Phonetica *22:* 129–159 (1970)

Vowel Length Variation as a Function of the Voicing of the Consonant Environment

MATTHEW CHEN

University of California, Berkeley, Calif.

Introduction: Defining the Problem

The present paper purports to deal with a sub-problem regarding vowel length[1]. Vowel length is known to vary owing to a number of factors [DELATTRE, 1962]. In this paper I propose to study just one of these, namely the voicing[2] of the consonantal environment, which has been shown to exercize the greatest influence on the syllable nucleus [HOUSE and FAIRBANKS, 1953].

Since from a preliminary survey of published sources, *preceding* consonants exhibit no readily discernible patterns of environmental influence on the duration of neighbouring vowels [PETERSON and LEHISTE, 1960], we shall restrict ourselves to the influence of the *following* consonants on the preceding vowel length. In a word, we shall attempt to examine the variations of vowel length as a function of the [±voice] feature of the following consonant.

Phoneticians have long observed the fact that in English vowels are generally longer before voiced consonants than before voiceless consonants. This fact seems reasonably well established as far as English is

[1] The research reported in the present paper was supported in part by the National Science Foundation Grants GS2108 and GS 1430. I am indebted to WILLIAM S. Y. WANG and PETER F. MACNEILAGE for their constructive comments on earlier versions of this paper and for their invaluable assistance on the various stages of its preparation. A special word of thanks goes to ROBERT KRONES, who was particularly helpful in the technical aspects of the experiments performed in connection with this paper. An earlier version of the present article appeared in Project on Linguistic Analysis (POLA), series 2, No. 9, pp. 1–54 (Berkeley 1969).

[2] By 'voicing' we mean the [±voice] feature, namely the presence or absence of 'voice'.

concerned [PETERSON and LEHISTE, 1960; ZIMMERMAN and SAPON, 1958]. The question remains, however, whether this kind of durational differential is (a) primarily a matter of linguistic structure, that is a learned *language-specific* speech habit characteristic of the phonological system of English[3]; or (b) conditioned by an inherent physiological feature of articulation or, in other words, by a physio-acoustic constant governing the duration of vowels. If indeed the length variation is *intrinsic* in the sense of WANG and FILLMORE [1961] and some inherent physiological factors underlie such durational variabilities, the next logical question is: What are they?

One way to answer the first question, perhaps indirectly, is by investigating whether or not the voicing of a subsequent consonant exerts the same or similar shortening or lengthening effect on the preceding vowel in other languages. If the answer is yes, then the positive findings would suggest the existence of some inherent physiological basis to account for a language-universal phenomenon. If, on the other hand, the findings are negative, then we would be inclined to regard similar durational differential as an idiosyncratic speech habit of the English-speaking community. We hope, therefore, to gain a cross-linguistic view of the phenomenon and from there to be able to make some inferences as to the presence or absence of inherent articulatory factors (section 1), and to inquire into the physiological basis, if any, underlying such phenomenon (section 2).

1. A Cross-Linguistic View of Vowel Length Variation

In order to gain a cross-linguistic view of vowel length under environmental influence by a following voiced or voiceless consonant, aside from English three other languages were chosen as samples for experiment. They were French, Russian and Korean.

1.1 Experiment

The word-list of each of the four languages consisted of minimal or near-minimal pairs. Care was taken to assure identical prosodic patterns (pitch, stress, rhythm, etc.) for each member of the pair in order to eliminate variability in duration due to suprasegmental

[3] Such appears to be WANG's view. Cf. WANG and FILLMORE, 1961.

features. All test words actually occur; no artificial syllables were used. Each word was read by a native speaker six times (3 times in isolation, 3 times in alternation with its counterpart of the pair) except English, in which case only 3 tokens were recorded for each word. Recording, spectrograms and measurements were made of English only for the purpose of comparison and control. In every case only one subject was used. The recordings were made on magnetic tape and then submitted to analysis by various devices described below.

Phoneticians are painfully aware of the absence of a satisfactory procedure for breaking down the continuous flow of speech into discrete linguistic units. 'A transcription of speech on the structural level as a sequence of phonemes is not upset by the fact that the speech wave stimuli contributing to the listener's identification of any phoneme may be spread out over several successive minimal units of the speech wave. The only trouble occurs if investigators insist on measuring or defining phoneme durations' [Fant, 1960: 208.] Since this paper has to do precisely with measuring phoneme durations, not surprisingly we ran into considerable trouble of both theoretical and practical nature. If complete accuracy and distinctness of segmentation is unattainable at the present stage, we shall at least aim at consistency in deciding the phonemic boundaries.

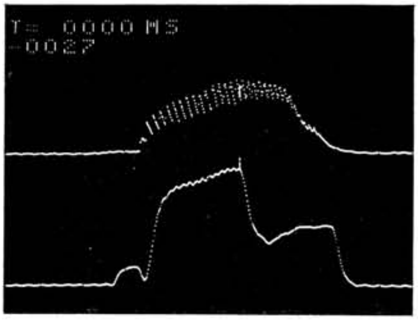
The theoretical basis of speech segmentation is laid down in the opening paragraph of Fant's classic work: 'The speech wave is the response of the vocal tract filter systems to one or more sound sources. This rule, expressed in the terminology of acoustic and electrical engineering, implies that the speech wave may be uniquely specified in terms of source and filter characteristics' [1960: 15]. Similarly, segmental boundaries can be specified in terms of temporally localized changes in the two basic categories of sound production, the source and the filter. These changes are: (a) change of the type of source (voice/noise); (b) intensity of the source (amplitude curve); (c) rapid change in the vocal tract filter function (formant transition); (d) simultaneous changes in both filter and source [cf. Fant, 1960: 207].

The above-mentioned cues of segmentation can be detected and displayed with varying degrees of effectiveness and accuracy by means of various types of instruments. In the experiments performed in relation to this paper, measurements were obtained from three kinds of graphic representation of speech sounds: oscillogram, screen display of the 'Pitch-program' on Linc-8, and spectrogram.

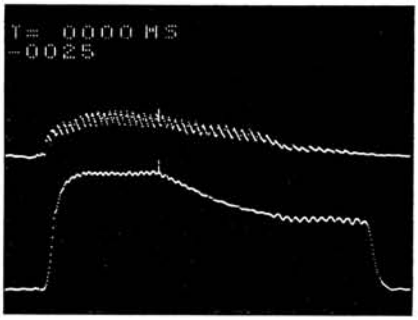
1.1.1 Oscillograms obtained from standard ink-jet Oscillomink model E show audio curve and pitch (F_0). Where alternation of source function (e.g. vowels flanked by voiceless segments) or drastic variation of pitch (e.g. voiced obstruents intervocalically) coincide with segmental boundaries, vowel duration may be measured with a high degree of accuracy. Oscillograms were used to determine the vowel durations tabulated in tables IX and X. In this case vowels were measured from the beginning to the end of the pitch contour, i.e. from the point where F_0 rises above, to the point where F_0 drops to the base-line. Initial and final stops were measured from the beginning of visible audio curve to the beginning of the vowel and from the end of the vowel to the end of visible audio curve respectively.

1.1.2 We have also explored the possibility of using the 'Pitch-program' as a segmentation technique. The 'PITCH' is a program designed by R. KRONES of the Phonology Laboratory, University of California at Berkeley, for a Linc-8 computer. The computer accepts analog voltage from the pitch extractor proportional to the fundamental frequency of a spoken utterance played from tape. The computer also samples voltage indicating the audio amplitude of the utterance. These voltages are displayed as a function of time on the scope screen. Frequency and time may be measured by moving a marker displayed on the screen (called a cursor) to any point of F_0 or amplitude curve [cf. KRONES, 1969]. Permanent records of screen display may be obtained by means of a polaroid camera (fig. 1).

