CONSONANT TYPES AND F0 IN KIOWA

ANNA JURGENSEN

UNIVERSITY OF CALIFONRIA, BERKELEY

Abstract

Since the work of House and Fairbanks (1953) and Lehiste and Peterson (1961) it has been established that the prevocalic voiced stops produce downward perturbations and prevocalic voiceless stops produce upward perturbations on the fundamental frequency (F0) of the following vowel. The effects of stops with other laryngeal settings, such as aspirated and ejective, are not as clear (Hombert 1978). Additionally, the effects on F0 in tonal languages may differ from those in nontonal languages, both in the magnitude of the effect on vowels of different tone and in the duration of the effect in general (Gandour 1974). Using pre-recorded elicitations in Kiowa, a tonal, Kiowa-Tanoan language of southwestern Oklahoma, the effects of prevocalic voiced, voiceless unaspirated, voiceless aspirated, and ejective stops on the F0 of the following vowel were measured and analyzed. Using WaveSurfer, the F0 of the vowel was measured every 20 ms for a duration of 100 ms. The measurements were categorically averaged for comparison between types of stop and tone level. The analysis shows that voiced and voiceless stops in Kiowa produce the predicted downward and upward perturbations, and provides evidence that aspirated stops produce a higher perturbation on vowels than do voiceless unaspirated, and that ejective stops produce a downward perturbation on F0 instead of a neutral effect, as has been previously assumed (Hombert 1978).

It has been well established that the laryngeal setting of a prevocalic stop has an intrinsic effect on the fundamental frequency (F0) of the vowel that follows. Though the reason for the phenomenon is not completely understood, it has been established that unaspirated stops produce an upward perturbation on the vowel, while voiced stops produce a downward perturbation, regardless of the place of articulation of the stop (Hombert 1978, House & Fairbanks 1953, Lehiste & Peterson, 1961). Research, however, has not shown the effects of stop consonants with other laryngeal settings, such as aspirated and ejective, to be as consistent. Data from Korean (Kim 1968 as cited in Hombert 1978) and Thai (Ewan 1976 as cited in Hombert 1978) demonstrate that voiceless aspirated stops can produce a higher F0 perturbation on vowels than do voiceless unaspirated stops in the language. However, there is evidence of other patterns, even the opposite result: data comparing the perturbation caused by English aspirated and French unaspirated voiceless stops (Hombert 1978) suggests that aspirated and unaspirated voiceless stops can have similar effects on F0 of the following vowel, while other data in Chinese (Ching X. X. & Yi Xu 2003), and Thai (Gandour 1974) show that unaspirated voiceless stops can produce higher F0 perturbations than their aspirated counterparts. Meanwhile what little data exists on the effects of ejectives on F0 suggests that their effect is neutral, but Hombert admits that information on the subject is lacking (1978). Furthermore, studies on the magnitude of the effect on F0 on different tones in tonal languages (Hombert 1978, Drury 1989, Gandour 1974) have shown that the magnitude of the effect can vary in different tonal environments, with more exaggerated effects by voiced consonants (with downward perturbations) on high tones and by voiceless consonants (with upward perturbations) on low tones.

In addition to the variance in magnitude of the effect on F0, it has been shown that the duration of the effect varies. It has been demonstrated that the F0 perturbation can act on the vowel for as little as 30 ms, as has been observed after voiceless consonants in Thai (Gandour1974), or over 100ms, such as in English (Hombert 1978). From these results the conclusion thus far has been that the duration of the effect is minimized in tonal languages (Hombert 1975), though some evidence in Punjabi (see Mirza 1990) challenges this assumption.

Kiowa, a tonal language with a rich inventory of stops, can provide us with more information about these perturbatory effects of prevocalic stops on F0. Kiowa contrasts four series of stops: voiced, voiceless unaspirated (heron referred to as "voiceless"), voiceless aspirated (hereon referred to as "aspirated"), and ejective stops at velar, alveolar, and bilabial places of articulation (Crowell 1949, Harrington 1928, Wakins 1984, Siversten 1956). In addition, Kiowa contrasts 3 tones (high, low, and falling), providing a number of environments in which the phenomenon of F0 perturbation can be tested[1] (Siversten 1956, Watkins 1984).

