
Volume 119, Issue 6, June 2009
ISSN 0024-3841
ELSEVIER
Lingua
International Review of General Linguistics
Special Issue
Tone and Intonation from a Typological Perspective
Guest-edited by S. Zerbian, L.J. Downing and F. Kügler
Lingua Tables of Contents delivered free to your desktop: Register for ScienceDirect Alerts
Available online at
ScienceDirect
www.sciencedirect.com

Available online at www.sciencedirect.com
ScienceDirect
Lingua 119 (2009) 827–845
Lingua
www.elsevier.com/locate/lingua

Kurtöp Tone: A tonogenetic case study

Gwendolyn Hyslop *

1290, University of Oregon, Eugene, OR 97403, USA

Received 19 February 2007; received in revised form 12 November 2007; accepted 17 November 2007
Available online 21 May 2008

Abstract

Languages of Asia are well-known for undergoing tonogenesis in recent history. The model established by Haudricourt (1954), Matisoff (1970), Thurgood (2002) and others is well-attested within the Tibeto-Burman language family, amongst others. While the phenomenon is well-documented in history, we rarely have the occasion to document tonogenesis in detail as it occurs. Though we know that loss of codas generally brings tonal contrasts into a language and from there voiceless consonantal onsets condition a high tonal split while voiced onsets condition a low tonal split, the details of such processes remain unknown. Kurtöp, a Tibeto-Burman language of Bhutan, is currently undergoing tonogenesis, providing the unique opportunity to analyze the phenomenon in detail. In this article we explore Kurtöp tonogenetic properties via comparative data and an acoustic study. We find that tone first entered Kurtöp following sonorant consonant onsets, spread to the palatal fricative and appears to be spreading following the remainder of the obstruents.

Keywords: Tonogenesis; Tibeto-Burman; Languages of Bhutan; Kurtöp; Areal influence; Sonority

1. Introduction

Some of the mechanisms and motivations underlying tonogenesis have been established over the last several decades by such pioneering work as Maspero (1912), Haudricourt (1954), Matisoff (1970), Mazaudon (1977), Hombert (1978), Matisoff (1999), Kingston (2004) and many others. Conventional wisdom suggests that tone usually enters a language via lost coda consonants which condition contour tones. Later, the tones may be split, with high register being diachronically conditioned by voiceless initials and low register being conditioned by voiced initials. Thurgood (2002) recently updated the model by arguing that voice quality plays a mediating role in tonogenesis. That is, between a contrast in voicing on a consonant and tone on a

* Tel.: +1 541 346 3906; fax: +1 541 346 5961.
E-mail address: glow@uoregon.edu.

0024-3841/$ – see front matter
doi:10.1016/j.lingua.2007.11.012

vowel, an intermediate stage of contrastive voice quality on the vowel is present. However, despite these advances in the study of tonogenesis, many questions remain unanswered. For example, the manner in which the sound change occurs remains unknown. We do not yet know whether the sound change happens across all segments at the same time or same rate, or whether certain segments are likely to undergo tonogenesis first.

The aim of this article is to fill the described gap in tonogenetic studies. Kurtöp, a Tibeto-Burman language of Bhutan, is currently undergoing tonogenesis and thus provides the opportunity to examine the process as it occurs. As we will argue, the results of this study suggest that Kurtöp is gradually acquiring tone following the consonantal onsets in word-initial position, commencing with the sonorants. We illustrate in detail that tone has entered the language following the sonorant consonants and is now proceeding through the system following the obstruents. We will demonstrate that tone has phonologized first in Kurtöp following the nasal and liquid consonantal onsets in initial syllables. The next step in the process was for Kurtöp to develop tone following the palatal fricative. The remainder of the obstruents in the language is now in place to undergo tonogenesis.

An investigation of tonogenetic properties of other languages may suggest that Kurtöp is not the only language to acquire tone in this manner. Lhasa Tibetan and arguably some Tai languages may have acquired tone in the manner described for Kurtöp, that is, by phonologizing tone following the sonorants before phonologizing tone following the obstruents.

Because many speakers of Kurtöp are also speakers of Dzongkha (the national language of Bhutan), which is tonal, one could argue that Kurtöp tonogenesis is a contact-induced phenomenon. However, regardless of whether tonogenesis in Kurtöp is contact-induced or motivated by other language-internal factors, Kurtöp tonogenetic properties merit further investigation.

This article is organized as follows. Section 2 provides background information on Kurtöp and describes the development of the synchronic tonal system over a period of time. Section 3 details the methodology, design and results of an acoustic experiment which examines the relationship between voicing of stop onset and pitch on the following vowel. Section 4 offers a discussion of the results. In section 5 we provide a summary and a brief conclusion is offered in section 6. We will conclude that Kurtöp has phonologized tone first following sonorant consonants, extended the tonogenesis to palatal fricatives, and that tone is now incipient following the other obstruents.

2. Kurtöp

Kurtöp is a Tibeto-Burman language spoken by approximately 10,000 people in the Kurtö region of Lhüntse province, surrounding Dungkar in northeastern Bhutan (van Driem, 1995), as illustrated in Fig. 1.

Kurtöp has been addressed very minimally in the literature. Van Driem (2001) offers a few lexical items and Michailovksy and Mazaudon (1994) provide the first phonemic analysis of Kurtöp. Their findings are for the most part corroborated by the present study, but for more details see Lowes (2006) and Hyslop (2008a). One difference in particular – the status of palatal fricatives – will be relevant for the present tonogenetic case study and will therefore be addressed in greater detail below.

In terms of genealogy, Kurtöp is an East-Bodish language within the larger Tibeto-Burman family. East-Bodish languages are considered to be closely related to, but not directly descended from Classical Tibetan (see Lowes, 2006; Hyslop, 2008a for more information on the relationship between Kurtöp and Classical Tibetan).

Fig. 1. The Kurtöp region in Bhutan.

2.1. Synchronic tonal system

Synchronically, tone in Kurtöp is contrastive on the first syllable following the sonorant consonants and the palatal fricative.[1] Tone in these instances is realized by either high or low pitch. The data in (1) below illustrate the contrast following sonorant consonants.

(1)

nɑ̀	'ear'	ɲèwɑ	'relatives'
nɑ́	'nose'	ɲéwɑ	'hell'
mì	'person'	ŋɑ̀	'drum'
mí:	'aim'	ŋɑ́:	'pillow'
wɑ̀ŋ	'hole'	jɑ̀ŋ	'to stand.'
wɑ́ŋ	'blessing'	jɑ́ŋ	'prosperous in wealth'
lɑ̀	'mountain'	rɑ̀	'root'
lɑ́	'month'	rɑ́	'hair'

Tone is also contrastive on the first syllable of words following the palatal fricative. This contrast is illustrated by the data in (2).