The Pitch-program incorporates the advantages of oscillograph plus the convenience of numerical display for frequency (in Hz) and duration (in msec) on the scope screen of the computer itself. Since vowels of all speech sounds characteristically carry the greatest amplitude [cf. SACIA and BECK, 1926], theoretically vowel boundaries ought to be specifiable in terms of the change in the source intensity. In this regard the instantaneous amplitude display as a function of time by means of the Pitch-program has a certain advantage over the conventional broad-band spectrogram which marks intensity by means of degrees of darkness. However, the change of source intensity is often difficult to localize, and in such cases segmentation remains somewhat impressionistic and arbitrary. One serious limitation which the Pitchprogram shares with Oscillograms is its inability to show rapid changes of the vocal tract transfer function during speech production.

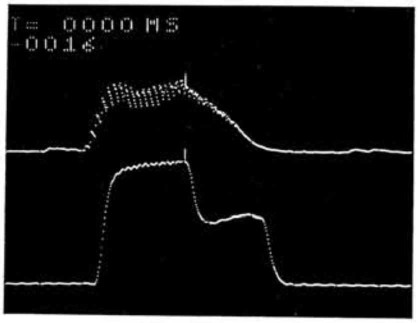


a

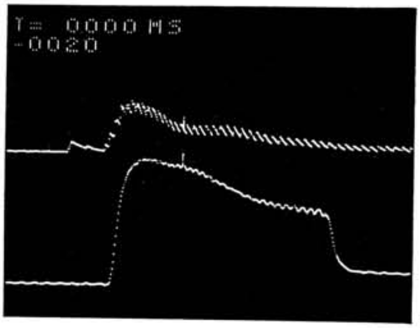


b

Fig. 1a. /bænk/.
Fig. 1b. /bæŋ/.

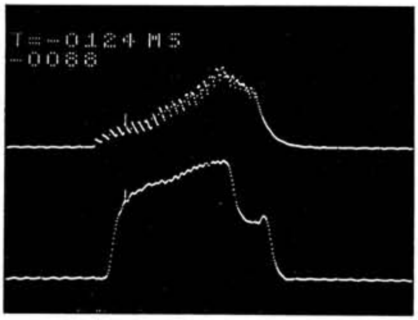


c

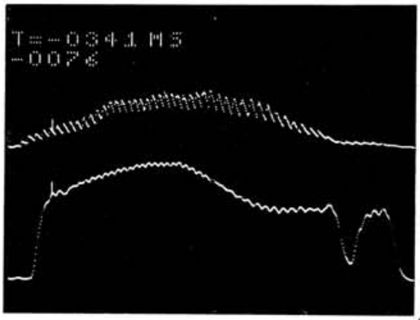


d

Fig. 1c. /kilt/.
Fig. 1d. /kild/.



e



f

Fig. 1e. /læp/.
Fig. 1f. /læb/.

Note: upper curve = amplitude, lower curve = F_0, cursor indicates suggested points of segmentation.

1.1.3 All considered, conventional spectrography still provides us with the most complete acoustic record of speech sounds. It is most effective in showing rapid changes in the vocal tract filter function, an acoustic correlate of the changes of the shape of the articulatory apparatus. Change of the type of source function is clear in most cases, though overlapping areas of periodic striation and random noise patterns occur. At best approximate F_0, an information not crucial for our purposes, can be obtained from narrow-band spectrograms[4].

For the sake of consistency, all measurements – except for the data used in section 2.4 – were made on a total of 376 broad-band spectrograms (300 Hz band-pass filter) produced by a Kay Sonagraph 6061-A. Oscillograms and the Pitch-program were used as subsidiary segmentation techniques.

Our criteria for segmentation based on spectrograms agree substantially with those established by FANT [1960: 21–24; 207–208] and PETERSON and LEHISTE [1960: 694–698]. We need only to point out some minor differences and certain preferences where FANT or PETERSON and LEHISTE offer alternate points of segmentation:

(a) The beginning and end of vowels flanked by stops are the beginning and end, respectively, of periodic striation in all formants. We regard the aspiratory interval as part of the preceding plosives [cf. FANT, 1960: 208 and PETERSON and LEHISTE, 1960: 694].

(b) Visible reduction of spectral intensity at about 500 Hz and, more generally, the diffuse spectral components in all frequency regions, serve as the primary cues of nasal segments[5]. These cues were found to be more reliable and more easily discernible than the onglide movements of formants regarded by PETERSON and LEHISTE [1960: 695] as primary cues of nasal boundaries.

(c) The vowel/liquid or liquid/vowel sequences present considerable problem. Typically a certain increase of intensity signals the transition from an initial liquid to the following vowel. In the case of the French /R/ followed by vowels, the vowels were measured from the onset of voicing in the lower frequencies although noise components continue in the higher region of the spectrum for 30–50 msec. Bounda-

[4] For the purpose of measuring F_0 oscillograms and scope screen displays of the Pitch-program are obviously superior.

[5] The reduction of spectral intensity in the region of 500 Hz is interpreted as anti-resonance due to the shunting effect of the nasal vacity system as a side channel to the vocal tract. The generally weak and diffuse spectral components are ascribed to the sound emitted from the nostrils [FANT, 1960: 149].

ries of liquids in postvocalic positions are even more difficult to determine. Formant movements in such cases are very smooth, and only approximate and sometimes somewhat arbitrary boundaries can be ascertained.

1.2 Results

The averages of vowel length of each test word are given in tables I–IV. The vowels in question are in italics. Each figure represents the average of 6 tokens (3 in the case of English). Table V summarizes tables I–IV and tabulates the variations, mean differences and ratios of vowel duration before voiced and voiceless consonants in English, French, Russian and Korean.

1.3 Tentative Conclusions

As attested by our data presented in the foregoing section (1.2), in all four languages studied a vowel is invariably longer before a voiced consonant than before an unvoiced one. Data based on only four languages (three of them being Indo-European) probably do not lend a sufficient basis for making sweeping generalizations. On the other hand, such invariable variations of vowel duration depending on the voicing of the following consonant can hardly be regarded as accidental. In fact, if we look beyond our own data and examine some of the evidences furnished by published sources, we would see that similar phenomenon is observed in a number of other languages as well (see table VI and fig. 2). The agreement of a number of widely divergent languages seems to bespeak a certain cross-linguistic validity of the statement that, ceteris paribus, vowels generally tend to become longer before voiced consonants and shorter before voiceless ones[6].

[6] This conclusion is further strengthened by similar data regarding intrinsic durational variations based on nonsense syllables spoken by native speakers of Chinese, Russian and German. Cf. Mohr, 1969. With regard to our Russian data, Miss Larissa Kukolshchikova (of Leningrad State University) remarked (through correspondence) that in Russian voiced obstruents are devoiced in word-final positions, so that /gleb/, /glub'/, etc. in our Table III are phonetically [glep] [glup'], etc. This fact does not necessarily invalidate our data; on the contrary, it is interesting to observe that the primary feature of voicing of the final stops in the underlying represent notions of /gleb/, etc. remains under the guise of the secondary feature of lengthening of the preceding vowel, even though voicing itself is absent phonetically. This is an instance of a secondary feature replacing a primary feature as the major distinguisher. Cf. Wang's (1968) comment quoted in footnote 13 of this paper.