Kiowa is a language of the Kiowa-Tanoan family with approximately 1,100 speakers in southwestern Oklahoma as of the 2000 census (Lewis 2009), and is one of the most well documented languages of its family (Sutton 2010). One such publication is the 1956 paper by linguist Eva Siversten: "Pitch Problems in Kiowa." In the study Siversten describes the Kiowa sound inventory and tonal system, and demonstrates Kiowa's tonal contrasts through minimal tonal pairs. The word lists of minimal tonal pairs Siversten used were the basis for the data used in this study, collected by Dr. Keith Johnson of University of California, Berkeley in June and July 2010. Dr. Johnson made these recordings with one female Kiowa speaker of Anadarko, Okalahoma. From the wordlists provided by Siversten (1956), words were elicited both individually and in context, imbedded in sentences created by the speaker.

---

[1] Additionally, Kiowa contrasts nasalized vowels with non-nasalized, as well as short and long vowel length (Siversten 1956, Watkins 1984), though these factors were not examined in this study.

Method

From the recordings provided by Dr. Johnson shorter .wav clips were created in Audacity for each of the cited words or elicited sentences to make Wavesurfer pitch analysis faster and the clip easier to observe in Wavesurfer. Then, from individual word and sentence sound files, the F0 contours of the vowels following the targeted stop segments (the 12 stops [b], [d], [g], [p], [t], [k], [p^h], [t^h], [k^h], [p'], [t'], [k']) were measured and recorded in a table. Additionally, each token was marked for either high or low tone (the reason falling tone was excluded is explained below), resulting in a total of 24 categories, each consisting of one of the 12 stops and either high or low tone, with all instances of downstepped high tone analyzed as high tone. Finally, the average frequency for each time interval for each category was calculated and plotted for comparison.

The F0 measurements themselves were made in Wavesurfer, using the ESPS pitch tracking method with an analysis window length of .0075s and a frame interval of .01s. The F0 measurements were taken in time increments of 20ms, from 20ms to 100ms after vowel onset. The vowel onset was defined as the moment in which both conditions of voicing and oral tract opening are met. For voiced stops, with a negative voiced onset time (VOT), and for those tokens of voiceless unaspirated stops with a VOT of 0, the stop burst was considered the onset of the vowel. For tokens of voiceless unaspirated stops with a positive VOT, and for voiceless aspirated stops and ejective stops which always have a positive VOT, the onset of voicing after the opening of the oral tract was considered the vowel onset. These conditions were determined for each clip through examination of both the waveform and the spectrogram in Wavesurfer.

Several problems surfaced in the analysis. First, Wavesurfer often produced questionable measurements at vowel onset (0 ms). Further, because manual measurements made on the

Stop	b		d		g		p		t		k	
Tone	H	L	H	L	H	L	H	L	H	L	H	L
Number of Tokens	28	29	49	66	41	60	29	19	44	50	48	51

Stop	p^h		t^h		k^h		p'		t'		k'	
Tone	H	L	H	L	H	L	H	L	H	L	H	L
Number of Tokens	18	9	45	41	42	25	11	20	44	27	41	17

Figure 1. Number of tokens across categories of stop consonant and high or low tone.

Wavesurfer spectrogram and waveform can only be measured to the nearest millisecond, manual measurements were too crude to be used for the purpose of this study.[2] The F0 measured contours, then, begin at 20ms after vowel onset instead of 0ms, the vowel onset itself.

Second, the number of recognizable falling tones in the recordings was very small for each stop, so the analysis only included observations on high and low tone, for which there was sufficient data. The possibility of studying the falling tone is mentioned below in the discussion.

