Oral Cavity Enlargement in Retroflex Stops

Ronald L. Sprouse[*], Maria-Josep Solé[#] & John J. Ohala[*]

[*]Department of Linguistics, University of California, Berkeley
[#]Department of English, Autonomous University of Barcelona, Spain
E-mail: ronald@berkeley.edu, mariajosep.sole@uab.es, ohala@berkeley.edu

Abstract

It is generally recognized that front-articulated stops are more compatible with voicing than back-articulated stops because the larger oral volume and compliance of front stops allows for longer trans-glottal airflow. Phonological patterns suggest that retroflex stops may be an exception to this pattern and may be more compatible with voicing than their dental/alveolar counterparts. An experiment was conducted in which an artificial leak was created in three speakers, and the effect of the abrupt closure of this leak on voicing was measured for [b d ɖ g]. The results of this experiment show that [b d g] follow the pattern of longer voicing for more front articulations. The retroflex was an exception – for two speakers voicing for [ɖ] persisted longer than for [d], and for one speaker [ɖ] voicing persisted longer than for [b d]. We believe that the greater surface area presented by the concave shape of the tongue during retroflexes as compared to dentals/alveolars allows for greater passive cavity expansion (i.e. compliance) and is a possible explanation of the observed pattern.

1 Introduction

The relative difficulty of maintaining voicing during a stop has been the object of extensive investigation (e.g., [1, 2]). The duration of voicing during the stop closure has been shown to vary, among other factors, with place of articulation of the consonant, labials retaining voicing more readily than velars – due, it is hypothesized, to differences in the amount of compliant surface area that could accommodate the glottal airflow [3]. But undercutting a simple pattern of "the more front the articulation, the easier it is to maintain voicing during a stop" is a pattern where retroflex stops are more favored for voicing than more forward-articulated apicals. For example, the voiced/implosive counterpart to the dental voiceless stop is a retroflex in a number of languages [4]. This is the case in Austronesian stop systems where the voiced stop is more retracted than its voiceless counterpart --sometimes retroflex but sometimes alveolar whereas the voiceless one is dental, i.e., apicoalveolar or retroflex /d/ *vs* apicodental /t/ [5]. Diachronically, when the implosive corresponding to a plain dental or alveolar has become retroflex or retracted, the glottalic feature tends to be lost [6:102]. Thus the common reflex of an earlier implosive is a retroflex [r] in Central Sudanic languages.

The main goal of this study was to obtain data on vocal tract volume and compliance for retroflex voiced stops vis-à-vis stops at other places of articulation, in order to elucidate this pattern.

2 Method

In order to estimate the overall compliance of the walls of the supraglottal cavity for various places of articulation, including retroflexes, a refined version of Ohala and Riordan's (1979) experimental procedure was used. In this procedure voicing during a stop closure is prolonged by means of an experimenter-controlled artificial leak inserted through the nasal cavity. The experimenter closes the leak, and the time from leak closure to devoicing is measured.

In the present experiment three speakers (speakers 1 and 2, native male speakers of American English, and speaker 3, a native female speaker of Catalan, all trained phoneticians) produced $V_1C{:}V_1$ utterances, where V= [i:] or [ɑ:] and C: was a prolonged [b d ɖ g]. (A fourth speaker was recorded, but the nasal venting was insufficient for that speaker to maintain adequate voicing during stop closures, and her results were not analyzed.) The same vowel was produced before and after the consonant to ensure that any coarticulation with the vowel was the same for each token.

The artificial leak was created by venting the oral pressure during the production of the initial vowel and the beginning of the stop closure by means of a nasal catheter (circular cross-sectional area of 4.76mm inner diameter, approximately 20cm long). The nasal vent allowed air to flow out during the stop closure and thus maintain sufficient pressure drop for voicing. The outer end of the catheter was coupled to a solenoid-activated valve that was triggered by a hand-operated toggle switch controlled by the experimenter. The experimenter held the valve open during the initial vowel and closed it at unpredictable times during the production of the stop. The speakers maintained a steady-state closure – delayed the stop release – until after voicing was extinguished. The other end of the solenoid valve was coupled to a Fleisch pneumotrachograph in order to record airflow through the valve so that the precise time of the valve closure could be determined.