(2)

çɑ̀	'what'
çɑ́	'meat'

[1] At present, two words have also been found with salient rising contour tones: [wǎ:] 'honey bee' and [jy̌:] 'rain', the origins of which will not be further pursued in this study.

Table 1
Written Tibetan correspondences for high tone following nasals and liquids in Kurtöp

Kurtöp	Gloss	Written Tibetan	Kurtöp	Gloss	Written Tibetan
ŋɑ̀	'drum'	རྔ་ *rŋa*	nɑ̀	'ear'	རྣ་བ་ *rna-ba*
ŋɑ́:	'pillow'	སྔས་ *sŋas*	nɑ́	'nose'	སྣ་ *sna*
rɑ́	'hair'	སྐྲ་ *skra*			
rɑ̀	'root'				

Following all other consonants tone is not contrastive. That is, in the environment following obstruents except the palatal fricative, tone is predictably high when following the voiceless obstruents and predictably low when following the voiced obstruents. This correlation is illustrated by the data in (3).

(3)

pɑ́	'pork slice'	bɑ̀	'target'
p^{h}ɑ́t	'leech'		
tɑ́	'axe'	dɑ̀	expletive
t^{h}ɑ́	weaving pattern		
ʈɑ́	'change of color'	ɖɑ̀	'to be good'
ʈ^{h}ɑ́ŋ	'to climb'		
cɑ́ro	'friend'	ɟɑ̀	'tea'
c^{h}ɑ́	'pair'		
kɑ́	'snow'	gɑ̀	'saddle'
k^{h}ɑ́	'language, mouth'		
tsɑ́	'nerves'		
ts^{h}ɑ́	'salt'		
sɑ́	'soil'	zɑ̀m	'bridge'

The voiceless palatal fricative is the only obstruent which does not have a voiced or aspirated counterpart and is further the only obstruent to occur preceding both high and low tone.

2.2. *Diachronic development of Kurtöp Tone*

Comparative evidence suggests that tone following the sonorants has developed in Kurtöp by means of historically present onset clusters in which the first member was a voiceless fricative /s/. The historical pattern of a complex onset is reflected in the Written Tibetan forms but Kurtöp words have reduced their onsets and added high tone to the nucleus. In Table 1 we compare Kurtöp forms with their cognates in Written Tibetan. Note that where Kurtöp forms have a high tone, an *s*- initial onset cluster is present in Written Tibetan.[2]

The sound change in which high tone is conditioned by an *s*- initial sonorant onset cluster can perhaps be envisioned in two steps. In the first step the *s*- sonorant cluster yields a voiceless

[2] Note we have not found a Classical Tibetan cognate for Kurtöp 'root'. Whether the Kurtöp form is innovative or is a retention remains unknown. The possibility that a cognate form existed in Classical Tibetan but does not appear in our sources cannot be ruled out either.

sonorant. The voiceless sonorant then conditions high pitch on the following vowel, according to the established model. At this point in the language a contrast would exist between voiceless and voiced sonorants, with high tone predictably following the voiceless sonorants and low tone predictably following the voiced sonorants.[3] Over time, a second sound change would occur in which high tone phonologizes following the voiceless series, low tone phonologizes following the voiced series, and the voicing distinction is neutralized in favor of voiced sonorants. Other motivations for tonogenesis following sonorants remain less clear.

Comparative data suggesting the source of high tone on other sonorants are displayed in Table 2. Possible phonetic motivations for the tonogenesis suggested by the data in Table 2 are less clear, though there is general agreement that the handful of Tibetan forms with initial <zl> represent some idiosyncratic Tibetan-internal development from earlier forms with **s*-. The sound change /db/ > /w/ with a high tone on the following vowel is also characteristic of modern Tibetan dialects.[4]

Table 2
Written Tibetan correspondences for Kurtöp words with a high tone following a liquid and approximate

Kurtöp	Gloss	Written Tibetan	Kurtöp	Gloss	Written Tibetan
lɑ̀	'mountain'	ལ་*la*	wɑŋ	'hole'	
lɑ́	'month'	ཟླ་*zla*	wɑ́ŋ	'blessing'	བདེ་འགྲོ་ *bde-ˎgro*

Note: No cognate for Kurtöp 'hole' has been found.

A third plausible means by which Kurtöp has obtained tone is areal influence. As Kurtöp has borrowed a large amount of its vocabulary from the national language Dzongkha (a tonal language), it may be more fitting to propose that Kurtöp tonogenesis is a contact-induced phenomenon. Under this hypothesis, as loan words with tones were borrowed into Kurtöp, tone eventually became a component of the phonology of Kurtöp. However, even in this scenario an explanation would need to be sought in order to motivate the current synchronic presence of contrastive tone in Kurtöp following only the sonorants and palatal fricative. Regardless of the source of tone in Kurtöp – via one of the possible acoustic motivations, borrowing, or a source not mentioned here – the fact remains that tone has first phonologized following the sonorant consonants. It is this observation we believe to be significant.

While the source for tone following the sonorant consonants may be debatable, the source for tone following the palatal fricative appears straight-forward. The tone following the palatal fricative has developed directly via the loss of contrast in voicing. Evidence for this development comes in two forms. First, Michailovksy and Mazaudon (1994) reported a voiceless and voiced palatal fricative in Kurtöp but no contrastive tone following either. The instances in which they report a voiced palatal fricative we find a voiceless palatal fricative with low tone. Note that

[3] A contrast between voiceless and voiced sonorants is not uncommon for Tibeto-Burman languages, and often the high tone follows only the voice less series. Dzongkha (van Driem, 1998), for example, has a voiceless rhotic and lateral which precede only high tone while high and low contrast following voiced sonorants.

[4] Contrastive high tone in Lhasa Tibetan and Dzongkha differs from Kurtöp in that it has developed by way of any onset cluster. The first step in the development of Dzongkha and Lhasa Tibetan tone involved the initial member of the onset cluster devoicing, which could have invoked a sound change similar to the one described above for the Kurtöp sonorants. That is, a voiceless initial could have perturbed higher pitch on the following vowel, which would have then phonologized as tone while the initial member in the cluster disappeared.

Table 3
Comparative palatal fricatives in Kurtöp and Tshangla

Kurtöp	Gloss	Tshangla	Kurtöp	Gloss	Tshangla
ɕòr	‘wine’	zu	ɕònbɑ	‘young’	ʑonmɑ

Michailovksy and Mazaudon (1994) collected their data in the 1970s. The approximately 30 years which have passed between the two studies may be taken to represent generational differences; that is, perhaps the generation of Kurtöp speakers represented by Michailovksy and Mazaudon (1994) had the voicing contrast in the palatal fricative but the generations considered today have neutralized the contrast in favor of tone on the following vowel.[5] The second line of support in favor of the argument that Kurtöp has neutralized a voicing contrast on the palatal fricatives in favor of a contrast in tone on the following vowel comes from comparative evidence, as shown by the data in Table 3.