Third, not all instances of voiced alveolar stop [d] were voiced. These instances of devoiced [d] were perceptually identical to the voiceless counterpart [t], and were identifiable only by a combination of the speaker's production of the voiced consonant [d] in other elicitations of the same word and Siversten's use of the voiced stop [d] in transcriptions (1956). Though the same occasional devoicing of a voiced stop was seen on the voiced velar stop [g], it was not nearly as common as it was on voiced alveolar stop [d]. All instances of devoicing on a stop with an underlying phoneme of voiced [d] were analyzed as voiced [d], and the effects of this decision can be seen in the results.

Lastly, because the wordlist was constructed to demonstrate tonal contrasts and not to elicit even numbers of tokens for each of the categories of stop and tone defined in this study, the distribution in the data differs for each category (see Figure 1). Velar stops had the widest

---

[2] Ladefoged (2003) states that it is necessary to measure a glottal pulse to the nearest tenth of a millisecond.

distribution across all laryngeal settings, while bilabial stops were by far the sparsest. Aspirated stops before low tones were also meager. Ultimately, many categories had at least 40 tokens, the exceptions being two velar stops [k'] and [kʰ] before low tones, the alveolar ejective [t'] before low tones, and all bilabial stops ([p'], [pʰ], [p], and [b]) before both high and low tones.

Results

The velar stops on a high tone, with the widest distribution in the data and therefore the largest number of tokens and also lacking the voiced/voiceless complication seen on voiced alveolar stop [d], displays most clearly the patterns of F0 perturbation across the various laryngeal settings of stops (see Figure 2). The comparison clearly shows the low average F0 and