Table I. Vowel length variations before [± voice] consonants. Each figure represents the average of 3 tokens

	English			
(1)	/læp/	155 ms	/læb/	300 ms
(2)	/fæt/	170	/fæd/	308
(3)	/læk/	158	/læg/	357
(4)	/æmpl/	148	/æmbl/	147
(5)	/sent/	123	/send/	147
(6)	/kilt/	102	/kild/	120
(7)	/bænk/	165	/bæng/	288

Table II. Vowel length variations before [± voice] consonants. Each figure represents the average of 6 tokens

	French			
(1)	/pɔ̃p/	400 ms	/bɔ̃b/	470 ms
(2)	/lapœR/	117	/labœR/	140
(3)	/gRat/	396	/gRad/	415
(4)	/lak/	451	/vag/	473
(5)	/butik/	304	/fatig/	305
(6)	/bɑ̃k/	453	/sɑ̃g/	520
(7)	/sa fwa/	246	/sa vwa/	310
(8)	/masif/	314	/masiv/	420
(9)	/dus/	399	/duz/	496
(10)	/laʃ/	459	/laz/	523

Table III. Vowel length variations before [± voice] consonants. Each figure represents the average of 6 tokens

Russian

(1)	/p*a*pa/	177 ms	/b*a*ba/	193 ms
(2)	/*i*tu/	79	/*i*du/	98
(3)	/b*a*ka/	91	/n*a*ga/	109
(4)	/st*e*p'/	152	/gl*e*b/	176
(5)	/gl*u*p/	120	/gl*u*b'/	123
(6)	/k*o*t/	170	/g*o*d/	227
(7)	/l'*o*t/	187	/l'*o*d/	195
(8)	/l*u*k/	103	/l*u*g/	139
(9)	/*u*trat'/	69	/*u*drat'/	101
(10)	/r*a*sa/	152	/r*a*za/	203
(11)	/vn*e*šnyj/	144	/n*e*žnyj/	195

Table IV. Vowel length variations before [± voice] consonants. Each figure represents the average of 6 tokens

Korean

(1)	/*ɔ*p^{h}ɔ/	92 ms	/*ə*bɔ/	139 ms
(2)	/k*ɯ*p^{h}al/	70	/k*ɯ*bal/	74
(3)	/*i*t^{ʔ}al/	86	*i*dal/	109
(4)	/k*a*t^{h}a/	81	/k*a*da/	95
(5)	/*a*ka/	109	/*a*ga/	137
(6)	/tʃ*ə*k^{h}ɔŋ/	127	/tʃ*ə*gɔŋ/	160
(7)	/k*a*tʃ^{h}i/	75	/k*a*dzi/	117

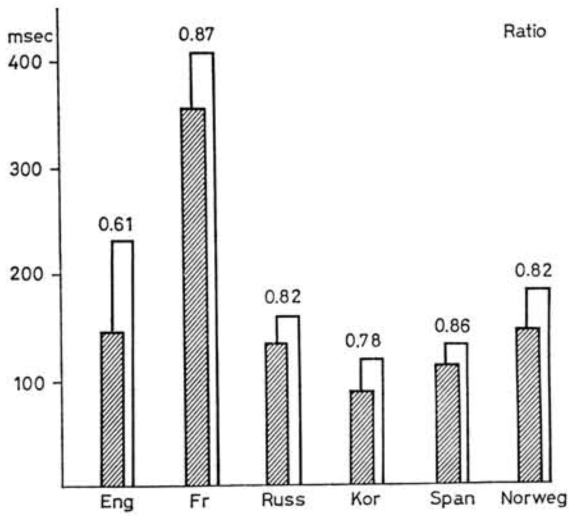


Fig. 2. Ratio of vowel duration before voiced consonants (white) to vowel duration before voiceless consonants (shaded).

Table V. A cross-linguistic view of vowel length variation as a function of the voicing of the following consonants

	Vowel length in msec			
	Before voiceless consonants	Before voiced consonants	Mean difference	Ratio
English	146	238	92	0.61
French	354	407	53	0.87
Russian	131	160	29	0.82
Korean	91	119	28	0.78

However, I must hasten to add that the voicing of the adjacent consonant influences its preceding vowel to different degrees in different languages. Vowel duration in English varies rather drastically depending on the following consonant, the vowel before a voiceless consonant being only 0.61 or less than 2/3, of its counterpart before a voiced consonant. This does not hold true in other cases. In French, for instance, the ratio is 0.87; in other words, the shorter vowel is only 13% less than its longer counterpart. And if we accept the data

Table VI. A cross-linguistic view of vowel length variation as a function of the voicing of the following consonants – additional data from published sources:

(1) = Peterson and Lehiste, 1960
(2) = Zimmerman and Sapon, 1958
(3) = House and Fairbanks, 1953
(4) = E. A. Meyer, Englische Lautdauer – Uppsala 1903, report in Fintoft, 1961
(5) = Zimmerman and Sapon, 1958
(6) = Fintoft, 1961

	Vowel length in msec			
	Before voiceless consonants	Before voiced consonants	Mean difference	Ratio
English:				
P–L (1)	197	297	100	0.66
Z–S (2)	145	228	83	0.64
H–F (3)	174	253	79	0.69
German (4)				0.90
Spanish (5)	109	127	18	0.86
Norwegian (6)	148	181	33	0.82

concerning Spanish [reported by Zimmerman and Sapon, 1958], the difference is reduced to a mere 18.2 msec.

In view of the above observations, we may tentatively conclude that (a) it is presumably a *language-universal* phenomenon that vowel duration varies as a function of the voicing of the following consonant, and (b) the extent, however, to which an adjacent voiced or voiceless consonant affects its preceding vowel durationwise is determined by the *language-specific* phonological structure.

2. *Intrinsic Factors of Vowel Length Variation*

Based on our findings and available literature on the subject, we have good reasons to assume as language-universal the variability of vocalic duration as a function of the [$\pm$voice] feature of the following consonant. This substantive phonological universal leads us to believe that some common inherent articulatory factor(s) must underlie the widely observable durational differential under the described environmental influence. The nature of such inherent factor is not immediately apparent. Some explanations have been offered by various authors. We shall discuss some of them.

2.1 Articulatory Distance

O. Jespersen suggested that 'the duration of a vowel is function of its articulatory "distance" to the adjacent consonants' [quoted by Lindblom, 1967: 22]. While Jespersen's theory of articulatory distance or a special version thereof provides a plausible explanation for the vowel length variation as a function of mandibular distance [Lindblom, 1967: 22], it does not seem to apply here. Strictly speaking the voiced/voiceless dichotomy does not involve difference in articulatory distance: the target position of buccal articulators is presumably identical for the voiced and the voiceless members of the consonantal pair, and the presence or absence of voice is produced by a separate articulator altogether.

In a transformed sense, however, assuming that vowels are normally voiced, vowels would be one degree more 'distant'[7] from a voiceless consonant than from a voiced one, because the transition from vowel to a voiceless consonant involves the distinction in one additional aspect of articulation, namely in the discontinuance of voicing. If our reasoning is correct, we would expect the vowel to be lengthened before a more distant voiceless consonant. But the contrary is true.

2.2 Articulatory Energy Expenditure

S. Belasco suggests that there is a cause and effect relationship between vowel duration and the voicing of a following consonant and attempts to explain the causal relationship by the 'force of articulation'. Belasco maintains that vowel duration varies inversely with the degree of physiological energy required to pronounce the following consonant. Thus, 'the anticipation of a consonant requiring a "strong" force of articulation will tend to shorten the preceding vowel since more of the total energy needed to produce the syllable is concentrated in the consonant. The opposite is true of course when the consonant has a weak force of articulation' [Belasco, 1953; see also accounts by Fintoft, 1961: 26, and Zimmerman and Sapon, 1958: 153].

Belasco's view resembles to some extent the 'energy expenditure' theory proposed by E. A. Meyer, according to whom, 'the temporal

[7] Obviously 'distant' is used here in a transformed sense of degrees of differentiation in terms of distinctive features understood as separately controllable aspects of articulation.

organization of speech sounds is determined by the amount of physiological energy that is consumed in producing them' [LINDBLOM, 1967: 22].

Implied in BELASCO and MEYER's hypothesis is the assumption that each syllabic unit would take relatively constant amount of energy to produce, so that the energy expended on the vowel would vary inversely as the energy consumed by the subsequent consonant. This is an interesting idea that can be tested [cf. KOZHEVNIKOV and CHISTOVICH, 1967, esp. chapter 3].

