Physiological effects on stop consonant voicing

Patricia A. Keating

Slightly expanded version of
paper presented at ASA meeting, May 1983, Cincinnati

Introduction

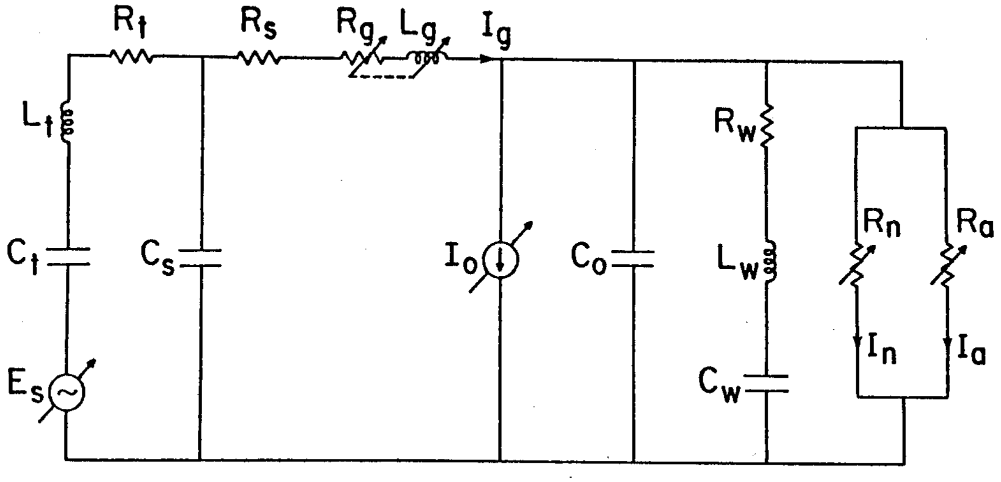
The effect of the voicing distinction on various temporal measures for stop consonants has been much studied. Phonologically voiceless stops are known to have longer Voice Onset Times, that is, more phonetic voicelessness, than phonologically voiced stops. They may also have less closure voicing and a longer total voiceless interval. However, other phonological distinctions besides voicing affect these same timing measures. So, for example, place of articulation is known to affect VOT and closure duration, and stress to affect VOT.

Various researchers, in noting these diverse effects, have offered suggestions that they are due to aerodynamic conditions, as determined by such things as vocal tract volumes, release velocities, and heightened subglottal pressure (e.g. Klatt 1975, Smith and Westbury 1975, Weismer 1980). Although some of these suggestions have been picked up by others and repeated as fact, to date there has been little evidence offered in their support. This paper reports preliminary results of an attempt to explore various possible explanations of temporal effects using an aerodynamic model.

Method

The model, shown in the figure below, is a computer implemented simulation of a circuit analog of the vocal tract derived from Rothenberg (1968) and similar to the model of Muller and Brown (1980). Details about values of circuit elements for this version are given in Westbury (1983). In such a model, voltage is taken as the analog of air pressure, and current as the analog of volume velocity airflow. Values for subglottal pressure, flow through the glottis, oral pressure, and flow through the oral constriction, among other things, are calculated for each small moment in time. Our interest here will focus on when aerodynamic conditions will permit voicing, other conditions being favorable: specifically, when the difference between subglottal pressure and oral pressure is sufficient to permit vocal cord vibration. It will be assumed that a difference of 2000 dynes/cm^2 is necessary to sustain voicing, while a difference of 3000 dynes/cm^2 is necessary to initiate voicing. All results of simulations reported below represent time elapsed until such pressure differences were obtained.

As a source of acoustic data to be modeled, six speakers each of California English, Stockholm Swedish, and Tokyo Japanese were recorded reading real words with stop consonants /b d g p t k/ in initial and medial positions, before nonhigh vowels but after an uncontrolled set of vowels. For English, the 196 words also systematically varied degree of stress; for Swedish, the 372 words also varied pitch accent and stress, and for Japanese the 48 words also varied pitch accent. From a computer oscillographic display, up to three measurements were made for each stop: duration of voiced closure, duration of voiceless closure, and duration of voicing lag. Measurements were averaged across speakers of a language for each word type.



Results

A topic of some interest in phonetics has been the maintenance of voicing during the closures of medial [b], [d], and [g] through passive expansion of the oral cavity, which keeps oral pressure relatively low. The effect of place of articulation on the duration of this medial voicing is illustrated in Table 1, which shows data for [b], [d], and [g] in two English stress conditions and one Swedish pitch-stress condition. The numbers represent msec of voicing after stop closure.

