Voice onset time and the realization of voiced stops in Kwak'wala

Connor Mayer
University of British Columbia

Although it has traditionally been described as contrasting ejective, voiceless aspirated, and voiced series of stops, there are no previous instrumental phonetic analyses of stop and affricate realizations in Kwak'wala. This study examines voice onset time in a single native speaker. VOT, the timing between the release of stop consonant closure and the onset of periodic voicing, has proven to be a reliable way of characterizing laryngeal state in the production of stops and affricates. In addition to providing a descriptive account, this study addresses several issues of broad phonetic interest, including the effects of context on VOT. Results also suggest that the realization of Kwak'wala voiced stops is more complex than previously thought: word-initial voiced stops tend to be produced without voicing, while intervocalic ones range from fully voiced to voiceless. This suggests a similar realization of stop voicing to English, which is characterized by an adduction of the vocal folds without mandatory vocal fold vibration. Data from more speakers is necessary to say whether this is typical across speakers, or the result of language transfer or obsolescence effects.

1 Introduction

Kwak'wala is an endangered Northern Wakashan language spoken on both sides of the Queen Charlotte Sound and northern Vancouver Island with approximately 250 native speakers. Although some linguistic work has been conducted on Kwak'wala, it has primarily been focused on the phonology and syntax of the language, and to the author's knowledge there have been no previous instrumental phonetic analyses. Following Ladefoged's suggestion that "[a]ny description of the phonetic structures of a language should include an account of the VOT" (2003), this paper will provide a phonetic analysis of voice onset time (VOT) in Kwak'wala. VOT is a measure of the time at which periodic voicing begins relative to the time of the release of a preceding stop or affricate, and has proven to be an effective way of characterizing the laryngeal configuration of these sounds (Lisker & Abramson 1964). Aspirated and ejective sounds display a positive VOT, voiced sounds display a negative VOT, and voiceless unaspirated sounds display a VOT close to 0.

Kwak'wala has traditionally been described as having a three-way distinction between voiced, voiceless aspirated, and ejective series of stops and affricates (Boas 1911; Grubb 1969; Lincoln & Rath 1960; etc.). A summary of these series is provided in table 1.

p'	t'	ƛ'	c'	kʲ'	kʷ'	q'	qʷ'
p	t	ƛ	c	kʲ	kʷ	q	qʷ
b	d	λ	ʒ	gʲ	gʷ	ġ	ġʷ

Table 1. The stop and affricate inventory of Kwak'wala

The present paper will give a descriptive account of the VOT of these sounds in a variety of contexts, particularly focusing on whether the series typically described as voiced is indeed voiced or voiceless unaspirated, as well as examining three cross-linguistic findings and predictions that have been made regarding contextual effects on VOT.

The first of these is the effects of word-initial vs. intervocalic context on VOT. The effects on voiced stops are fairly well understood: because the vocal folds are vibrating prior to closure for intervocalic stops, they will typically display a greater degree of voicing than word-initial voiced stops. The effects of initial vs. intervocalic context on voiceless and ejective stops is somewhat less clear. Previous research has shown, however, that stops occurring at the edges of prosodic constituents will tend to display initial/final strengthening: in terms of VOT, this equates to a more positive VOT for aspirated stops and a more negative VOT for voiced stops (see Cho & Jun (2000) for a summary). Thus it is likely that voiceless and ejective sounds will display a longer VOT word-initially than intervocalically when the word-initial context occurs at a prosodic boundary. Voiced stops that occur word-initially are more difficult to predict: although it is possible that initial strengthening might result in a more negative VOT, there is also an inherent aerodynamic difficulty in beginning voicing with no preceding vowel, which could result in a less negative VOT.

The second contextual effect is that of following vowel on VOT. Previous research examining this issue suggests that voiced and voiceless stops exhibit longer negative and positive VOT respectively before high, close vowels such as /i/ and /u/. Voicing is relatively difficult to maintain in stops: because air cannot escape from the oral cavity, the pressure differential across the larynx quickly falls below the 2cm/H_2O required for voicing to occur. Both passive (Westbury 1983) and active (Bell-Berti 1975) expansion of the vocal tract is often used to help prolong voicing. Ohala and Riordan (1979) attempted to determine how long voicing could continue in stops with only passive expansion of the vocal tract, finding that stops preceding /u/ and /i/ allowed voicing to continue for longer than stops preceding other vowels. They suggest that the reason for this is the greater pharyngeal cavity volume in these vowels, which allows voicing to continue for a greater period of time before the subglottal and supraglottal pressures equalize, resulting in a more negative VOT. Ohala (1981) writes that voiceless stops tend to undergo an opposite process, with a greater positive VOT before high, close vowels. Although the reason for this is not completely understood, it has been suggested that the tighter constriction in high close vowels, particularly for /i/, offers greater resistance to air flowing from the oral cavity, lengthening the time it takes to achieve a suitable cross-laryngeal pressure differential suitable for voicing. It is likely, therefore, that voiced stops will display a more negative VOT and voiceless stops a more positive VOT before the vowels /u/ and /i/. It is unclear whether following vowel context has any effect on the VOT of ejective sounds.