Tshangla is a Tibeto-Burman language spoken to the east of the Kurtöp language area in Bhutan and Arunachal Pradesh in India. In many instances where Tshangla has a voiced palatal fricative, Kurtöp has a voiceless palatal fricative with low tone on the following vowel. This, in conjunction with the fact that for at least one variant of Kurtöp in the 1970’s (Michailovksy and Mazaudon, 1994) a voiced palatal fricative corresponds with a voiceless low-toned palatal fricative in the dialect of Kurtöp represented in this study, suggests that tonogenesis following the palatal fricative is more recent than the genesis of tone following the sonorants.[6,7]

While data from Tshangla (Andvik, 2003) and Kurtöp (Michailovksy and Mazaudon, 1994) provide synchronic evidence for the source of tone following the fricatives, evidence for the source of tone following the sonorants is found only in comparison with written forms of Tibetan, which we presume represent an older synchronic state of Classical Tibetan. Variation amongst the palatal fricatives exists between Kurtöp and neighboring languages and has been noted in a prior publication on Kurtöp (Michailovksy and Mazaudon, 1994), while variation amongst the sonorants is not found in either. Thus, we argue that the development of

[5] We do not need to suggest that tonogenesis has completed for the palatal fricative in the past approximately 30 years. Though spoken by a small community, Kurtöp purports a handful of mutually intelligible dialects. Michailovksy and Mazaudon (1994) do not mention where in Kurtö their speakers come from and therefore we do not know which variety of Kurtöp they spoke. It is entirely plausible that the dialect represented in their study is different from that discussed here. As we have not completed a full dialect survey, the possibility remains that some dialects of Kurtöp have retained a voicing contrast for palatal fricatives.

[6] In instances where tone develops via a contrast of voice it is often the case that phonation is an intermediate contrast as Thurgood (2002) articulates. However, it is not clear this is the case in Kurtöp even though it appears to be in Dzongkha. For example, in Dzongkha, historically voiced consonants may be followed by high or low tone. In instances with low tone, a salient feature is breathy voice on the following vowel, often with a concomitant devoicing of the initial (van Driem, 1998, personal field notes). Such salient breathy voice is not audibly present in Kurtöp, though no acoustic measures for breathy voiced such as H1–H2 or H1–F2 (Gordon and Ladefoged, 2001) have been taken.

[7] Note that the direction of the change in voicing is reversed from that proposed for the sonorants. That is, while we posited a mediating stage of voiceless sonorants which voiced as part of the tonogenesis process, here we see a consonant become voiceless with tonogenesis. We do not see this as a problem; in both instances the neutralization of a voicing contrast is in favor of markedness—voiceless sonorants are more marked than voiced sonorants and voiced obstruents are more marked than voiceless obstruents.

tone as triggered by the palatal fricative has come after the phonologization of tone following the sonorants.

3. Experimental study

In this section we investigate the correlation between tone (measured as fundamental frequency) and voicing (measured as voice onset time[8]). The focus of this study was on the production of the stop consonantal onset and vowel in monosyllabic words. Mean fundamental frequency was computed across the duration of the vowels, taking mean and standard deviation of *f*0 at the approximate vowel midpoint. Mean and standard deviation of voice onset time (VOT; Lisker and Abramson, 1964) of stops was also measured. The goals were (1) to determine whether the observation that high and low tones correlated with voiceless and voiced obstruents,[9] respectively, held true across the entirety of the vowel; (2) determine whether the high and low tones would represent statistically distinct categories; (3) determine mean and standard deviation of VOT for the three voicing categories of stops (voiceless, aspirated, voiced) and mode for the voiced series; and (4) ascertain whether the VOT means represent significantly disparate categories.

If Kurtöp obstruents are undergoing tonogenesis we would predict high tone to phonologize following voiceless obstruents, low tone to phonologize following the voiced series, and that the voiced series of obstruents would be devoicing. For the present experiment, then, we predict that *f*0 measurements will display significantly distinct categories on vowels following voiceless versus voiced stops. If Kurtöp voiced stops are collapsing with the voiceless series we might expect some utterances of voiced stops to be realized as voiceless. If voiced stops were at times realized as voiceless stops we would find VOT values associated with the voiceless series of stops alongside the negative VOT values expected for the voiced stops. Therefore, if Kurtöp were neutralizing a contrast in voice on stops we predict this would manifest a very high standard deviation from the mean VOT of voiced stops and possibly a bimodal distribution.

3.1. Speakers

Data from male native speakers of Kurtöp were recorded and analyzed. The first speaker, P.C., was in his 20s at the time of the study and is from Tabi, within Dungkar. K.W. is the second speaker, is in his 40s, and is from Thuke, within Dungkar. Both speakers P.C. and K.W. are also fluent speakers of Dzongkha and English. The speakers were chosen due to their proximity to the researcher (both resided in the western United States at the time of study) and it is by accident that both happen to be males but of different generations. However, the fact that speaker K.W. is approximately 20 years older than speaker P.C. will be of interest when we discuss the findings in terms of sound change.

---

[8] While we have not done perception studies confirming this observation, it is our impression that voice onset time is the primary cue to voicing in Kurtöp. We are basing this conclusion on two observations. First, acoustic measurements not mentioned in this article have shown no salient distinction in other possible cues, such as duration of closure or vowel length preceding voiceless versus voiced consonants, for example. Second, mean VOT is statistically significant for the categories of voiceless unaspirated compared to voiceless aspirated (cf. section 3.3.2, this article), suggesting that Kurtöp could also employ VOT as a means by which to distinguish the voiced category from the other two categories of voice.

[9] In order to simplify wording, here and throughout the remainder of sections 3 and 4, when we refer to 'obstruents' we are excluding the palatal fricative, which has already devoiced and triggered tonogenesis.

3.2. Methodology

A total of 1,041 monosyllabic stop-initial tokens were recorded and analyzed acoustically for *f*0 on the vowel and VOT of onset. In order to control for possible word stress or tone variation in multi-syllabic words, only monosyllabic words have been chosen for this study. The attempt was made to design a list of tokens which were equally balanced for place (bilabial, dental, retroflex, palatal, velar), and voicing (voiceless, aspirated, voiced) of stop, while also controlling for the quality of the following vowel (non low front, non low back, low). Because vowel quality can minimally influence *f*0, it was hoped that by controlling for quality, the current study would rule out the possibility that vowel quality had influenced the results. We examined only stop consonants in this study but expect the results we find would extend to the entire category of obstruents which have not already undergone tonogenesis. This expectation is based on our impressions and observations that high and low tones also follow voiceless and voiced affricates and fricatives.