For subject 1 oral air pressure was recorded by a catheter inserted into the pharynx via the other nostril and connected to a pressure transducer. For subjects 2 and 3 it was not possible to use the other nostril, and oral pressure was recorded with an oral catheter placed between the lips and routed behind the molars via the buccal sulcus (the space between the teeth and cheeks) A laryngograph recorded vocal fold vibration during the vented and blocked conditions, and a microphone recorded the audio signal. The laryngograph, oral pressure, and audio signals were all useful for determining when devoicing occurred.

In total four signals were recorded using National Instruments PCI-6013 data acquisition hardware and the Matlab Data Acquisition Toolbox (20kHz sample rate per channel and 16 bits/sample): 1) audio signal (AKG C520 microphone and M-Audio AudioBuddy microphone preamp); 2) laryngograph (Laryngograph Ltd. portable); 3) oral air pressure (Biopac TSD160C pressure transducer); 4) airflow through the nasal catheter and solenoid valve (Fleisch pneumotachograph with Biopac TSD160A pressure transducer).

The airflow and oral pressure signals were digitally filtered with a 400Hz lowpass filter, a cutoff that allowed for the simultaneous visualization of relatively slow-changing aspects of the aerodynamic signals and the effects of voicing. A 20Hz highpass digital filter was applied to the laryngograph signal to minimize the effects of larynx movement.

The solenoid closure time was determined by calculating the second order differential of the airflow signal. The minimum of this differentiated signal corresponds to the moment when the airflow is decreasing most rapidly. Comparison with the oral pressure recordings confirms that oral pressure also rises rapidly at this moment (see Figure 1).

Figure 1 also shows that the most rapid decrease in airflow occurred near the beginning of the valve closure. Full closure of the solenoid valve occurred approximately 10ms later. We chose the beginning of the closure phase (indicated by the vertical line in Figure 1) as the defined closure time since the resistance presented by partial closure of the valve was sufficient to extinguish voicing for some tokens. .

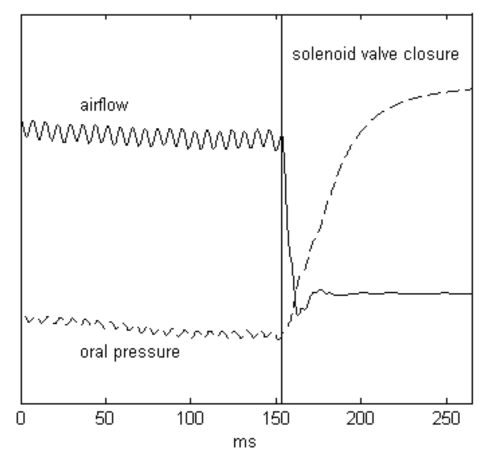

Figure 1**:** *Solenoid valve closure at the vertical line during utterance 'idi'.*

Measurements of the end of voicing were performed by hand. A peak detector was used to identify the last glottal pulse of the laryngograph signal, but the detector was not very accurate and all tokens were visually inspected and corrected as needed. In some cases voicing was clearer in the oral pressure signal, and for these cases the pressure signal was used to find the last glottal pulse. To a much lesser extent spectrograms of the audio signal were also examined to verify the absence of voicing.