The third contextual effect to be examined is the effect of place on VOT. It has been noted many times in the literature that VOT is positively correlated with the posteriority of the place of articulation (see the summary in Cho & Ladefoged 1999). Voiced stops tend to have a less negative VOT in more posterior sounds due to the smaller volume between the glottis and oral closure, which makes it difficult to maintain voicing. It has also been observed that voiceless stops tend to have a more positive VOT in posterior sounds, although the reasons for this are less clear. Explanations range from aerodynamic (eg. higher pressure posterior to the closure and a larger mass of air anterior to the closure in more posterior stops, resulting in a delay in achieving sufficient transglottal pressure sufficient for voicing; a larger articulatory contact area for posterior stops resulting in a slower release due to the Bernoulli effect; higher pressure behind more posterior constrictions causing slower glottal constriction after release), physiological (slower movement of the tongue dorsum), and temporal (more anterior stops tend to have longer closure duration, resulting in a shorter VOT to achieve uniform duration across stops). Regardless of the causes, the correlation between posteriority and VOT has been shown to be fairly robust cross-linguistically, and it is likely that more posterior stops in Kwak'wala will display a higher VOT.

To summarize, this paper will provide a descriptive account of VOT in Kwak'wala stops and affricates and assess the claim that the voiced series is indeed voiced and not voiceless unaspirated. It will also address the effects on VOT of word-initial vs. intervocalic context, following vowel, and place of articulation, predicting that word-initial stops will display a longer VOT than intervocalic ones, VOT will tend to be greater before high vowels, and VOT will increase with the posteriority of the place of articulation.

2 Methods

A single native speaker of Kwak'wala originally from Kingcome Inlet acted as the language consultant for this study. A list of Kwak'wala words familiar to the speaker was constructed with her assistance. This list contains all of the sounds in table 1 occurring as the onset in syllables with primary stress before the vowels /u/, /i/, and /a/ both word-initially and intervocalically. It was not possible to control for preceding vowel in medial context. Although every effort was made to ensure the list was as complete as possible, due to phonotactic or practical reasons some gaps were inevitable. Phonotactically, the distinction between rounded and unrounded stops is neutralized before /u/ (whether this is a case of rounded stops becoming unrounded or unrounded stops becoming rounded is not clear, although the former seems more likely): thus there are no tokens containing /kʷ'u/, /kʷu/, /gʷu/, /qʷ'u/, /qʷu/, or /ġʷu/. Additionally, due to not being able to find words known to the speaker, there are also no word-initial tokens containing /p'u/, /p'a/, or /t'a/, and no intervocalic tokens containing /qwa/. With these gaps, there are 128 words in the list, 63 word-initial and 65 intervocalic.

Recordings were made using a Marantz Solid State PMD660 Recorder with a Sennheiser MK66 short shotgun microphone or an Isomax EMW Lavalier Microphone, depending on the location of the elicitation. English glosses of the stimuli were presented to the speaker in pseudo-random order, with the speaker being asked to provide a Kwak'wala translation in citation form. Citation form was chosen to ensure a stable prosodic context for all tokens and because the variety of words used made it impossible to design a single natural carrier phrase capable of incorporating all of them that did not simply include the word in citation form (eg. "The word I am going to say is ____"). In cases where a word different from the target was produced or the speaker had difficulty remembering the Kwak'wala word, the written form of the Kwak'wala target was used as a prompt. Over a series of elicitations in a sound attenuated room at the university or at the speaker's home, an average of five repetitions of each word were collected. Due to certain tokens being discarded as a result of background noise or inconsistent production and an excess of tokens for other words, the number of tokens per word ranges from 4 to 8. A total of 669 tokens were collected.

The tokens were analyzed using Praat 5.1.31. Release points of consonants and the onset of periodic voicing were marked by hand, and the values of the time of the release subtracted from the time of the onset of period voicing were calculated using a script. In the case of voiced stops, only voicing contiguous to the release was considered. Statistical analyses were done using R 2.9.1. A statistical significance level of $p = 0.05$ was used for all tests.