In order to substantiate the theory of 'force of articulation', it would be necessary to have accurate figures of speech organ muscle potentials (e.g. by means of electromyographic recordings) and to rank these figures against the relative durations of the vowels of the corresponding syllabic units. If BELASCO had documented such data in support of his hypothesis, he did not present it in his cited article. In point of fact, experimental evidence, if any, seems to disprove such a hypothesis. ZIMMERMAN and SAPON [1958] reports that 'an attempt to rank the duration of the vowels in accord with the notion of force of articulation as suggested by BELASCO yielded negative results'.

2.3 Perceptual Distance

Viewed from the perception end, the variation of vowel length in inverse proportion to the closure time of the following consonant might reflect the speaker's tendency to maximize the perceptual distance between voiced and voiceless consonants by maximizing the difference in the ratio between the duration of a vowel and that of its neighbouring consonant. The simultaneous reduction of vowel duration and the prolongation of the closure for the following voiceless stop would heighten the perceptual effect of voicelessness. Conversely, the simultaneous lengthening of vowel and shortening of the consonantal closure would concur as acoustic cues of voicing [LISKER, 1957]. This hypothesis is not altogether implausible in the light of the following observation by P. DENES [1955: 763]:

'... when, however, the perception of "voicing" is plotted as a function of the *ratio* (my emphasis) of the duration of consonant and vowel a consistent, single curve results and the scatter of the points about this curve is small ... It can be seen that the perception of "voicing" of the final consonant increases as the ratio of the durations of final consonant to preceding vowel decreases.'

As we suggested in section 1.3, whereas the vowel length duration is a language-universal phenomenon, the extent to which the vowel duration varies is determined by language-specific phonological characteristics. The remarkable durational differential in English vowels (mean difference: 92 msec; ratio: 0.61) may be seen as a perceptual device serving as distinctive function in the phonological system of the English language.

The situation is entirely different with regard to the other tested languages. The data concerning these languages exhibit a difference in both ratio and absolute values of vowel duration that border on the DL (difference limen; also referred to as JND or 'just noticeable difference') of human perception. Various experiments have been performed on the DL of duration by Stott [1935], Henry [1948] and, more recently, by Ruhm *et al.* [1966] and Huggins [1968]. The Weber ratios and absolute DLs established in the recent studies are much smaller than were found in earlier experiments. Ruhm *et al.*, for instance, reports a Δ T/T ratio of 0.0260 or an absolute DL of 2.6 msec for a reference duration of 100 msec. Henry on the other hand established the Δ T/T ratio and the absolute DL to be 0.196 and 21.56 msec respectively, for a comparable reference duration of 110 msec. Regarding the obvious discrepancy in the DLs established by earlier investigators as against those determined by more recent experiments Lehiste [1969: 18] comments:

'The research technique employed in the more recent study [Ruhm *et al.*] was probably more conducive to testing the limit of the auditory sensitivity of the subjects. It is quite likely that in a speech situation, where a great amount of external noise is present, the perception would not be as acute. It is perhaps a reasonable assumption that the difference limens established by Ruhm *et al.* represent the limit of perceptibility under optimal conditions, whereas it appears likely that in a speech condition, the just noticeable differences established by Henry and Stott may apply.'

Assuming then the figures established by Stott and Henry to be the standards applicable to normal speech conditions, taking into account the not always favourable signal-to-noise ratios and other interferences, we have combined the findings of Stott and Henry into a composite table (table VII) showing the Weber ratios (column B)

Table VII. Comparison of mean differences of vowel duration in Korean, Spanish, Russian, Norwegian, English and French, with Weber ratios and absolute DLs for comparable reference durations. (H) = HENRY; (S) = STOTT

A Reference duration in msec	B Weber ratio $\Delta T/T$	C Absolute DL in msec	D Mean difference in msec	E Av. vowel duration in msec	
110 (H)	0.196	21.56	28	105	Korean
			18	118	Spanish
			29	145	Russian
175 (H)	0.188	32.90	33	165	Norwegian
200 (S)	0.142	28.4	92	192	English
400 (S)	0.120	48.0	53	380	French

and DLs (column C) for reference durations ranging from various languages are arranged according to the average vowel duration (column E)[8], which can be regarded as equivalent to the reference duration in the STOTT and HENRY experiments. The mean difference of vowel duration as a function of the voicing of the following consonant is reproduced (from tables V and VI) in column D.

When we compare columns C and D we see that, aside from the conspicuous exception of English the mean difference of vowel duration falls below the DL (as in the cases of Spanish and Russian) or hovers just above the DL for the nearest comparable reference durations (Norwegian, Korean and French). Alternately, we may plot on a semilogarithmic graph paper HENRY and STOTT's DLs as a function of reference duration (fig. 3) and compare the mean difference of vowel of various attested languages. The resulting graph shows that Spanish and French fall below – and Russian coincides with – the absolute DL curve extrapolated from HENRY's data, while Korean and Norwegian barely rise above it. English, again, constitutes a clear exception. The significance of the relatively small mean differences is further reduced by the consideration that in normal speech perception of voiced or voiceless consonants using vowel duration as secondary acoustic cue is an absolute and not a discriminatory judgment; consequently the absolute DL in the perception of vowel duration

[8] The average vowel duration in column E of table VII represent the mean vowel durations before voiced and voiceless consonants indicated in tables V and VI.

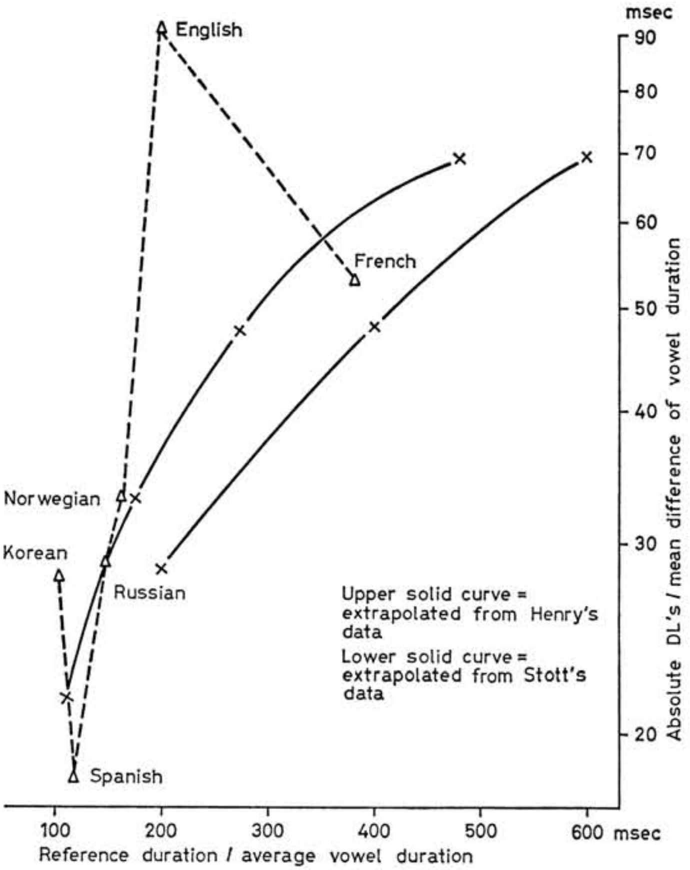


Fig. 3. DL's and mean differences of vowel duration in various languages (connected by broken line).

Table VIII. Variations of average time of closure for voiced and voiceless consonants

Language	Average closure time in msec			
	for voiceless consonants	for voiced consonants	mean difference	ratio
English	140	88	52	0.63
Korean	124	54	70	0.45

NB. In order not to confuse issues, the figures represent, in the case of English, the averages of closure times for *stops only*. The intervening nasals and /l/ in such pairs as /sent/ vs. /send/ might complicate the issue.

would be not the mean difference between vowels preceding voiced consonants and vowels before voiceless consonants, but rather the standard deviation from the average vowel length. This would reduce the durational differential by a factor of 2 to 1. In this sense the variation of vowel length in all languages cited (except English) would fall way below the absolute DLs. In the light of the above evidences, perceptual distance does not seem to be a satisfactory explanation for the intrinsic durational differential of vowels[9].

