, i.e., the maximally varying vowel means, in each type of consonant environment shown. The characteristic influences of the voiced and voiceless groups upon both duration and frequency, the contrasting effects of cognate environments, and the distinct differences in duration between stop-plosives and fricatives are readily apparent.

A concise illustration of the major findings of the experiment is found in Fig. 7 which shows the means of the pooled vowels for the 12 consonant environments. The ordinate is frequency, the abscissa is duration, and the diameter of the dots is proportional to relative power. Substantial intercorrelations are evident between all dimensions. It will be observed that vowels in voiced environments are, in general, longer in duration, lower in fundamental frequency, and greater in power than are the same vowels when in voiceless environments. Within voicing groups, clusters corresponding to manner of production may also be found.

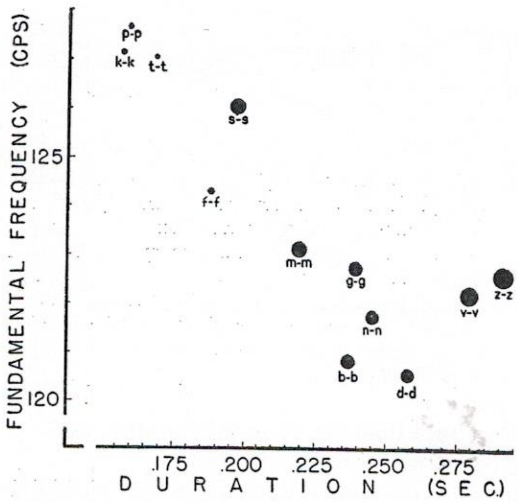

FIG. 7. Relationships between duration, fundamental frequency, and relative power of vowels in various consonant environments. Relative power proportional to dot diameter. Vowels pooled.

III. SUMMARY AND CONCLUSIONS

In a study of the influence of consonant environment upon the secondary acoustical characteristics of vowels, the subjects spoke 72 different consonant-vowel-consonant syllables in each of which the vowel was both preceded and followed by the same consonant. Twelve representative consonants were combined with six representative vowels. Acoustical measurements of the duration, fundamental frequency, and intensity of the vowel of each syllable were made, and analyzed with special reference to variance attributable to the articulatory characteristics of the consonants. Following were the major results.

1. Consonant environment significantly influenced all three acoustical characteristics of the vowels. Of the types of consonant influences studied, the effects of voicing were greatest. In the comparisons of voiced and voiceless consonant environments, vowels in voiced environments, with few exceptions, were longer in duration, lower in fundamental frequency, and greater in relative power.

2. Manner of production was the second most influential consonant characteristic. Its effect upon the duration and relative power of vowels was more consistent than upon fundamental frequency, although all three varied significantly.

3. Place of articulation appeared to be the least important of the consonant characteristics, but its influence may have been obscured by the conditions of the experiment.

4. When all consonant environments were pooled, significant differences between vowels were found in all three acoustical characteristics. From vowel to vowel, duration and fundamental frequency varied in a manner systematically related to the usual conceptions of vowel physiology, while variations in relative power were atypical.

5. When changes in the acoustical characteristics of the vowels were examined in relation to variations of both consonant environment and vowel, the influence of neither factor obliterated that of the other.

6. Variations of the three acoustical characteristics in response to changing consonant environments were substantially intercorrelated.