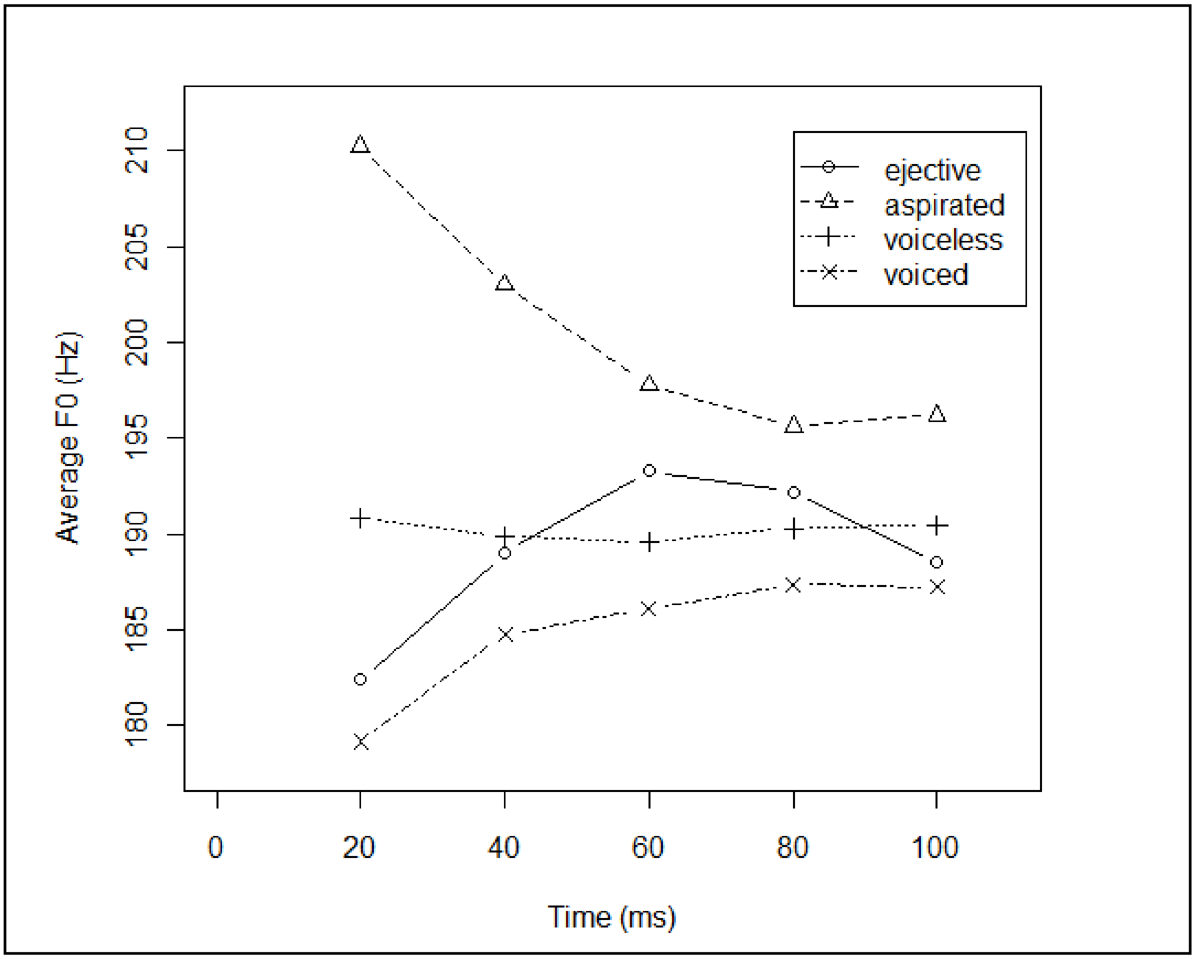

Figure 2. comparison of F0 contours on high tone following velar stops

Voiced		Voiceless		Aspirated		Ejective	
Average F0 (Hz)	F0 Range (Hz)	Average F0 (Hz)	F0 Range (Hz)	Average F0 (Hz)	F0 Range (Hz)	Average F0 (Hz)	F0 Range (Hz)
185	179-187	190	190-191	201	196-210	189	182-193

Figure 3. Average F0 (Hz) for each category of velar stop on vowels with high tone.

rising contour expected for voiced consonants and the high average F0 and the falling contour expected for voiceless consonants (see also Figure 3). Also as expected, the magnitude of the effect of the voiced consonant on a vowel of high tone is greater than that of the voiceless counterpart on vowel of high tone. This seen in steeper rise of the F0 contour after the voiced top [g]: a rise of 8 Hz (from 179 to 187 Hz) over 80 ms of time, compared to a fall of 1 Hz (from 191 to 190 Hz) following the voiceless counterpart [k]. The comparison also shows the aspirated stop [k^h] producing a much higher F0 and steeper falling contour than is produced by the voiceless, unaspirated stop [k]. Finally, and most interestingly, the effect of the ejective stop [k'] is clearly not neutral, as had been stated by Hombert (1978). Instead, the data shows a low perturbation with a steeply rising contour that peaks 60 ms after vowel onset, and gradually falling thereafter.

This patterning of perturbations (in which aspirated stops exhibit a very strong raising perturbation on F0, voiceless stops a less exaggerated raising perturbation, and voiced stops a lowering perturbation) is interesting for its contribution to the debate about the cause of the perturbation phenomenon itself. The patterns seen here are consistent with an aerodynamic theory for F0 perturbation, which theorizes that higher airflow and subglottal air pressure after voiceless stops, and especially after those aspirated, induces a strong Bernoulli Effect between the vocal folds which leads to a higher rate of vocal fold vibration (Hombert 1978). Further, it theorizes that high oral pressure at vowel onset following a voiced stop results in a lower transglottal pressure drop and thus a lower frequency (Hombert 1978). The exaggerated lowering effect seen on F0 following ejective stop [k'] may support the theory as well, given that ejective stops are articulated by compressing the air in the pharynx and intraoral cavity, which results in high supraglottal air pressure. However, the timing of glottal and oral gestures in the articulation of ejective stops must also be considered. Just as critics of the aerodynamic theory have provided experimental evidence that, while airflow at the release of aspirated stops is significantly higher than that at the release of voiced stops, by the time of voice onset the airflow following aspirated stops is actually lower than that of voiced stops, it is possible that the high supraglottal air pressure that exists at the release of an ejective stop is no longer significant at vowel onset (Ohala 1978). While it is not unreasonable to consider the possibility that supraglottal air pressure is still significantly high at the time of voice onset, the conditions of the supraglottal and subglottal pressure throughout the release of stops and the onset of voicing must first be examined before any definitive conclusions can be made regarding this data's support of an aerodynamic cause of perturbation.