2.4 Compensatory Temporal Adjustment

Our data show that the occlusion time before the release of voiceless stops is appreciably longer than that before the release of voiced stops. In English, for instance, the ratio of the shorter closure for voiced stops to the longer closure for voiceless stops is approximately 2:3. The difference is even more striking in Korean, for which the ratio is less than 1:2 (see table VIII). If vowel duration were kept constant, the (C)VC syllable would vary considerably in time depending on the presence or absence of voicing in the final consonant. However, the same data also show that the vowel duration varies inversely with the closure time of the following consonant. Thus the longer the closure time for voiceless consonant, the shorter the vowel length, and vice versa[10]. One might wonder, therefore, if a compensatory temporal reorganization of sequential motor commands might not be at work here to prevent the time differential of syllabic units from becoming too drastic and to assure a relatively even flow of syllables. Thus Kozhevnikov and Chistovich [1967: 107] consider the variability of vowel duration in question as 'the assumed effect of the compensation of differences in the duration of closure by differences in the duration of the adjacent vowels; however, this compensation was not full'. An analogous situation is described by B. Lindblom. In his recent article [1967: 21] Lindblom writes:

[9] A qualification here may be necessary. Although the vowel length variation *per se* does not seem pronounced enough as a perceptual cue; in conjunction with other concommitant cues, however, the intrinsic variation of vowel length may be a contributing factor in the listener's discrimination between voiced and voiceless consonants.

[10] Kozhevnikov and Chistovich, 1967, esp. ch. 3, contains similar data regarding Russian.

Table IX. Segmental durations of $/pai(\begin{Bmatrix}p\\k\end{Bmatrix}(t))/$

	Initial stop	Vowel	Final stop(s)	Total duration
/pai/	31	378	–	409
/paip/	36	152	155	343
/paik/	33	137	171	341
/paipt/	29	149	243	421
/paikt/	30	129	230	389 msec

Each figure represents the average of 9 tokens except the case of /paip/, where only 8 tokens were measured (the 9th being rejected as dubious).

'The primary reason why open vowels tend to be longer than close vowels is that the extent of mandibular movement is larger for open vowels even during the occlusion of the initial consonant and that the mandibular offglide movement from the vowel into the final consonant progresses so slowly that the contact between the articulators for the consonant is delayed in the context of the open vowel ... However, there is some evidence ... that the lip gestures proper can be reorganized in such a way so as to compensate for the contribution of the jaw movement to the opening of the lips for the initial [b].'

And on p. 26:

'Comparing the words with open and close vowel pairs it is seen that for some of the utterance paires, there is reorganization of the "labio-muscular" components of lip separation x_u–x_l. The effect of these compensatory adjustment is to prevent the open vowels from becoming even longer[11].'

The assumption underlying the above hypothesis is that the duration per syllabic unit is relatively constant and that the vowel length varies inversely as the following consonant. In order to test this hypothesis, we carried out the following experiment: magnetic tape recordings of utterances of the form $[pai(\begin{Bmatrix}p\\k\end{Bmatrix}(t))]$ were analyzed by means of an Oscillomink E which displays both pitch contours above 40 Hz as well as oscillograms. Measurements of segmental durations were made and average length of each segment was obtained (see table IX).

[11] In relation to the sequential organization of utterances on a higher level than phonemes cf. MACNEILAGE's [1968] article on the serial ordering of speech sounds.

Table X. Contrastive lengths of syllables with single and double final stops

	Initial stop	Vowel	Final stop(s)	Total duration
/pai$\begin{Bmatrix}p\\k\end{Bmatrix}$/	35	144	163	342
/pai$\begin{Bmatrix}p\\k\end{Bmatrix}$t/	30	139	236	405
Mean difference	5	5	73	63
Ratio	0.86	0.97	0.69	0.84

In order for the hypothesis of compensatory temporal adjustment to work, first the total duration of syllable units must be relatively constant and, secondly, the final stops must be lengthened at the expense, so to speak, of the contiguous vowel, or vice versa. Neither supposition is borne out by our data. To begin with, if we compare [pai$\begin{Bmatrix}p\\k\end{Bmatrix}$] and [pai$\begin{Bmatrix}p\\k\end{Bmatrix}$t], there is a statistically significant variability of the order of 63 msec in the total syllabic duration. More importantly, the vowel duration remains constant (exhibiting a mean difference of mere 5 msec) while the final double stops [$\begin{Bmatrix}p\\k\end{Bmatrix}$t] is considerably stretched out (mean difference: 73 msec; see table X). These data, though limited, seem to argue against compensatory temporal adjustment as a hypothetical explanation of the vowel durational differential.

2.5 Laryngeal Adjustment

In *The sound pattern of English* [1968: 301] Chomsky and Halle propose vocal cord adjustment rate as an explanation for the lengthening of vowels before voiced obstruents. This hypothesis is based on an earlier study by Halle and Stevens [1967] on the mechanism of glottal vibration. Essentially the hypothesis put forward by Halle and Stevens is this: Vocal cords may vibrate in two ways: (a) spontaneously, in response to unimpeded air flow, as during vowel production; (b) nonspontaneously, when there is a radical constriction or total closure as during voiced obstruents. Oral constriction or closure causes the supraglottal pressure behind the constriction point to build up and consequently reduce the pressure drop across the glottis during phonation. In order to maintain continuous vocal cord vibration in the face of reduced pressure drop across the glottis, glottal opening must be widened. HALLE and STEVENS [269] reports that 'the vocal cords are separated so that the glottis remains open during the entire vibratory cycle'. This configuration of the larynx involves a delicate and precise adjustment of the position of the vocal cords and can be achieved only rather slowly. The longer laryngeal adjustment time required for the nonspontaneous voicing of a subsequent voiced consonant would then necessitate an increased duration of the preceding vowel.

Logically, HALLE and STEVENS' theory appears to be rather attractive. Since we have isolated the [±voice] feature of the consonantal environment as the conditioning factor of vowel length variability, it is reasonable to expect that the physiological explanation of such durational differential would probably lie in the articulatory mechanism involved in the voicing of the consonantal environment. Now the mechanism primarily and directly responsible for voicing is obviously the larynx. Thus, a priori, the laryngeal mechanism seems a most likely place to find an answer to the intrinsic variation of vowel duration.

One way of testing the validity of the laryngeal adjustment theory is this: if laryngeal adjustment rate is indeed responsible for the vowel length variation, then in a vowel-sonorant-obstruent string, we would expect the sonorant alone to vary in length, the vowel remaining constant. The obvious reason is that in such a sequence the vowel is separated from the [±voice] obstruent and is, therefore, shielded from the immediate effect of the laryngeal adjustment which takes place between the sonorant and the obstruent. We put this prediction based on HALLE and STEVENS' theory to experimental verification. The experiment involved a test list of 32 words (16 pairs) of the form: vowel-sonorant-consonant. Three recordings were made of each of the 32 words. 96 spectrograms were made from these recordings, and measurements were obtained and averaged through the customary procedures. The results are tabulated in table XI and summarized in table XII.

Apparently, our prediction based on HALLE and STEVENS' hypothesis was not borne out by facts. Contrary to our expectation, the

Table XI. Duration of vowels and sonorants before [± vcd] consonants

	Duration in msec		
	Vowel	Sonorant	Total
	(A) vowel + nasal sequences		
lumper	92	56	148
lumber	118	102	220
sent	218	51	269
send	245	133	378
linker	102	37	139
linger	119	94	213
cinch	156	33	189
singe	192	142	334
hence	179	38	217
hens	213	111	324
	(B) Vowel + [l] sequences		
help	52	128	180
helb*	106	229	335
kilt	76	134	210
killed	105	231	336
bulk	59	168	227
bulg*	95	214	309
belcher	39	157	196
Belgian	55	175	230
false	97	203	300
falls	125	251	376
	(C) vowel + [r] sequences		
harp	73	124	197
harb*	106	189	295
cart	118	138	256
card	157	169	326
Burke	83	175	258
berg	105	271	375
perch	107	124	231
purge	141	186	327
surf	53	164	217
serve	61	194	355
course	107	131	238
cores	135	192	327

NB: starred (*) words = artificial syllables.