3 Results

The mean VOTs of word-initial and intervocalic forms by following vowel are shown in tables 2 and 3. Dashes reflect the gaps mentioned in the above section.

3.1 Statistics

Statistical comparisons were done using a one-way ANOVA with post hoc Tukey's Honestly Significant Differences tests in cases where there were more than two factors, namely for all comparisons except for word-initial vs. intervocalic position.

3.2 Manner of articulation

	a	i	u
p’	-	78	-
t’	-	106	113
k'^{j}	100	112	124
k'^{w}	135	138	-
q’	90	95	84
q'^{w}	97	113	-
ƛ’	132	148	116
c’	119	120	136
p	53	78	74
t	68	75	38
k^{j}	102	100	110
k^{w}	106	100	-
q	68	78	86
q^{w}	100	68	-
ƛ	112	112	113
c	102	135	98
b	5	4	-2
d	8	12	-2
g^{j}	35	24	42
g^{w}	42	10	-
ġ	7	6	18
$ġ^{w}$	34	20	-
λ	60	70	55
ʒ	74	64	62

Table 2: Mean VOT (ms) of word-initial stops and affricates by following vowel.

	a	i	u
p’	50	43	66
t’	80	58	70
k'^{j}	98	60	96
k'^{w}	75	98	-
q’	55	128	112
q'^{w}	68	113	-
ƛ’	88	102	94
c’	154	116	98
p	33	44	58
t	28	74	50
k^{j}	78	112	92
k^{w}	105	88	-
q	65	76	70
q^{w}	-	112	-
ƛ	104	98	102
c	80	123	105
b	-102	-70	-96
d	-67	-74	-92
g^{j}	-43	-76	-45
g^{w}	-114	-2	-
ġ	-70	-54	-32
$ġ^{w}$	-30	2	-
λ	-32	-40	-75
ʒ	-40	-56	-60

Table 3: Mean VOT (ms) of intervocalic stops and affricates by following vowel.

Significant differences were found in initial position between ejective and aspirated stops ($p < 0.001$), ejective and voiced stops ($p < 0.001$), aspirated and voiced stops ($p < 0.001$), ejective and aspirated affricates ($p < 0.02$), ejective and voiced affricates ($p < 0.001$), and aspirated and voiced affricates ($p < 0.001$). Medially, significant differences were found between ejective and voiced stops ($p < 0.001$), aspirated and voiced stops ($p < 0.001$), ejective and voiced affricates ($p < 0.001$), and aspirated and voiced affricates ($p < 0.001$). No significant differences were found between ejective and aspirated stops ($p = 0.636$) or affricates ($p = 0.741$) in medial position. These results are summarized in figures 1 and 2.

3.3 Word-initial vs intervocalic

Comparing word-initial with intervocalic sounds, the intervocalic tokens tended to show significantly lower VOT. Significant differences were found between word-initial and intervocalic ejective stops ($p < 0.001$), ejective affricates ($p < 0.01$), aspirated affricates ($p < 0.05$), voiced stops ($p < 0.001$), and voiced affricates ($p < 0.001$). Although aspirated stops tended to have lower VOT in medial position, the difference was not significant ($p = 0.064$).

3.4 Following vowel

Comparing VOT by following vowel, no significant differences were found in the ejective or voiced series. In the aspirated series, there were significant differences in VOT between /a/ and /i/ for stops ($p < 0.05$) and affricates ($p < 0.05$), with the sounds preceding /i/ having a greater VOT. These results are summarized in figures 3-5.