Yet another interesting aspect of the F0 perturbation is seen in the data: the duration of the effect on the vowel. While most of the data in tonal languages has revealed shorter durations for the perturbation on F0 (as brief as 30 ms in Thai, see Gandour 1974), leading to the conclusion that the perturbation on F0 in tonal languages is shorter than in nontonal languages such as English, the data here shows significant differences between the contours even at 100ms after vowel onset: a duration comparable to that documented in English (Hombert 1978). Though this may initially seem anomalous, equally long durations of perturbation have also been observed in the tonal language Punjabi (Mirza 1990).

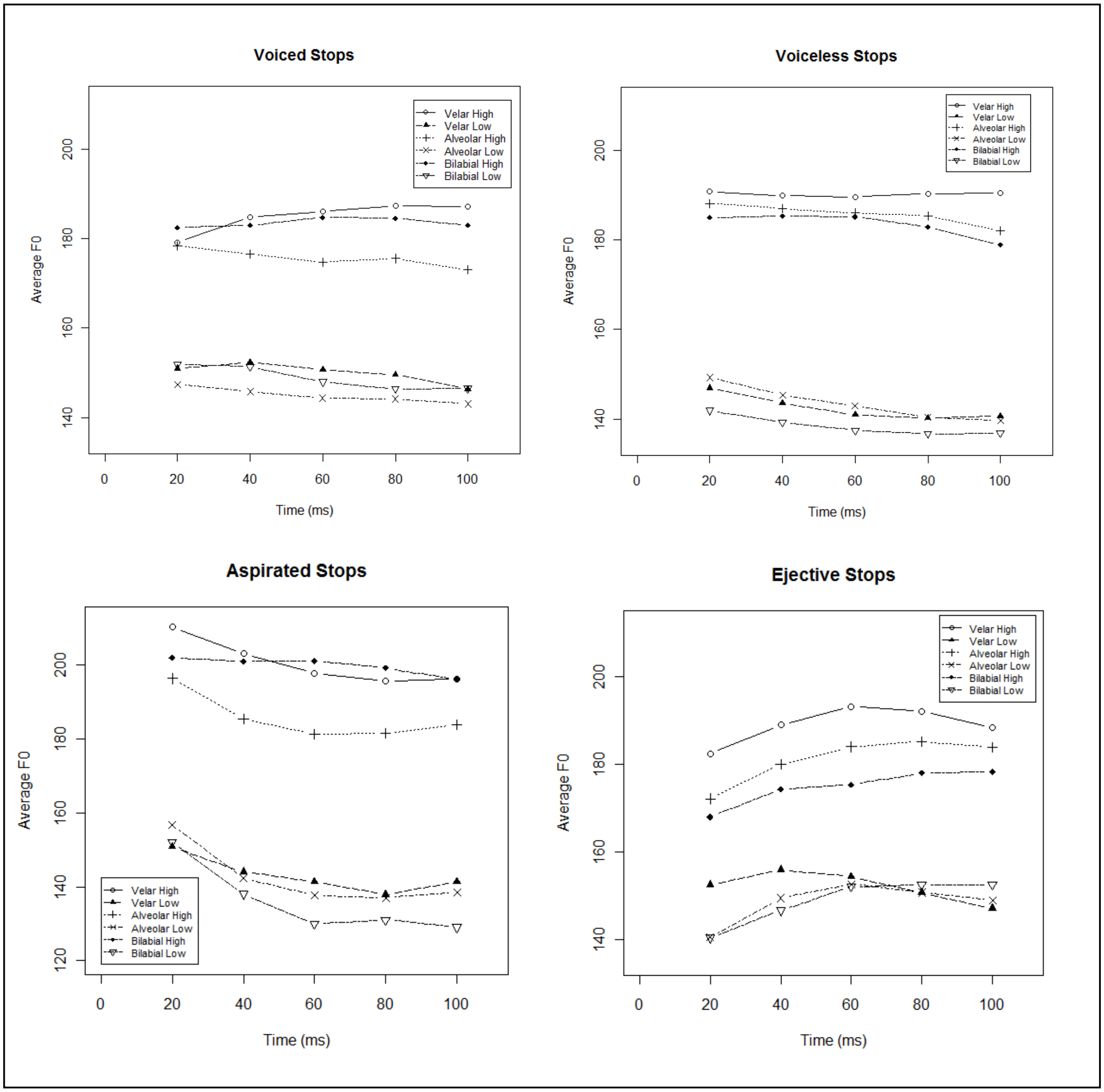

Figure 4. F0 (Hz) contours by laryngeal setting of the stop.

Moreover, the contours observed from the data of the other categories of stops and tones, despite complications and a smaller number of tokens, still generally resemble the contours that velar series produce on high tones (see Figure 4): that is, aspirated stops produce a steeper falling contour than do voiceless stops, and ejectives produce a contour that steeply rises in the first 60ms of the vowel (the exception in the data is the F0 contour for vowels of low tone following

	t	d	p	b	k	g
average F0 (Hz), high tone	186	175	183	183	190	185
average F0 (Hz), low tone	163	158	167	166	168	163

Figure 5. Comparison F0 following voiced and voiceless stops, averaged across the time period 20ms to 100ms after vowel onset.

bilabial ejectives: a category with too few tokens to be conclusive). In these results, however, the complications with devoicing of a voiced stop become apparent. While the voiced velar stop [g] produces the characteristic rising contour of F0 on a high tone, the rising contour is otherwise not demonstrated by the other voiced stops, or even by voiced stop [g] on a low tone. Examination of the average F0s following both the voiced and voiceless stops raises even more questions: while, as expected, the average F0s following the voiced stops [d] and [g] were lower than the average F0s following the corresponding voiceless stops [t] and [k], the average F0s following voiced and voiceless bilabial stops [b] and [p] are practically the same on both high