Table XII. Summary of data on the duration of vowels and sonorants before [± vcd] consonants

Following consonant	Duration in msec			Ratio		
	Vwl	Nas	Tot	Vwl	Nas	Tot
___ [–vcd]	129	43	172	0.73	0.47	0.59
___ [+vcd]	177	116	239			
	Vwl	[l]	Tot	Vwl	[l]	Tot
___ [–vcd]	65	158	223	0.65	0.64	0.65
___ [+vcd]	99	246	345			
	Vwl	[r]	Tot	Vwl	[r]	Tot
___ [–vcd]	90	143	233	0.76	0.66	0.70
___ [+vcd]	118	217	335			
	All sonorants pooled:			All sonorants pooled:		
	Vwl	SON	Tot	Vwl	SON	Tot
___ [–vcd]	95	115	210	0.73	0.60	0.66
___ [+vcd]	131	193	324			

lengthening or shortening effect of the consonantal environment was not limited to the immediately preceding sonorant alone, but rather spread to the vowel segment as well; vowel duration varied notably even when separated from the obstruents by an intervening sonorant. As a matter of fact, the sequence of vowel-sonorant behaved as a unit durationwise. In other words, the voicing of the consonantal environment exercised durational influence on the vowel-sonorant sequences as a whole. Both vowels and sonorants varied by very similar proportions. This finding seems to weaken somewhat the explanatory power of the laryngeal adjustment hypothesis[12].

Two related observations seem to cast some doubt on the hypothesis under discussion. From experience (cf. for instance table IX) we know that vowel is longest in open syllables, decreases slightly when followed by voiced consonants and is shortest before a voiceless obstruent. If the delicate adjustment of vocal cords accounts for the lengthening of vowels before voiced consonants, it does not seem to explain why vowels tend to prolong even followed by silence [cf. WANG, 1968]. 'We may further observe that', as WANG points out in the same article

[12] Admittedly, the data on vowel-sonorant sequences alone do not entirely rule out the plausibility of the laryngeal adjustment hypothesis. It is possible, for example, to argue that vowels and sonorants constitute homogeneous continua, and that the influence of the consonantal environment spans over the whole continua across segmental boundaries.

[1968: W35], 'vowel duration is affected in the same way by the "voicing" of the following consonant, even in whispered speech [SCHARF, 1964]. Since presumably voicing is not realized in the whispered consonant, it may be argued that the duration differential cannot be due to laryngeal adjustments[13]'.

The limited EMG (electromyographic) data we obtained through the Audio and Speech Research Laboratory, V. A. Hospital, San Francisco, provided us a more direct way of observing the activities of the laryngeal muscles in speech production. The word list for one particular EMG experiment included three pairs of monosyllabic words: 'tap/tab, tat/tad, tack/tag'[14]. The EMG printout recorded signals from four groups of laryngeal muscles: (1) posterior cricoid-arytenoid (PCA), (2) interarytenoid (IA), (3) thyroid-arytenoid (TA), and (4) cricoid-thyroid (CT) muscles, along with recordings of (5) sub-glottal pressure, (6) air flow, and (7) audio curve.

The critical information that concerns us most is of course the EMG signals of the PCA muscles which are directly responsible for the regulating of the opening of the glottis. If indeed, as HALLE and STEVENS argued, voiced consonants require a delicate adjustment in the width of the opening between the vocal cords, we would expect the PCA muscles to be active relatively long before a voiced consonant because of the necessary time for the delicate positioning of the vocal cords. On the other hand, we would expect the same muscle to show a higher but somewhat later EMG peak before a voiceless consonant, since for the voiceless consonant the glottis is wide open and relatively little time is needed for the simple abduction of the vocal cords. The EMG recordings we have fail to show any difference either in timing or in intensity of signals from the PCA muscles. In the absence of distinct EMG signals from the PCA muscles, we examined the

[13] WANG was careful to point out: 'The situation is more complicated, however. We know that many primary phonetic features are habitually accomplished by secondary features during normal speech: as the primary consonantal feature here (be it [voice], [tenseness], or whatever) is accomplished by the secondary vowel feature of [long], and as the primary tone feature is accompanied by the secondary amplitude feature (...). As we change from the normal mode into some mode of production, e.g. whispering, some primary features are lost. The secondary features remain, even when their causes are no longer present and replace the primary features as the major distinguishers. The situation is quite parallel to many historical instances of phonological change where a primary feature becomes lost and a secondary feature takes up its role' [WANG, 1968: W35].

[14] We want to express here our thanks to Dr. T. A. SHIPP, director of the Audio and Speech Research Laboratory, V. A. Hospital, San Francisco, for inserting these words designed for our purposes.

recordings for visible decrease of EMG signals for other groups of muscles responsible for the closure of the glottis. Such signs were not obvious either[15].

2.6 Rate of Closure Transition

Another hypothesis that deserves some thought is the suggestion that vowel duration variability might be a result of the different speeds of the transition from vowel to consonantal closure[16]. To simplify the matter, let us for the moment consider vowel+stop sequences. The VC sequences of this type may be represented as figure 4. We may hypothesize that A and A′ start at about the same time for (a) and (b) sequences. In other words, Tl and Tl′ are approximately equal. However, it is well known that a voiceless consonant is articulated with open glottis, whereas a voiced one is made closed glottis. As a result, the intraoral pressure during a voiced consonant closure is relatively low, since the pressure is built up by the air of the mouth cavity *alone;* in the case of a voiceless consonant occlusion, the intraoral pressure is considerably higher, since the volume of air of the mouth *and* lungs is increased. Take the /b, p/ pair for instance:

'As a consequence of these aerodynamic events, the pressure impulse (time integral of pressure) acting on the closed lips from inside the mouth is much greater during the voiceless [p] than during the voiced [b]. It is therefore reasonable to assume that the muscular effort needed in order to prevent the lips from being blown apart by the pressure at the implosion of a [p] should be greater than the corresponding effort for a [b]' [ÖHMAN, 1967: 47].

From the anticipatory effect of the differences of muscular effort in the closed position for voiced and voiceless consonants we may further infer that the transition from vowel to a voiceless consonant closure

[15] Considering the difficulty of positioning the electrodes in the laryngeal muscles, there was a shade of doubt regarding the accuracy of the EMG recordings of the PCA muscles in this particular experiment. The limited data presented here are not intended as definite proof. A more carefully controlled experiment with a greater number of samples will be necessary to ascertain the significance of our preliminary findings.

[16] This hypothesis was first suggested to me by PETER F. MACNEILAGE. Similar opinion has been expressed by KOZHEVNIKOV and CHISTOVICH [1967: 108]: 'It is natural to think that it is a side effect of the differences in the articulation of voiceless and voiced consonants. It seems to us that the cause of the effect is very elementary. The movement of the lip or tongue accomplished upon the closure of voiceless consonants are conducted with great force. Correspondingly the speed of movement in the case of voiceless consonants is greater and the moving organ reaches the final position (closure) in a lesser interval of time.'

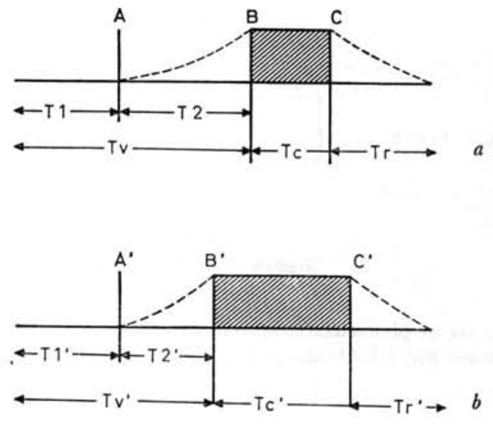


Fig. 4. Closure transition.