Both speakers produced each target word in the study four times: three times in isolation and a fourth time in the carrier phrase shown below in (4).

(4)	*ngai*	*dangnin* ———	*lab-mi*
	1ST.ERG	yesterday ———	say-PFCTV
	‘I said ———	yesterday’	

Each utterance of the word was included in the acoustic analysis, yielding a total of four tokens for every word. List intonation was often associated with the three words in isolation; a rising contour was often present on the first token and a falling contour was often present on the third. This was true regardless of whether the word began with a voiceless, aspirated or voiced onset. Because we were interested in mean *f*0, and not contour of the pitch, we did not exclude any tokens on the basis of intonation. By systematically including each of the four utterances for a given word, rather than choosing one utterance, for example, we were able to increase the overall number of tokens for each category. We assume that by systematically including all tokens in our analyses, any effect intonation might have on mean *f*0 would be consistent throughout the voicing categories and therefore not affect the overall results.

However, due to unforeseen difficulties in gathering the data, there were some gaps and the data were not completely balanced for place, voicing and vowel quality. Of considerable importance is the fact that for speaker K.W. voiced retroflex tokens were entirely lacking. At times the speakers repeated an incorrect word, in which case the word was not counted. During the recording a few iterations were omitted, also reducing the number of tokens in a given category. Sometimes additional words in a particular category were recorded, leading to categories with a greater number of tokens. In total, 610 (155 words) tokens were analyzed for speaker P.C. and 431 (108 words) tokens were analyzed for speaker K.W. Despite the lack of precise balance in the tokens analyzed for this study, we believe the main argument of this article remains tenable. The number of tokens for each place of articulation combined with voice category is listed by speaker in Table 4.

All recordings were done using a head-mounted Shure brand microphone, placed approximately 3 cm from the speaker’s mouth. The data were recorded at a sampling rate of 22.05 kHz into a Marantz PMD 660 flash digital recorder and saved as .wav files on a computer. All acoustic analyses on the tokens were carried out using Praat (Boersma and Weenink, 2007) phonetics software.

Table 4
Total number of tokens analyzed in acoustic study, organized according to voicing type for each speaker

Speaker	# Voiceless tokens	# Aspirated tokens	# Voiced tokens	# Combined tokens
P.C.	194	223	194	611
K.W.	142	155	133	430
Total	336	379	326	1041

Voice onset time was measured between the first voicing cycle and the initial release of the stop. We computed the measurement by hand, using the computer cursor to identify the initial release of the stop and the first voicing cycle. Frication was sometimes present in the velar and especially in the palatal stops. This frication was always included in the measurement of VOT. Fundamental frequency (*f*0) on each vowel was measured using a script at eight equidistant points on the vowel, beginning with the first glottal pulse.

3.3. Results

The results of this acoustic study demonstrate that (1) the high tone following voiceless stops and low tone following voiced stops is maintained across the duration of the vowel and that (2) these tones are statistically significant categories. The study also (3) calculated mean and standard deviation for VOT of all three stop types (voiceless unaspirated, voiceless aspirated, voiced); and while the standard deviation for the voiced categories, especially, was quite high, the results of this study also show that (4) VOT measurement for each stop type is a statistically significant category. However, the results of this study also illustrate a trend for the voiced category of stops to be merging with the voiceless category of stops. The results for fundamental frequency following each stop type will be considered first, followed by an examination of the VOT results.

3.3.1. Fundamental frequency

Figs. 2 and 3 represent the fundamental frequency on vowels following voiceless unaspirated, voiceless aspirated and voiced tokens for speakers P.C. and K.W., respectively. Recall that speaker P.C. is one generation younger than speaker K.W.

Figs. 2 and 3 offer a visual illustration of the fact that both speakers demonstrate a clearly disparate *f*0 on vowels following voiced stops compared to when the vowel is following voiceless

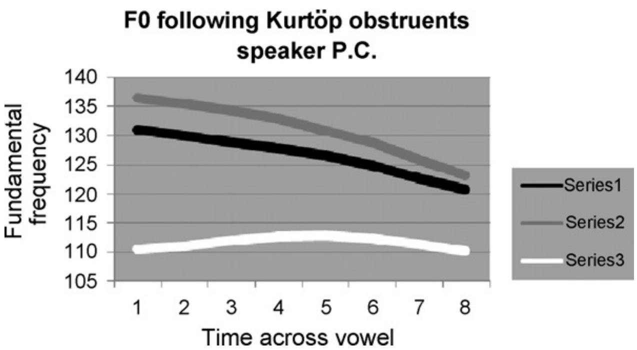

Fig. 2. Mean *f*0 (610 tokens) on vowels following obstruents for speaker P.C. Series 1 represents mean *f*0 on vowels following aspirated stops; series 2 represents *f*0 following voiceless stops and series represents *f*0 following voiced stops.

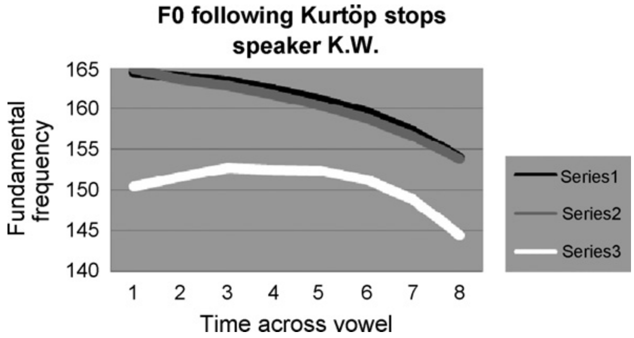

Fig. 3. Mean *f*0 (431 tokens) on vowels following stops for speaker K.W. Series 1 represents mean *f*0 on vowels following aspirated stops; series 2 represents *f*0 following voiceless stops and series 3 represents *f*0 following voiced stops.

stops. Note also that this difference is maintained across the entire length of the vowel, not neutralizing at some point, as would be expected in a simple intrinsic difference. Hombert (1978), for example, found that English speakers showed a great degree of individual differences with respect to fundamental frequency following voiceless versus voiced obstruents, but the averaged result showed a drastic decrease of the intrinsic pitch difference over time. Our results illustrate that the difference in *f*0 is consistently maintained across the entire duration of the vowel for both speakers.