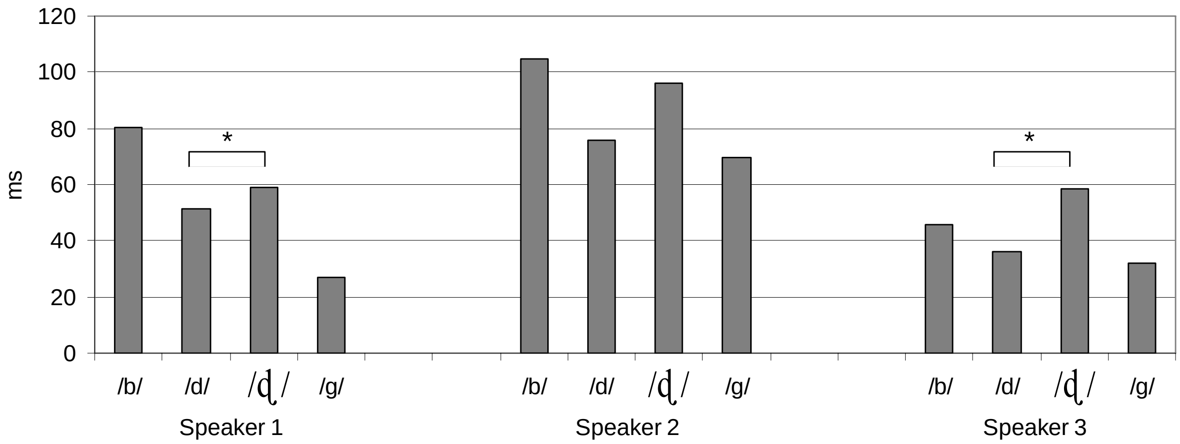

Figure 2: *Median duration of voicing (in ms) in intervocalic stops due to passive supraglottal compliance for the 3 speakers.*

Table 1. *Descriptive statistics. Values in ms.*

	Speaker 1				**Speaker 2**				**Speaker 3**			
	N	*Mean*	*SD*	*Median*	*N*	*Mean*	*SD*	*Median*	*N*	*Mean*	*SD*	*Median*
/b/	45	84.87	37.83	80.15	34	110.40	28.82	104.7	34	44.81	71.71	45.85
/d/	42	49.87	20.34	49.27	28	75.03	31.21	75.72	49	39.14	20.45	36.15
/ɖ/	66	64.66	29.86	59.22	35	88.14	35.56	95.85	55	75.88	48.06	58.55
/g/	33	36.31	25.2	27.0	36	68.88	40.74	69.45	68	36.94	24.94	31.80

3 Results

The results for the three speakers are presented in Figure 2, which shows median duration of voicing after the valve closed for the different places of articulation. Figure 2 shows that, with one exception, the further back the stops were articulated, the earlier did voicing cease, as previously reported in the literature. The exception was the retroflexes, which in speakers 1 and 2 showed voicing longer than alveolars and in speaker 3 had voicing lasting longer than any other place of articulation. Median values of voicing duration, rather than the mean, are plotted in Figure 2 because the distribution of values for voicing duration was positively skewed, i.e., there were a few very long values, of the order of 200ms. The mean values and SD, in addition to the median, for the individual consonants are reported for each speaker in Table 1, although due to the skewing the SD must be interpreted with caution. Speakers 2 and 3 showed unusually long voicing values for velar stops, as compared to speaker 1 and as compared to values reported in [3]. We attribute the relatively long voicing values for velars to the lack of a complete seal around the catheter recording the oral pressure for these speakers. The catheter bent round the back molars into the oral cavity interfered with the articulation of back stops, and an inadequate seal may have created a leak resulting in longer voicing values.

Two-factor ANOVAs were performed for each individual speaker to evaluate the effects of place of articulation (labial, alveolar, retroflex, velar) and vowel context ([ɑː] vs. [i]) on duration of voicing after the valve closed. The main effect of place of articulation was significant at the $p<0.01$ level for the three speakers ($F(3, 176)= 22.04$, Speaker 1; $F(3, 123)= 10.09$, Speaker 2; $F(3, 225)= 21.38$, Speaker 3). The main effect of vowel context was only significant for speaker 1 ($F(1, 176)= 15.67$, $p<0.01$), with [ɑː] showing longer voicing duration than [iː], contrary to what has been previously reported in the literature [3, 7]. Interaction effects were significant at the $p<0.01$ level for speakers 2 and 3 ($F(3, 123)=4.11$, ($F(3, 225)= 4.14$, respectively), indicating that certain places of articulation showed longer voicing in the [ɑː] context, and other places in the [iː] context, with no consistent pattern across the two speakers.