(a) = vowel + voiced stop
(b) = vowel + voiceless stop
A = starting point for movement toward closure
B = closure
C = release
T1 = steady-state portion of vowel
T2 = transition from vowel to consonantal closure
Tv = total vowel duration
Tc = closure time
Tr = transition from closure to open position
Prime ['] indicates voiceless counterparts.

(i.e. from A′ to B′) would be faster than the transition from vowel to a voiced consonant closure (i.e. from A to B). In other words, T2′ would be shorter than T2. This inference makes sense since F = ma (force = mass × acceleration) and consequently $a = \frac{F}{m}$. Now, supposing m (e.g. of lips) to be constant, a (of labial movement) would vary in direct proportion to F (muscular effort closing the lips). If the above reasoning is correct, the differential between Tv and Tv′ would be a function of the differential between T2 and T2′, which is, in turn, a function of the differential between muscular efforts for making closure B and B′[17].

[17] Incidentally, the same principle may also account for the variability in closure time of voiced and voiceless consonants (Tc and Tc′). The releasing motion of voiced stops requires a greater force than that of voiceless stops which is aided by the greater intra-oral pressure [cf. ÖHMAN, 1967: 47]. By the same force-acceleration principle, Tr would be shorter than Tr′. As an anticipatory effect of a longer Tr, release C begins earlier than C′, and Tc is accordingly shortened.

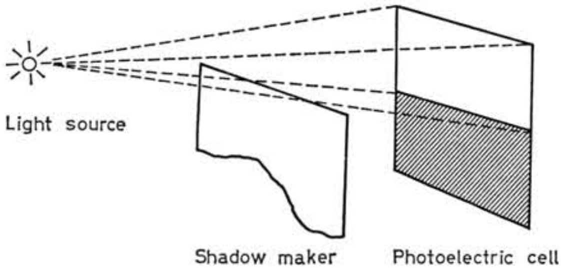


Fig. 5. Basic elements of photoelectric system for tracking the displacement of a light (attached to the lower lip) [cf. OHALA *et al.*, 1968: 137, fig. 2].

Recent photoelectric methods of transducing lip and jaw movements [cf. OHALA *et al.*, 1968] offer us a simple way of testing the closure rate as an explanation of vowel duration differential. The device for tracking lip movements consists of three basic elements: (a) light source: a tiny bulb attached to the lower lip of the speaker; (b) a 'shadow-maker'; and (c) a photocell (fig. 5). Vertical displacement of the light cast a greater or smaller shadow on the photocell causing it to develop a smaller or greater amount of voltage respectively. The varying voltages are fed into a 6-channel ink-jet Oscillomink E, which displays the curve of the vertical lip displacement as a function of time along with the corresponding audio curve (fig. 6). The speed of lower lip movement in articulating the final stops is specified as $\frac{\Delta D}{\Delta T}$ (T = time; D = degree of displacement) at the point of maximum velocity. By giving T a constant value (T = 1), we can represent the velocity of closing lip movement by just indicating the values of D (table XIII). The actual time of the lower lip moving toward bilabial closure was measured from the lowest point of vowel opening to the bilabial contact signaled by the audio curve (table XIV)[18].

Table XIII shows lip closure movement is typically faster for voiceless ($\frac{\Delta D}{\Delta T} = 0.68$) by a ratio of 0.82. The mean difference of time

[18] After the bilabial closure as indicated by the audio curve, the lower lip was shown to continue to rise, presumably as a result of the lagging mandibular movement [cf. LINDBLOM, 1968].

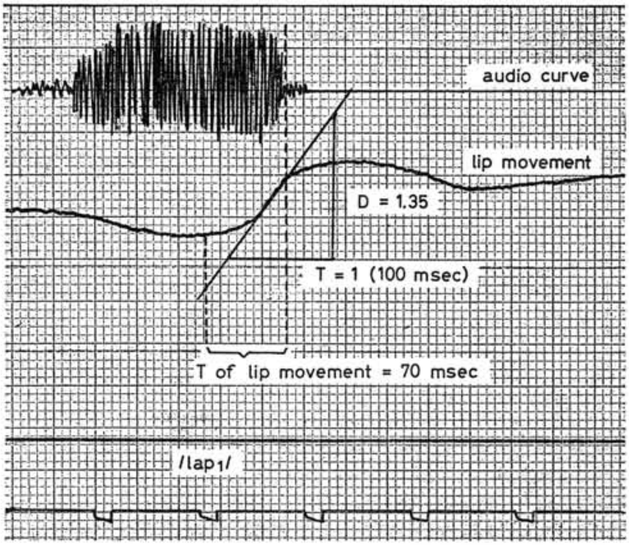


Fig. 6a. Sample of recordings of a photoelectric device for transducing lower lip movement toward consonantal closure /læp/. Note: the slope = line tangent to the curve at the point of maximum velocity.

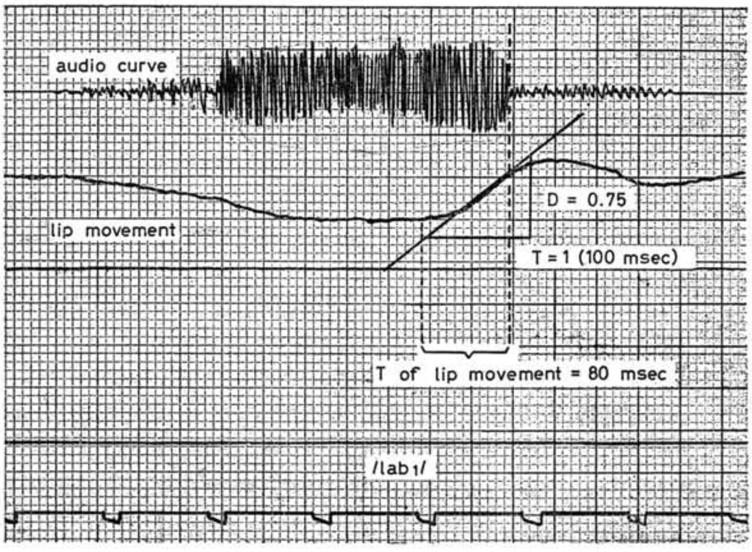


Fig. 6b. Sample of recordings of a photoelectric device for transducing lower lip movement toward consonantal closure /læb/.

Table XIII. Velocity $\left(=\frac{\Delta D}{\Delta T}\right)$ of lip movement toward final voiced/voiceless stop

Voiceless final stop		Voiced final stop	
lap	1.10	0.76	lab
nap	0.87	0.82	nab
cup	1.00	0.56	cub
pup	0.82	0.68	pub
rip	0.47	0.56	rib
bop	0.98	0.81	Bob
av.	0.87	0.69	

Table XIV. Time (in msec) of lower lip movement toward voiced/voiceless final stop

Voiceless final stop		Voiced final stop	
lap	63	80	lab
nap	63	82	nab
cup	55	88	cub
pup	57	108	pub
rip	52	58	rib
bip	58	102	bib
bop	78	97	Bob
av.	61	88	

needed for making the voiced and voiceless bilabial stops is 27 msec[19]. A mean difference in the order of 27 msec correlates very closely with the mean difference in vowel duration for Spanish, Korean, Russian and Norwegian (18, 28, 29 and 33 msec respectively), but less well for French (53 msec). As for English, the far greater variation (92 msec) may be accounted for by the added perceptual function of vowel length in the listener's differentiation of voiced vs. voiceless consonants.