Statistical analysis also confirms an effect of consonant voice type on *f*0 at the near mid point of the vowel (time interval 4). A univariate analysis of variance showed [$F(2,610) = 60.95$, $p < .001$] for P.C. and [$F(2,429) = 26.261$, $p < .001$] for K.W. A Tukey's HSD post hoc test confirmed that *f*0 mean values at the fourth interval on the following vowel for the categories of voiceless unaspirated and voiceless aspirated were both statistically distinct from the *f*0 measure following the voiced category of stops ($p < .01$). The difference of *f*0 following voiceless unaspirated stops versus that following voiceless aspirated stops for P.C., the younger speakers, was significant ($p = .017$) but not significant for K.W., the older speaker ($p = .946$). This difference is also evident in a visual comparison of Figs. 2 and 3.

3.3.2. *Voice onset time*

Mean and standard deviation of voice onset time for voiceless unaspirated, voiceless aspirated and voiced stop tokens were calculated for both speakers. The results are illustrated in Table 5.

Glancing at the values underneath the mean columns for the speakers we see evidence that voiceless unaspirated, voiceless aspirated, and voiced stops are separate categories. For speaker K.W. the mean VOT of voiceless unaspirated stops was +39.65 ms; the mean VOT of voiceless aspirated stops was +71.75 ms; and the mean VOT of voiced stops was −41.12 ms. These values are similar to those reported for the younger speaker P.C. For speaker P.C. the mean VOT of voiceless unaspirated stops was +36.62 ms; the mean VOT of voiceless aspirated stops was +89.00 ms; and the mean VOT for voiced stops was −50.11 ms.

In terms of the current study, however, as we are interested in sound change, it will be useful to consider the variation within these means. A large degree of variation could be suggestive of a change in progress. Considering the data presented in the columns underneath standard deviation in Table 5, we note that both speakers tended to display the least variation for the voiceless unaspirated category, with an overall standard deviation of 27.16 for speaker K.W. and 23.93 for speaker P.C. Considering each place of articulation separately, we note that the lowest reported

Table 5
VOT summary for P.C. and K.W. Number, mean, standard deviation, and range values are shown for each speaker and each stop type (place × voice)

		Speaker—P.C.				Speaker—K.W.			
POA	Voice type	*N*	Mean	S.D.	Range	*N*	Mean	S.D.	Range
Labial	Unaspirated	48	19.04	8.7	9 to 57	12	25.26	3.57	19 to 31
	Aspirated	60	77.42	18.66	41 to 133	36	65.02	16.82	33 to 105
	Voiced	44	−55.63	56.56	−160 to 33	36	−74.27	40.14	−176 to −14
Dental	Unaspirated	28	26.93	22.72	11 to 109	32	22.63	5.66	14 to 42
	Aspirated	24	73.45	22.62	33 to 123	28	59.05	16.06	36 to 90
	Voiced	20	−74.23	75.56	−253 to 16	24	−62.23	30.43	−142 to −21
Retroflex	Unaspirated	28	31.68	13	14 to 67	36	22.23	6.97	12 to 50
	Aspirated	28	75.23	15.63	47 to 108	32	60.62	14.09	35 to 88
	Voiced	40	−55.08	57.94	−137 to 38				
Palatal	Unaspirated	27	70.61	23.11	44 to 146	28	75.27	33.44	44 to 150
	Aspirated	44	115.9	27.03	77 to 187	28	99.31	22.31	75 to 172
	Voiced	40	−33.87	57.2	−115 to 65	36	−32.48	61.22	−156 to 72
Velar	Unaspirated	35	46.22	13.25	19 to 76	33	50.17	16.2	19 to 78
	Aspirated	30	97.99	19.35	55 to 133	35	76.97	16.53	38 to 117
	Voiced	31	−41.47	65.55	−211 to 52	36	−11.27	49.96	−104 to 76
Total	Unaspirated	166	36.62	23.93	9 to 146	141	39.65	27.16	12 to 150
	Aspirated	186	89	26.97	33 to 187	159	71.75	22.23	33 to 172
	Voiced	179	−50.11	61.5	−253 to 65	132	−41.12	55.74	−176 to 76

The results are representative of 610 tokens for speaker P.C. and 429 tokens for speaker K.W.

standard deviation was 3.57 (speaker K.W. labials) and the highest was 33.44 (speaker K.W. palatals).

The aspirated category of stops displayed slightly more variation than the unaspirated category. Amongst aspirated stops, the overall standard deviation for speaker K.W. was 22.23 and for speaker P.C. the overall standard deviation was 26.97. In the aspirated category, the lowest standard deviation of 14.09 was found amongst speaker K.W.'s retroflex stops while the highest standard deviation of 27.03 was found amongst speaker P.C.'s palatal stops.

For both speakers the voiced category of stops displayed the greatest amount of variation, with a standard deviation of 55.74 for speaker K.W. and 61.5 for speaker P.C. Amongst the voiced stops, speaker K.W.'s dentals displayed the least amount of variation with a standard deviation of 30.43, while the same set – dentals – for speaker P.C. displayed the most variation, with a standard deviation of 75.56.

Despite the large standard deviations, the difference between the three groups is statistically significant for both speakers: [$F(2,610) = 60.95$, $p < .001$] for speaker P.C. and [$F(2,429) = 349.504$, $p < .001$] for speaker K.W. A Tukey HSD post hoc test confirmed $p < .01$ for both speakers for each of the three possible pairwise comparisons of the three stop types.

On the other hand, it could be argued that voiced segments inherently display more variation than voiceless segments. However, a consideration of VOT in other languages suggests the variation in Kurtöp is unusual. For example, Lisker and Abramson (1964) reported mean and range of VOT in eleven languages, two of which report a three-way contrast in voicing similar to

Kurtöp. Thai and Armenian both contrast voiceless unaspirated, voiceless aspirated and voiced stops. The ranges reported by Lisker and Abramson (1964:396) for the Thai and Armenian voiced category were always negative. That is, the Thai and Armenian voiced stops in their study were always prevoiced. In Kurtöp, however, with the exception of speaker K.W.'s dental stops, the voiced categories always revealed iterations with positive VOT values. Histograms displaying the distribution of VOT for each stop type for each speaker will enable us to visually compare the findings across categories. Consider Figs. 4 and 5.

These histograms illustrate that the distribution of VOT of voiced stops indeed differs from the distribution of VOT of the other stop categories. Let us consider first the relationship between the distribution of voiceless aspirated (series 1) and voiceless unaspirated (series 2) values. There is some overlap between these two categories for both speakers. However, the histograms offer a visual illustration of the findings that distinct means were found for the voiceless aspirated and voiceless unaspirated categories for both speakers, despite the overlap.