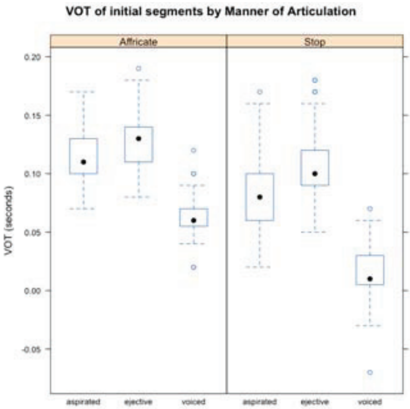

Fig. 1

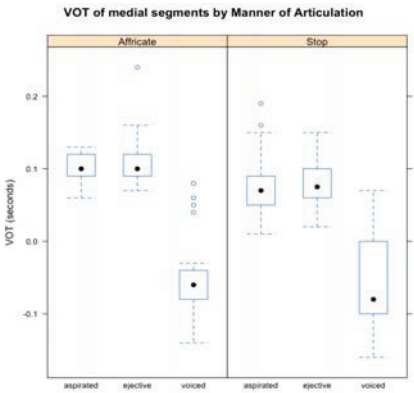

Fig. 2

3.5 Place of articulation

The comparisons of VOT by place of articulation could only be made on stops, as both the affricates in Kwak'wala are alveolar. Ejective stops displayed significant differences in VOT between labial and uvular ($p < 0.001$), labial and alveolar ($p < 0.001$), and labial and velar ($p < 0.001$) places. Aspirated stops displayed significant differences in VOT between labial and velar ($p < 0.001$), labial and uvular ($p < 0.001$), alveolar and velar ($p < 0.001$), alveolar and uvular ($p < 0.001$), and velar and uvular ($p < 0.01$) places. Voiced stops showed significant differences in VOT between labial and uvular ($p < 0.05$) places. These results are summarized in figures 6-8.

4 Discussion

4.1 Manner of articulation

The ejective, aspirated, and voiced series of stops and affricates in Kwak'wala displayed significant differences in VOT in all cases except for ejective vs. aspirated intervocalically. While the role of VOT as perhaps the most salient cue to distinguish voiced, voiceless, and aspirated stops has been

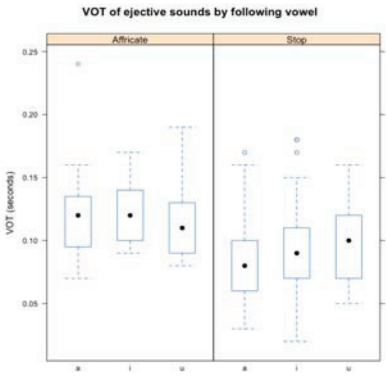

Fig. 3

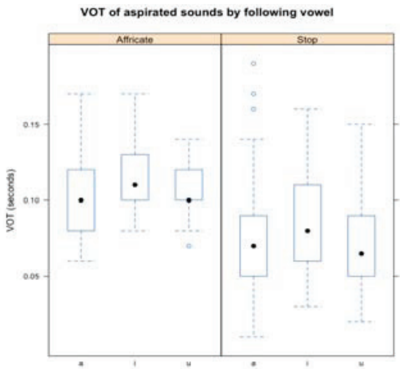

Fig. 4

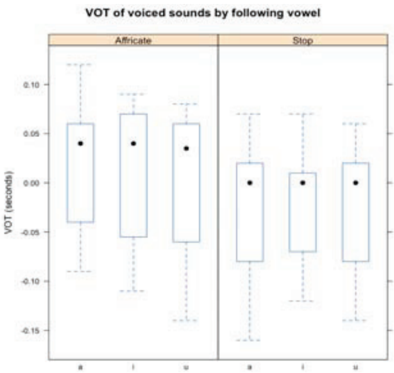

Fig. 5

well established in the literature, these results also suggest that, in Kwak'wala, VOT might be a usable cue to distinguish ejectives from aspirates word-initially, even if it is not the most salient one. The neutralization of the VOT distinction medially also raises questions as to its effect on the functional load and perceptual space of medial affricates vs. ejectives, a possible topic for further research.

4.2 Word-initial vs intervocalic context

All of the series of stops and affricates displayed significantly lower VOT in intervocalic context compared to word-initial context with the exception of aspirated stops. In the case of voiced stops, the reason for this is obvious: it is easier to maintain vocal fold vibration in obstruents when the sounds in the surrounding context have vibrating vocal folds. This is evinced by phonological processes such as intervocalic voicing, which is quite common cross-linguistically. Hence, with few exceptions, intervocalic voiced sounds tended to be fully voiced by the speaker, resulting in a negative VOT, while word-initial ones tended to be unvoiced or have very marginal prevoicing. The one exception from this is intervocalic uvular stops: these were quite often unvoiced. This is consistent with observations by linguists such as Ohala and Riordan (1979) who state that more posterior sounds will be more difficult to voice due to the smaller volume of air between the glottis and closure, as well as the smaller area of compliant surfaces (ie. the walls of the vocal tract) that could expand to prolong voicing. The realization of voiced stops in Kwak'wala will be discussed further below.