EMG data seem to lend some further support to the hypothesis of closure transition as the cause of vowel differential: 'The EMG peaks

[19] In an experiment performed by KOZHEVNIKOV and CHISTOVICH [1967: 182], measurements obtained from continuous photooscillographic recordings of lip movements yielded the following results: the velocity of the transition from vowel opening to lip closure was approximately 125 and 140 mm/sec for voiced and voiceless occlusives respectively. These figures represent utterances in 'full style', i.e. in citation forms; the corresponding figures for 'conversational style' are approx. 90 and 110 mm/sec respectively.

associated with the implosion of the [p] are in general of 20% greater amplitude than the implosion peaks of the [b]' [Öhman, 1967: 38, fig. I-B-18; cf. MacNeilage, 1968: 11].

We do not have as yet supporting data regarding occlusives other than bilabial stops, and none for fricatives and affricates. Nevertheless, based on the data available to date, the rate of closure transition seems to be the best, if partial, explanation of vowel length variability.

Summary

Measurements obtained from 376 spectrograms taken from recordings of four languages led us to conclude that it is a language-universal that vowel duration varies as a function of the voicing of the following consonant. The assumed language-universal called for some inherent articulatory factor. We subsequently discussed some of the hypothetical explanations regarding the underlying physiological factors affecting vowel duration under conditions described. Evidences based on original experiments or published data were used to evaluate the claims by various authors. The hypothesis of closure transition seems to be, at the present stage, the best, if partial, explanation of vowel length variation.

Zusammenfassung

Die Vokaldauer als Funktion der Stimmhaftigkeit der umgebenden Konsonanten

376 Spektrogramme, von Bandaufnahmen vier verschiedener Sprachen hergestellt, wurden ausgemessen. Aus diesen Messungen folgt, daß Schwankungen der Vokaldauer auf Grund der Stimmhaftigkeit oder Stimmlosigkeit des nachfolgenden Konsonanten eine allgemeingültige Eigenschaft der Sprache ist. Eine solche Spracheigenschaft verlangt eine inhärente Artikulationsursache zur Erklärung. Einige frühere Hypothesen über verschiedene physiologische Ursachen, die die Vokaldauer unter den gegebenen Umständen beeinflussen könnten, werden besprochen. Beweismaterial aus eigenen Versuchen und aus anderen Veröffentlichungen wird zur Bewertung dieser Erklärungsversuche herangezogen. Die Annahme einer Lautverschlußübergangszeit scheint zurzeit die beste, wenn auch unvollständige Erklärung der Vokaldauerschwankungen zu sein.

Résumé

La variation de durée des voyelles selon la sonorité de la consonne suivante

L'analyse spectrographique de 376 voyelles prononcées par quatre parleurs de langues différentes a mené l'auteur à constater que la variation de durée des voyelles selon la sonorité de la consonne suivante est une propriété universelle des langues, et que cette propriété nécessite une explication articulatoire inhérente. L'auteur a discuté ensuite quelques hypothèses concernantes les faits physiologiques à base des variations dans la durée des voyelles sous les conditions décrites. Les données expérimentales de l'auteur et des autres

expériences déjà publiées ont été employées dans l'évaluation des conclusions de plusieurs auteurs. L'hypothèse basée sur la transition de fermeture semble être, au présent, la meilleure explication, quoique incomplète, de la variation de durée des voyelles.

References

BELASCO, S.: The influence of force of articulation of consonants on vowel duration. J. acoust. Soc. Amer. *25:* 1015–1016 (1953).

BRANDSTÄTER, H. J.: Vokaldauer und Positionseinfluß beim Sprachvergleich. Z. Phon. *18:* 417–419 (1965).

DELATTRE, P.: Some factors of vowel duration and their cross-linguistic validity. J. acoust. Soc. Amer. *34:* 1141–1143 (1962).

DENES, P.: Effect of duration on the perception of voicing. J. acoust. Soc. Amer. *27:* 761–764 (1955).

FANT, G.: Acoustic theory of speech production ('s-Gravenhage 1960).

FINTOFT, K.: The duration of some Norwegian speech sounds. Phonetica *7:* 19–39 (1962).

HALLE, M. and STEVENS, K. N.: On the mechanism of glottal vibration for vowels and consonants. QPR, Research Laboratory of Electronics, MIT, No. 85.267–271 (1967).

HOUSE, A. S.: On vowel duration in English. J. acoust. Soc. Amer. *33:* 1174–1177 (1961).

HOUSE, A. S. and FAIRBANKS, G.: The influence of consonant environment upon the secondary acoustical characteristics of vowels. J. acoust. Soc. Amer. *25:* 105–113 (1953).

KLATT, D. H.: Articulatory activity and air flow during the production of fricative consonants. QPR, Research Laboratory of Electronics, MIT, No. 84.257–259 (1967).

KLATT, D. H.; STEVENS, K. N., and MEAD, J.: Studies of articulatory activity and airflow during speech. Ann. N.Y. Acad. Sci. *155:* 42–55 (1968).

KOZHEVNIKOV, V. A. and CHISTOVICH, L. A.: Speech: articulation and perception. English translation distributed by Joint Publications Research Service, Washington, D.C. (4th printing) (1967).

KRONES, R.: How to use the Pitch-program. Monthly Internal Memorandum, Phonology Laboratory, University of California, Berkeley, January issue, pp. 55–77 (1969).

LEHISTE, I.: Suprasegmentals (Columbus, Ohio, MS 1969).

LEHISTE, I. and PETERSON, G.: Transitions, glides and diphthongs. J. acoust. Soc. Amer. *33:* 268–277 (1961).

LINDBLOM, B.: On vowel reduction. The Royal Institute of Technology, Stockholm (Licentiatavhandling) (1963). – Vowel duration and a model of lip mandible coordination. STL-QPSR 4.1–29, The Roayl Institute of Technology, Stockholm (1968).

LISKER, L.: Closure duration and the intervocalic voiced-voiceless distinction in English. Language *33:* 42–49 (1957).

MACNEILAGE, P. F.: The serial ordering of speech sounds. Project on Linguistic Analysis, Phonology Laboratory, University of California, Berkeley, 2.8, p. M1–52 (1968).

MOHR, B.: Intrinsic variations of accoustic parameters of speech sounds. Project on Linguistic Analysis, 2.9: M1–44 (1969).

OHALA, J.; HIKI SHIZUO; HUBLER, S., and HARSHMAN, R.: Photoelectric methods of transducing lip and jaw movements in speech. Working papers in phonetics, University of California, Los Angeles, *10:* 135–141 (1968).

ÖHMAN, S.: Peripheral motor commands in labial articulation. STL-QPSR 4.30–63, The Royal Institute of Technology, Stockholm (1967).

PETERSON, G. and LEHISTE, I.: Duration of syllable nuclei in English. J. acoust. Soc. Amer. *32:* 693–703 (1960).

ROSITZKE, H.: Vowel length in general American speech. Language *15:* 99–109 (1939).

SACIA, C. F. and BECK, C. F.: The power of fundamental speech sounds. The Bell system technical journal *5:* 393–403 (1926).

SCHARF, D. J.: Vowel duration in whispered and in normal speech. Language Speech *7:* 89–97 (1964).

WANG, W. S.-Y.: The many uses of F_0. Project on Linguistic Analysis, Phonology Laboratory, University of California, Berkeley, 2.8, pp. W1–35 (to appear in Papers on Linguistics and Phonetics in Memory of PIERRE DELATTRE) (1968).

WANG, W. S.-Y. and FILLMORE, CH. J.: Intrinsic cues and consonantal perception. Journal of Speech and Hearing Research *4:* 130–136 (1961).

ZIMMERMAN, S. A. and SAPON, S. M.: Note on vowel duration seen crosslinguistically. J. acoust. Soc. Amer. *30:* 152–153 (1958).

Author's address: Dr. MATTHEW CHEN, Phonology Laboratory, Department of Linguistics, University of California, *Berkeley, CA 94720* (USA).