Table 1

	ENGLISH		SWEDISH
	_ main stress	_ 2ary stress	_ main stress
b	55	42	71
d	40	39	57
g	31	24	43

Although such effects used to be thought due to differences in cavity size behind the oral constriction (Smith and Westbury 1975), it is now recognized that the contribution of this parameter is quite small compared to the contribution of the surface area of the cavity walls. Rothenberg (1968), Muller and Brown (1980), and Ohala (1983) all agree that the differences in compliant wall area across place of articulation *should* produce the observed differences in voicing duration, but none of them have presented supporting data from modeling. This is the goal of our first modeling exercise. The three places of articulation were simulated at three degrees of wall impedance following Ishizaka et al. (1975). No subglottal or glottal differences were assumed. Differences in place of articulation can be represented as differences in cavity volume, surface area of walls as it influences impedance parameters, the dimensions of the oral constrictions, and the speed of the closing gesture. However, independent manipulations of these variables indicate that the largest effect by far on closure voicing is due to the wall surface variables. Table 2 gives the duration of voicing to be expected from simulations for [b], [d], and [g] when the walls are like lax cheeks, when they are like moderately tensed cheeks, and when they are like the neck walls. The linguistic data in Table 1, with which the simulations in Table 2 can be compared, suggest wall values that were slightly more tensed than the tense cheek values used here. But the point is that for any given setting, place of articulation differences will produce acoustic differences in voicing maintenance: fronter places, with greater surface area, have more voicing.

Table 2

	lax cheeks	tense cheeks	neck
b	145	54	25
d	124	46	21
g	105	36	13

[b], [d], and [g] are not the only stops to have voicing during closure: for most speakers, especially of English, [p], [t], and [k] will have at least one or two pitch periods of closure voicing. Table 3 shows durations of such closure voicing in English before reduced vowels, in Japanese between High and Low pitches, and in Swedish in Accent 1 words between reduced and stressed vowels. As Rothenberg suggested, the vocal cords' opening gesture for voiceless segments will allow some vibration, presumably breathy, before their separation makes voicing impossible. Will wall surface area differences across place of articulation produce the voicing duration differences observed?

Table 3

	ENGLISH	JAPANESE	SWEDISH
	_ reduced V	H _ L	reduced V _ stressed V
p	13	10	16
t	7	10	12
k	6	5	8

The first column in Table 4 below shows (as closure voicing in msec) the result of simulations in which all three stops have a constant glottal gesture beginning at closure and walls like tense cheeks. These simulations indicate that the differences due to wall surface area are quite small, with only a 3 msec difference between labial and velar. The differences in the acoustic data in Table 3, while also small, are larger, and may indicate some additional mechanism. One such mechanism is suggested by observations on the timing of the glottal gesture relative to consonant closure and to consonant release. Lofqvist (1980) and others (e.g. Lofqvist and Yoshioka 1981) have noted that for aspirated stops the glottal gesture appears to be timed to begin at the moment of stop closure and to reach its peak opening value within 20 ms before the moment of release. That is, the time to peak glottal area is proportional to closure duration. Extending this observation to the case at hand, we can note that stops at different places of articulation differ in closure duration: fronter places have longer closures. As long as peak glottal area does not vary across place of articulation, then fronter places of articulation will have more time for the vocal cords to travel the same distance, that is, slower glottal gestures. With these assumptions, the right order of magnitude of voiced closure will result; the second column in Table 4 shows the durations of closure voicing in msec that result when the labial closure duration (and therefore the time to peak glottal area) is 10 msec longer than the alveolar, and 20 msec longer than the velar.

Table 4

	wall effects only	add varying glottal gesture
p	15	15
t	13	12
k	12	9

Consider next the effect of stress on medial [b d g] voicing. First, English [b d g] before reduced vowels are typically voiced throughout their closures, which are quite short. But if there is a break in voicing, as is common for stops before non-reduced vowels, then there is more stop closure voicing before a more stressed vowel. Table 5 shows durations of closure voicing for Swedish and English [b d g] as affected by the stress on a following vowel.