Post-hoc analyses (Tukey HSD) to follow up the significant main effect of place of articulation indicate that retroflex stops showed significantly longer voicing than alveolars for speakers 1 and 3. The differences did not reach significance for speaker 2 due to the large standard deviation. Other significant differences were those between labials and alveolars (speakers 1 and 2); labials and retroflex (all speakers); labials and velars (speakers 1, 2), and velars and retroflex (speakers 1, 3). Non-parametric tests (Kruskal-Wallis) were also used to compare the medians of the different places of articulation, with the same results obtained for the ANOVAs and pairwise comparisons.

4 Discussion

The longer duration of voicing in retroflex stops indicates a larger passive expansion in the volume of the oral cavity during their constriction than for the more front-articulated alveolars. We think a possible explanation for these findings is that the concave tongue shape for the retroflex (shown in x-rays analysis [8]) may provide a greater surface area of the tongue that can be displaced in response to the impinging air pressure, compared to alveolars. The increase in oral volume allowed by the greater compliant surface area may accommodate more glottal airflow before pressure rises sufficiently to extinguish voicing.

In addition to the differences found during the prolonged steady-state stops of the experimental conditions, another factor may favor a longer voicing in natural connected speech: some forward movement of the tongue apex during the retroflex articulation. That is, the movement over time of retroflex sounds may by itself change vocal tract volume during the stop closure. It has been reported that the tongue tip first bends back to form a retroflex articulation and then slides forward during the closure phase enlarging the oral cavity [9]. In this way, the time course of the retroflex articulation usually has some component of cavity enlargement which may also account for a longer duration of voicing. Such cavity enlargement during the constriction is an essential characteristic of implosives in order to maintain voicing.

If voicing is considered an essential characteristic of implosives, and retroflexion facilitates voicing by enlarging the oral cavity and prolonging voicing, then retroflexion can be considered a mechanism to facilitate and enhance voicing in apical implosives. Hence the common covariation between implosivization and retroflexion in the observed phonological patterns.

Finally, in order to further test this explanation and estimate the volume of the expansion, variations in oral cavity volume during the stop closure due to active (e.g., larynx lowering for implosives) and passive expansion (e.g. for different places of articulation) and their effects on voicing are currently being simulated with an aerodynamic model.

References

[1] J. Ohala. The origin of sound patterns in vocal tract constraints. In P. F. MacNeilage (ed.), *The production of speech* (pp.189-216). New York: Springer-Verlag, 1983.

[2] J.R. Westbury & P. A. Keating. On the naturalness of stop consonant voicing. *Journal of Linguistics* 22:145-166, 1986.

[3] J. Ohala and C. Riordan. Passive vocal tract enlargement during voiced stops. In J. J. Wolf & D. H. Klatt (eds.), *Speech communication papers*. New York: Acoust. Soc. of Am. 89-92, 1979.

[4] J.H. Greenberg. Some generalizations concerning glottalic consonant, especially implosives. *International Journal of American Linguistics*, 36(2):123-145, 1970.

[5] F. Ozanne-Rivierre. The Proto-Oceanic consonantal system and the languages of New Caledonia. *Oceanic Linguistics*, 31 (2): 191-207, 1992.

[6] A.N. Tucker. The Eastern Sudanic languages. Oxford Press, 1940.

[7] D. Pape, C. Mooshammer, P. Hoole & S. Fuchs. Devoicing of word-initial stops: a consequence of the following vowel? *Proceedings of the 6th International Seminar on Speech Production*. Sidney, 207-212, 2003.

[8] P. Branderud, H.J. Lundberg, J. Lander, H. Djamshidpey, I. Wäneland, D. Krull & B. Lindblom. X-ray analyses of speech: Methodological aspects. http://www.ling.su.se.staff/peter/Pb_Bli.html

[9] Rothenberg, M. 1968. *The breath-stream dynamics of simple-released-plosive production*. Basel: Karger.