and low tones (see Figure 5). Perhaps a closer examination of the first 20ms of the vowels, omitted here due to technical limitations, will reveal a more significant difference between the contours following voiced and voiceless stops. If not, these findings raise interesting questions about the production and perception of voiced versus voiceless stops in Kiowa. First, is this devoicing particular to this individual speaker or is exhibited by other speakers as well? If it is a common process in Kiowa, are the small differences in F0 (in these averages, 5-11Hz difference if a difference occurs at all) perceptible to a Kiowa speaker? Additionally, what is the primary cue that a listener uses to determine if the underlying phoneme is the voiced or voiceless stop: is the voiced onset time the primary cue (and the perturbations on F0, if perceived, redundant cues: see Abramson & Lisker (1985) and Whalen, Abramson, Lisker, & Mody (1993)), or are the differences in F0 not only perceived, but also used as the primary cues (see Pearce (2009))?

Discussion

Though this study encountered some complications, the findings are promising and show a need for further research beyond the previously suggested research questions dealing with F0 perturbation, perception, and primary/redundant cues. The examination of tone provided here can be refined and repeated for a better understanding of what appear to be interesting phenomena: the unusually long duration of effect on the F0 of vowels, and the particular F0 contour that is produced by ejective stops. Using a more accurate method of measuring fundamental frequency, especially at vowel onset, will create a more complete contour and perhaps lend better understanding to the voiced/voiceless stop problem that was encountered. Additionally, using word lists created with the aim of eliciting large number of tokens from all stop/tone categories (including the effects on vowels of falling tone) will render more balanced and complete data. The parameters for categories can also be changed or expanded, as Kiowa also contrasts short versus long vowel length and nasalized versus non-nasalized vowels. Finally, future study should work with more than one speaker to gain a better understanding of patterns in the language in general.