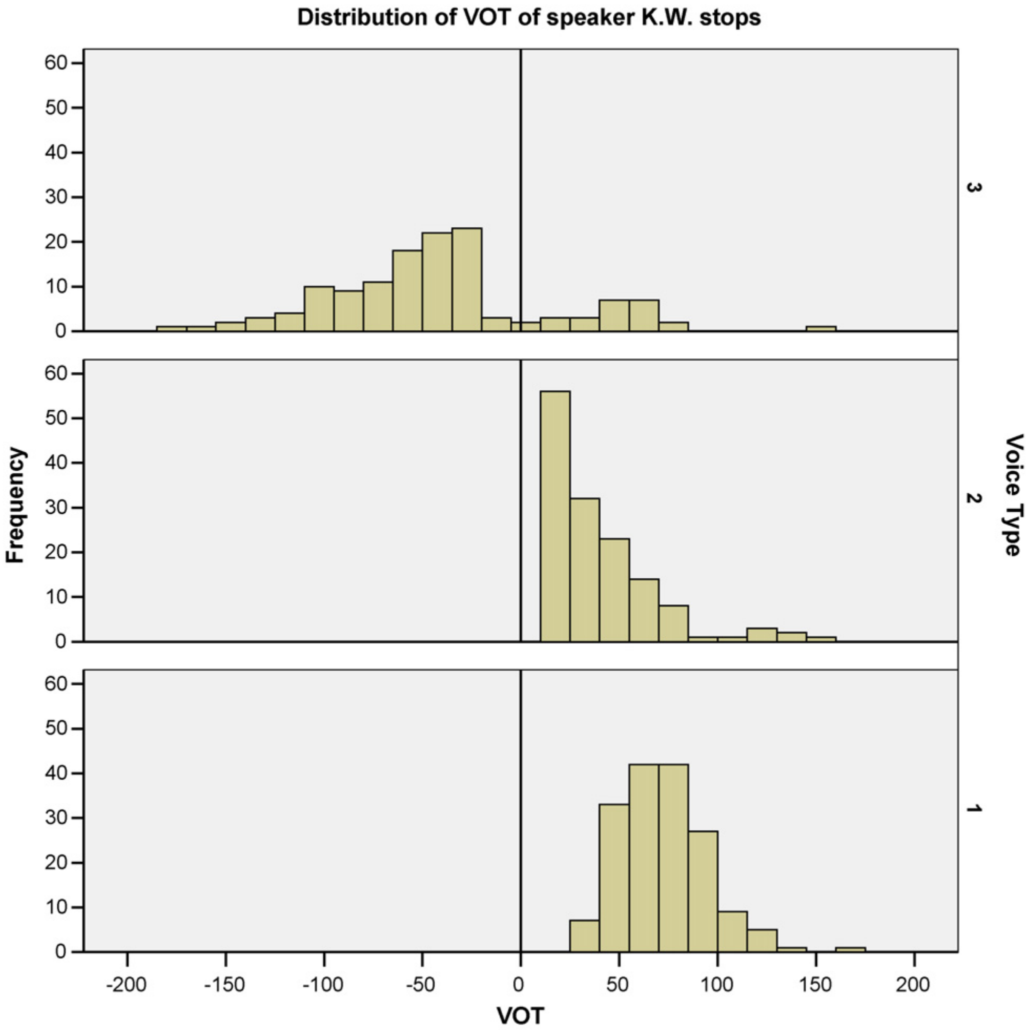

Fig. 4. Histogram displaying distribution of VOT values for speaker K.W. stops. Series 1 represents voiceless aspirated stops; series 2 represents voiceless unaspirated; and series 3 represents the voiced series of stops. Each bar represents an interval of 15 ms.

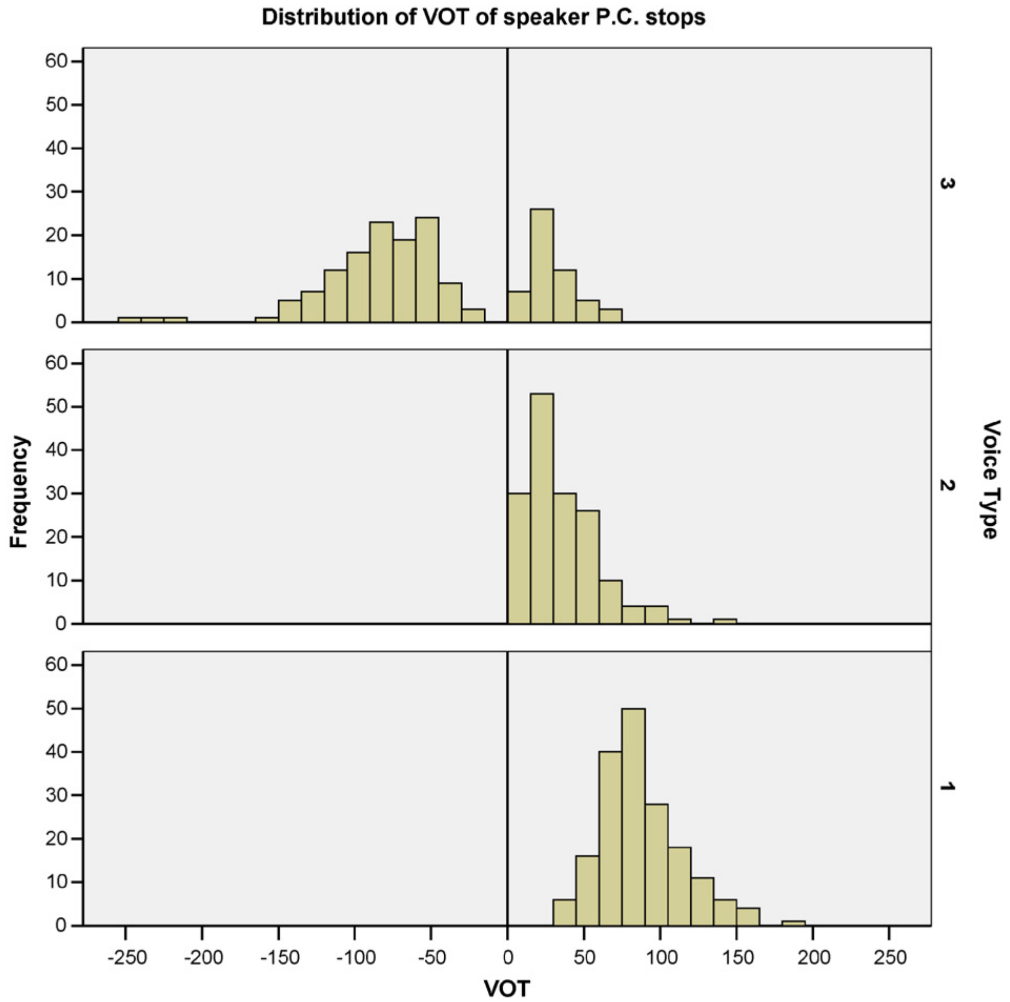

Fig. 5. Histogram displaying distribution of VOT values for speaker P.C. stops. Series 1 represents voiceless aspirated stops; series 2 represents voiceless unaspirated; and series 3 represents the voiced series of stops. Each bar represents an interval of 15 ms.