It is not completely clear why the ejective and aspirated series, with the exception of aspirated stops, displayed shorter VOT intervocalically. As stated in the introduction, a possible explanation is the strengthening of articulation and salient acoustic cues at prosodic boundaries. Because these words were elicited in citation form, word-initial sounds necessarily fell at the

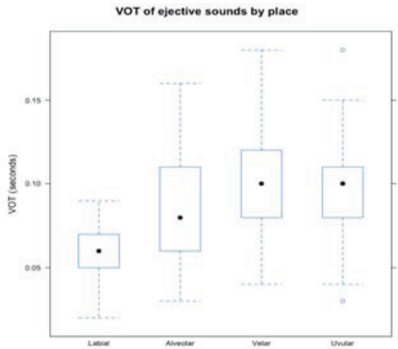

Fig. 6

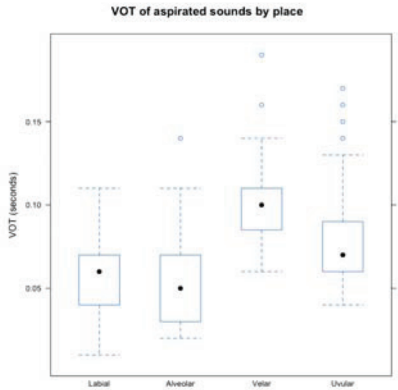

Fig. 7

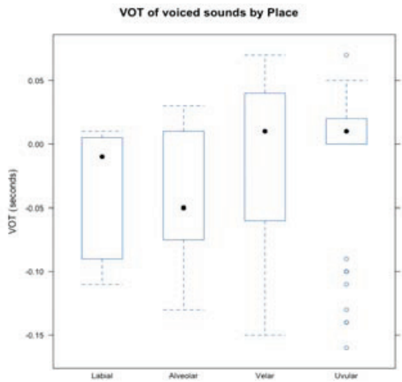

Fig. 8

highest prosodic boundary of the utterance, while intervocalic sounds did not align with any prosodic boundary. Further studies, both articulatory and acoustic, should be done to address this issue, examining specifically the realizations of these sounds at the boundaries of various levels in the prosodic hierarchy.

4.3 VOT differences by following vowel

No significant differences for ejectives and voiced stops were observed based on the following vowel. Although it is unclear if vowel context has predictable effects on ejective consonants, and no clear pattern seems to exist in the data, this finding is contrary to the prediction made by Ohala (1983) that voiced stops should have a more negative VOT before high vowels. Indeed, no clear pattern can be observed in the VOT measurements for voiced stops. Aspirates, on the other hand, display significant differences between following /i/ and /a/, with stops and affricates followed by /i/ showing a higher VOT. Although the means for the VOT of aspirated stops before /u/ and /a/ are almost identical, the fact that aspirates before /i/ have a significantly higher VOT than before /a/ supports Ohala's (1981) prediction that voiceless stops before high vowels will have a greater VOT.

4.4 VOT differences by place of articulation

In the data comparing VOT based on place of articulation, we can see a general tendency for more anterior stops to have lower VOT than posterior stops as predicted by Cho & Ladefoged (1999). Not all of these differences are significant, however. Ejectives show significant differences between labial and all other places; aspirates show significant differences between labials and velars, labials and uvulars, alveolars and velars, and alveolars and uvulars. Voiced sounds show significant differences between labials and uvulars. Thus the correlation of posteriority with VOT holds to some degree in all series, although the labial~uvular contrast is the only significant difference that occurs in all of them, and alveolar stops have a higher mean VOT than labials in both the aspirated and voiced series. The fact that velar stops in Kwak'wala are always palatalized or rounded results in a higher VOT for these sounds, which could be part of the reason that velars and uvulars never display significant differences from each other.

4.5 The realization of voiced stops in Kwak'wala

The final goal of this paper was to assess the claim that Kwak'wala has a series of voiced stops. The results suggest that this claim is indeed valid, but that the feature [+voice] is realized in a somewhat surprising way. Voiced stops in Kwak'wala are generally unvoiced word-initially and fully voiced intervocalically, with the exception of uvulars. This realization of voicing shows a pattern that is strikingly similar to the pattern which is characteristic of

English: in fact, although this voicing pattern where voiced stops are produced in a variety of ways heavily dependent on context does occur in other languages, it is sometimes referred to as 'voicing as in English' (Keating 1984). This variable realization has raised questions about what exactly the feature [+voice] in English entails phonetically, and Flege (1982) conducted a study using electroglottography to address this issue. He found that although there was substantial variability in whether initial voiced stops were produced with prevoicing, speakers consistently displayed laryngeal settings that were *sufficient* for voicing, even when the stops were produced as unvoiced: that is, speakers' vocal folds were adducted to a position where voicing could occur with the appropriate transglottal pressure differential, but in contexts when this differential is difficult to achieve, namely word-initially, prevoicing did not necessarily occur. This is compelling evidence that the feature [+voice] for stops can be realized in different ways across languages. In English it is enough to have a laryngeal setting sufficient to produce voicing, but voicing does not need to be present in the acoustic output, whereas in languages like French, [+voice] requires modal voicing to be present in the acoustic output (Keating 1984).