Table 5

	SWEDISH		ENGLISH	
	_ main stress	_ other V	_ main stress	_ other V
b	71	56	55	45
d	57	31	40	34
g	43	39	31	28

Main stress increases the duration of voicing. To see why this rather unexpected result should hold, consider how stress is effected through increased respiratory muscle activity, resulting in greater subglottal pressure. Data of Ladefoged and colleagues (Ladefoged 1967) shows that there is muscle activity before a stressed vowel, that is, during the closure of the preceding consonant. These data, and similar data of Lieberman (1967), show an increase in subglottal pressure of from 1 to 5 cm aq, with peak pressure after the onset of the vowel, and relatively smooth increases and decreases of about 100 msec around that peak. Simulations indicate that a peak in subglottal pressure will lag a respiratory force peak by about 20 msec. Table 6 compares the effect (msec closure voicing) on labials of such a muscularly-induced boost in subglottal pressure when it begins at the moment of closure and lasts longer, and when it begins half-way through closure and is somewhat shorter, with the ordinary case we saw before for b. Under the two boost conditions, both subglottal and oral pressure will be higher during closure, but the subglottal pressure will be proportionately higher, such that closure voicing will be increased in duration by about 5 msec, on one simulation, or will last longer than the consonant, on the other.

Table 6

Long P_s boost	Short P_s boost	No P_s boost
100+	59	54

The finding that stress on a following vowel favors voicing, all things being equal, does further work for us. Table 7 compares lag VOT in initial stops in English before main stressed, secondary stress, and reduced vowels. Stress on the vowel following [b], [d], and [g] also decreases the VOT value, as has been noted by Lisker and Abramson (1967).

Table 7

	_ main stress V	_ 2ary stress V	_ reduced V
b	10	10	13
d	13	16	18
g	24	27	27

An initial /b/ was simulated under two conditions, with and without the extra muscularly-induced boost in subglottal pressure. In the case without, VOT is about 8 msec, but if we give an extra respiratory push, subglottal pressure will be relatively higher than the oral pressure, and voicing begins at 6 msec, that is, 2 msec earlier. It's not clear that the magnitude of the simulated effect is sufficient compared to the data, but the direction of the effect found is encouraging.

Conclusions

Although more work clearly remains to be done, we have seen that certain observations about acoustic effects of place of articulation and stress may be accounted for as consequences of other physiological variables. The effects of place and stress on the duration of closure voicing for [b d g] may derive from independently necessary factors, such as wall area and respiratory force. The effect of place on [p t k] may involve a more arbitrary factor, the starting time and speed of the glottal gesture. Further work, with more attention to small cross-language differences, may motivate these interarticulator timing differences. At the same time, it may elucidate how the physical properties of the speech production system constrain acoustic variation.

Acknowledgement

This research was supported by a grant from NIH to Peter Ladefoged.

References

Ishizaka, K., J. C. French, and J. L. Flanagan (1975). "Direct Determination of Vocal Tract Wall Impedance", IEEE Trans. Acoust., Speech Signal Process. ASSP-23 (4), 370-73.

Klatt, D. H. (1975). "Voice-onset time, frication, and aspiration in word-initial consonant clusters", J. Sp. Hear. Res. 18, 686-706.

Ladefoged, P. L. (1967). Three Areas of Experimental Phonetics (Oxford University Press, London).

Lieberman, P. (1967). Intonation, Perception, and Language (MIT Press, Cambridge).

Lisker, L., and A. S. Abramson (1964). "A cross-language study of voicing in initial stops: acoustical measurement", Word 20, 384-422.

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

Lofqvist, A. (1980). "Interarticulator programming in stop production", J. Phon. 8, 475-490.

Lofqvist, A., and H. Yoshioka (1981). "Interarticulator programming in obstruent production", Phonetica 38, 21-34.

Muller, E. M., and W. S. Brown, Jr. (1980). "Variations in the Supraglottal Air Pressure Waveform and their Articulatory Interpretation", in Speech and Language: Advances in Basic Research and Practice, Vol. 4, edited by N. Lass (Academic Press, New York), pp. 317-389.

Ohala, J. J. (1983). "The Origin of Sound Patterns in Vocal Tract Constraints", in The Production of Speech, edited by P. F. MacNeilage (Springer-Verlag, New York), pp. 189-216.

Rothenberg, M. (1968). The Breath-Stream Dynamics of Simple-Released-Plosive Production, Bibl. Phonetica 6.

Smith, B. L., and J. R. Westbury (1975). "Temporal control of voicing during occlusion in plosives", paper presented at April ASA meeting.

Weismer, G. (1980). "Control of the voicing distinction for intervocalic stops and fricatives: some data and theoretical considerations", J. Phon. 8, 427-38.

Westbury, J. R. (1983). "Enlargement of the supraglottal cavity and its relation to stop consonant voicing", J. Acoust. Soc. Am. 73, 1322-36.