The distribution of voiced VOT values, on the other hand, demonstrates much more overlap with the voiceless categories. Further, the distribution of voiced stop VOT values is much greater than either the voiceless aspirated or the voiceless aspirated category. VOT values for voiced stops are distributed widely between values of mainly −150 ms and +100 ms, with a few outliers. For the younger speaker P.C., this overlap is even more pronounced, with a second mode apparently overlapping closely with the mode illustrated for the voiceless unaspirated stops.

Because we are particularly interested in the possibility that the category of voiced stops is merging with the category of voiceless stops as part of tonogenesis, let us consider exclusively the distribution of voiced VOT for both speakers in Figs. 6 and 7.

Fig. 6 displays the frequency of VOT values of voiced stops for the older speaker, K.W. The majority of the utterances have a negative VOT. It is noteworthy, however, that a second mode appears to be developing in the positive VOT range, around +50 ms. This trend is even stronger

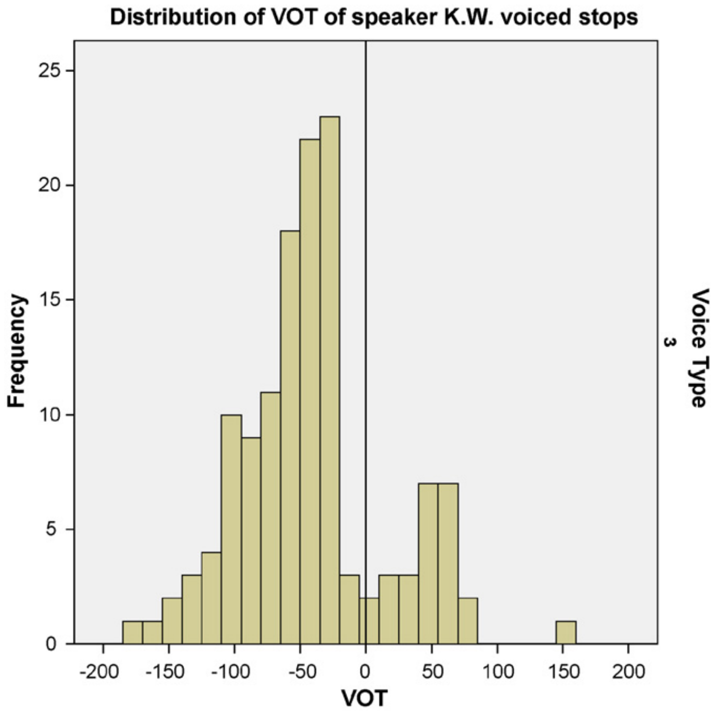

Fig. 6. Histogram displaying distribution of VOT values for speaker K.W. voiced stops. Each bar represents an interval of 15 ms.

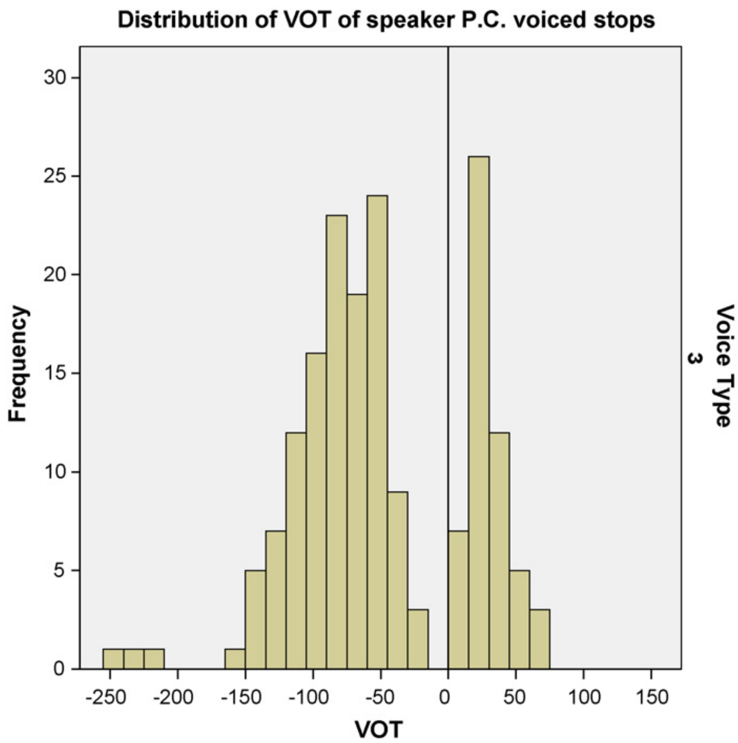

Fig. 7. Histogram displaying distribution of VOT values for speaker P.C. voiced stops. Each bar represents an interval of 15 ms.

for the younger speaker, P.C., as illustrated in Fig. 7. This figure shows a clear bimodal distribution, with one mode at approximately −70 ms and the other mode around +30 ms. Recall that mean VOT values amongst voiceless unaspirated stops were recorded between +35 ms and +40 ms for both speakers.

4. Discussion

The acoustic study has confirmed the impression that high tone follows voiceless stops and low tone follows voiced stops. We expect this difference to be true of the entire series of obstruents which have not already undergone tonogenesis. The difference in *f*0 between the two tones ranged from approximately 12 Hz for the older speaker K.W. to approximately 20 Hz for the younger speaker P.C. The difference between high and low tone was maintained across the entirety of the vowel.

Mean *f*0 was calculated at the approximate midpoint (time interval four) for both speakers and a univariate analysis of variance confirmed the two mean values (high tone following voiceless and low tone following voiced) to be statistically distinct. We argue these results suggest that tone is a salient property of these words' production. That is, we argue that these results suggest high tone has phonologized following voiceless obstruents and low tone has phonologized following voiced obstruents.

In this experimental study we also computed mean and standard deviation of VOT of voiceless unaspirated, voiceless aspirated, and voiced stops. Because we are interested in researching the possibility that the voiced category of obstruents is devoicing and merging with the voiceless unaspirated stops as part of the tonogenesis process, we will discuss only the voiceless unaspirated and voiced results here.