It is impossible to say based on the current study whether or not this manifestation of [+voice] is characteristic of Kwak'wala or is the result of language obsolescence or transfer effects from English. The language consultant who participated in this study is a native speaker of Kwak'wala, but has lived in Vancouver for several decades. Her primary means of communication in daily life is English, and she only speaks Kwak'wala on the phone to members of her family. Thus it is entirely possible that this realization of [+voice] reflects a subtle phonetic influence from English, similar to a phenomenon observed by Babel (2008) in Pauite, where younger generations of native speakers who used English as their dominant language replaced the palatalized retroflex sibilant with a sibilant closely resembling English /s/. Although a cross-generational study of this type would be impossible to do in Kwak'wala due to the demographics of native speakers, examining VOT in more speakers, particularly speakers for whom Kwak'wala plays a more consistent role in daily life, could shed light on this question. Furthermore, a study using an electroglottogram would be useful to determine whether Kwak'wala speakers are indeed displaying similar patterns of laryngeal timing as found by Flege (1982).

5 Conclusion

This study has provided a descriptive account of voice onset time for the Kwak'wala ejective, aspirated, and voiced series of stops and affricates. In addition, it has addressed several issues of broad phonetic interest: the effects of word-initial vs intervocalic position on VOT, following vowel on VOT, and place of articulation on VOT. It found that intervocalic sounds tended to display significantly lower VOT than word-initial sounds, aspirated stops exhibit significantly longer VOT before /i/ than before /a/, and anterior sounds tend to have lower VOT than posterior sounds, although there is not a clear pattern of significant differences. Furthermore, this study has shown that our language consultant uses a realization of voicing on stops that is strikingly similar to that found in English. Whether or not this is the characteristic realization of voicing on Kwak'wala or the effect of language obsolescence or transfer effects from English remains to be seen. Numerous questions remain, and only further research with a greater number of speakers can hope to address them.

References

Bell-Berti, F. (1975). Control of pharyngeal cavity size for English voiced and voiceless stops. *JASA*, 57: 456-461.

Boas, F. (1911). Kwakiutl. In *Handbook of American Indian Languages I*. 423-557. (Bureau of American Ethnology-Bulletin 40).

Babel, M. (2008). The phonetic and phonological effects of moribundity. In Gorman & Nguyen (eds.) *Pennsylvania Working Papers in Linguistics*, v. 14.2.

Cho T., & P. Ladefoged (1999). Variation and universals in VOT: evidence from 18 languages. *Journal of Phonetics*, 27: 207-229.

Cho, T., & S. Jun (2000) Domain-initial strengthening as featural enhancement: Aerodynamic evidence from Korean. *Chicago Linguistics Society 36.* 31-44.

Flege, J.E. (1982). Laryngeal timing and phonation onset in utterance-initial English stops. *Journal of Phonetics*, 10: 177-192.

Grubb, D. McC. (1969). A Kwakiutl phonology. M.A. thesis, University of Victoria.

Keating, P. (1984). Phonetic and Phonological Representation of Stop Consonant Voicing. *Language*, 60(2): 286-319.

Ladefoged, P. (2003) *Phonetic data analysis*. Oxford, UK: Blackwell.

Lincoln, N. J. & J. Rath (1980). North Wakashan Comparative Root List. *National Museum of Man Mercury Series, Canadian Ethnology Service Paper no. 68*. Ottawa: Museums of Canada.

Lisker, L., & A. S. Abramson (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20: 384-422.

Ohala, J. & C. Riordan (1979). Passive vocal enlargement during voiced stops. *Speech communication papers presented at the 97th meeting of the ASA*.

Ohala, J. (1981). Articulatory constraints on the cognitive representation of speech. In T. Myers, J. Laver, & J. Anderson (eds.), *The cognitive representation of speech.* Amsterdam: North Holland, 111-122.

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

Westbury, J. (1983). Enlargement of supraglottal cavity and its relation to stop consonant voicing. *JASA*, 73: 1322-1336.