References

Abramson, A. S. & Lisker, L. (1985). Relative power of cues: F0 shift versus voice timing. In. Fromkin, V. (1985). *Phonetic linguistics: Essays in honor of Peter Ladefoged.* 25-33.

Anderson, S. R. (1978). Tone features. In Fromkin, V. A. (Ed.). 133-175.

Ching X. X. & Yi Xu (2003). Effects of consonant aspiration on Mandarin tones [abstract]. *Journal of the International Phonetic Association, 33*, 165-181. doi:10.1017/S0025100303001270

Crowell, E. (1949). A preliminary report on Kiowa structure. *International Journal of American Linguistics* 15(3). 163-167.

Drury, J. (1989). Consonant types and tone in Siamese revisited [abstract]. *Acoustical Society of America 85* (S1), S52-S53. doi:10.1121/1.2027017

Ewan, W. G. (1976). *Laryngeal behavior in speech*. Ph.D. dissertation, University of California, Berkeley.

Fromkin, V. A. (1978). *Tone, A linguistic survey.* San Diego, California: Academic Press, Inc.

Gandour, J. T. (1974). Consonant types and tone in Siamese. *Journal of Phonetics, 2,* 337-350.

Halle, M. & Stevens, K. (1971). A note on laryngeal features. *Quarterly Progress Report*, MIT Research Laboratory of Electronics, *101*, 198-213.

Harrington, J.P. (1928). *Vocabulary of the Kiowa language*. Washington: U.S. Government Printing Office.

Hombert, J-M. (1978). Consonant types, vowel quality, and tone. In Fromkin, V. A. (Ed.). 77-111.

House, A. & Fairbanks, G. (1953). The influence of consonant environment upon the secondary acoustical characteristics of vowels. In Lass, N. (Ed.). 79-99.

Kim, C.W. Review of Lieberman 1967. *Language*, *44*(4), 830-842.

Ladefoged, P. (2003). *Phonetic data analysis: An introduction to fieldwork and instrumental techniques.* Malden, MA: Blackwell Publishing.

Ladefoged, P. & Maddieson, I. (1996). *The sounds of the world's languages.* Malden, MA: Blackwell Publishing.

Lass, N. (Ed.) (1974). *Experimental phonetics.* New York: MSS Information Corp.

Lehiste, I. & Peterson, G. E. (1961). Some basic considerations in the analysis of intonation. *The Journal of the Acoustic Society of America*, *33*(4). 419-425. doi:10.1121/1.1908681

Lewis, M. P. (Ed.) (2009). *Ethnologue: Languages of the world*, Sixteenth edition. Dallas, Tex.: SIL International. Retrieved from: http://www.ethnologue.com/.

Mirza, J.S. (1990). F0 perturbnation effect of prevocalic stops on Punjabi tones. *Proceedings of the Australasian Speech Science and Technology Association*, 90, 400-405.

Ohala, J.J. (1978). Production of tone. In Fromkin, V. A. (Ed.). 5-39.

Pearce, M. (2009). Kera tone and voicing interaction. *Lingua* 119 (6), 846-864.

Sivertsen, E. (1956). Pitch Problems in Kiowa. *International Journal of American Linguistics, 22*(2), 117-130.

Sutton, L. (2010). Noun class and number in Kiowa-Tanoan: Comparative historical research and respecting speakers' rights in fieldwork. In Berez, A. L., Mulder, J., & Rosenblum, D. *Fieldwork and Linguistic Analysis in Indigenous Languages of the Americas* (57-90). Honolulu: University of Hawai'i Press.

Watkins, L.J. (1984). *A grammar of Kiowa*. Lincoln: University of Nebraska Press.Whalen, D. H., Abramson, A. S., & Lisker, L. & Mody, M. (1993). F0 gives voicing information even with unambiguous voice onset times. *Journal of the Acoustic Society of America* *93*(4), 2152-2159.