For the older speaker K.W., the mean VOT of voiceless unaspirated stops was +39.65 ms (S.D. 27.16). For the same speaker, the mean VOT of voiced stops was −41.12 ms (S.D. 55.74). Findings were similar for the younger speaker P.C., who had a mean VOT of +36.62 ms (S.D. 23.93) for voiceless unaspirated stops and a mean VOT of −50.11 ms (S.D. 61.5) for voiced stops. This large amount of deviation from the mean for both speakers is indicative of the great deal of variation within the realization of both speakers' category of [+voice] as VOT. Figs. 4 and 5 displayed the distribution of voiced VOT values for both speakers by means of a histogram. For the older speaker K.W., we found one mode around −60 ms with a possible second mode emerging around +50 ms. Fig. 5 illustrated an even stronger trend for the younger speaker P.C. to have a bimodal distribution, with one mode centered around −70 ms and the second emerging around +30 ms. Based on these findings we can speculate that the large amount of variation within the voiced category of stops is attributed to the trend for voiced stops to be realized with positive VOT values. We argue this tendency suggests that a negative VOT is no longer the primary characteristic in the production of the voiced series in stops.

A bimodal trend is indeed what we would expect if the category of voiced stops was merging with the category of voiceless stops in favor of a contrast in tone on the following vowel. Sound change is not instantaneous and it thus follows that when voiced stops are merging with voiceless stops there would exist an intermediate stage in which some iterations are voiceless and some are voiced. As the language places phonemic importance on tone and decreases the importance in voicing, voiced stops are free to devoice and merge with their less marked voiceless counterparts. Such variation within the voiced category of stops is precisely what the VOT study has shown. That is, these results illustrate a trend for ''voiced' stops to be realized with the positive VOT associated with voiceless stops.

The combined findings of the experimental study suggest that tone has phonologized following obstruents; high tone has phonologized following voiceless obstruents and low tone has phonologized following voiced obstruents. The results further suggest that a three-way contrast in voicing of stops is collapsing in favor of a two-way contrast, and specifically that the voiced series is merging with the voiceless unaspirated series. These findings are consistent with the argument that Kurtöp obstruents are undergoing tonogenesis.

We must be clear that the nature of these findings represent the synchronic state of Kurtöp as spoken by speakers K.W. and P.C., and cannot be taken as absolute predictors of future sound change. The findings in our acoustic study indeed suggest that *f*0 is a salient acoustic cue on the vowel following voiceless versus voiced stops. Our study also suggests that the contrast between voiced and voiceless unaspirated stops is neutralizing, with the voiced category merging with that of the voiceless unaspirated. However, it is clear that absolute neutralization has not taken place as voiced stops are still often produced with a negative VOT. We are unable to say whether or not voiced stops will completely merge with voiceless unaspirated stops in the future.

Despite the fact that we are not able to know the future outcomes of sound change, we can consider the results of both speakers separately, as representing subsequent generations of synchronic stages of Kurtöp, and make a prediction. While both speakers displayed the results described in this study (disparate tone following voiced versus voiceless stops and the tendency to realize voiced stops with the VOT associated with voiceless unaspirated stops), the younger speaker P.C. consistently displayed results more consistent with the notion that tone, and not voicing, was the important category. That is, the difference in Hertz of high versus low tone following voiceless unaspirated versus voiced obstruents was greater for speaker P.C. (25–10 Hz) than for the older speaker K.W. (15–8 Hz). The younger speaker P.C. also showed a stronger trend to produce voiced stops with the VOT associated with voiceless unaspirated stops.

Indeed, if our speakers can be taken to represent different stages of Kurtöp diachronically, then we have further evidence that tonogenesis is occurring in the language. Crucially, both speakers have disparate tones associated with the vowels following voiceless and voiced stops. The older speaker K.W. illustrated a tendency to realize voiced stops with the VOT associated with voiceless stops. The younger speaker is even more pronounced in displaying this trend. If this trend continues, perhaps we can expect the next generation of Kurtöp speakers to completely neutralize the contrast of voiced versus voiceless stops altogether.

5. Summary

Kurtöp provides a unique opportunity to examine a gradual tonogenesis in progress and in doing so we have seen that tone has entered the language following the sonorant consonants. The fact that tone first phonologized in Kurtöp following sonorant consonant onsets is illustrated by the synchronic state of the language. Comparative evidence suggests that, at least in the case of the nasals, high tone has been conditioned by **s*- initial members of sonorant onset clusters in a historically attested stage of the language. The conditioning environment for high tone following the remaining sonorant consonants remains unclear but we feel this does not detract from the central point, which is that in the synchronic state of the language tone is contrastive only following the sonorants.

In the time since the development of tone following the sonorant consonants, the palatal fricative has collapsed its contrast in voicing in lieu of a tonal contrast on the following vowel.

Our main source of evidence for this ordering (that is, that tonogenesis following the palatal fricative came after the development of tone after sonorant onsets) is the fact that there has been variation reported regarding the palatal fricative in Kurtöp (Michailovksy and Mazaudon, 1994) and in the related language Tshangla (Andvik, 2003), where the voiced segment is still present. However, such variation is not found amongst the sonorants in Tshangla or in previous descriptions of Kurtöp. Recall that Michailovksy and Mazaudon (1994) reported a voiced palatal for Kurtöp where we find only the voiceless palatal fricative with ensuing low tone but otherwise found the same tonal system. Whether the variation represents a completed sound change, dialectal or speaker differences, while an interesting question, does not affect the fact that there has been variation reported for this segment while none has been observed amongst the nasals.

The experimental study and results described in section 3 suggest that the remainder of the obstruents is also in place to undergo tonogenesis. Pitch, measured at the rough midpoint, on vowels following the voiceless obstruents is statistically higher than when following voiced obstruents. Graphs representing pitch on vowels following voiceless versus voiced obstruents provided a visual representation of the fact that tone is higher (10–25 Hz) across the entire length of the vowel following voiceless stops compared to when following voiced stops. The study also examined VOT and found that the voiced obstruents displayed greater variation than would be expected, suggesting that VOT is no longer the primary cue for voiced segments. More importantly, VOT values for both speakers' voiced stops suggest a bimodal distribution, supporting the idea that the category of [+voice] is being replaced by the category of [−voiced]. The fact that the bimodal distribution is more exaggerated for speaker P.C. is consistent with his trend for tone to be a more distinct category. That is, it appears that tonogenesis amongst the obstruents is further progressed for the younger speaker P.C. than for the older speaker K.W.

To summarize, for Kurtöp, it appears tonogenesis is a gradual process; tone has developed first following the sonorant consonants, spread to the palatal fricative and now is spreading to the remainder of the obstruents. Kurtöp is unusual in that it is developing tone for the first time following onsets, rather than codas, which is the most common pathway for tone to enter a language for the first time. As Kingston (2004) notes, initial tonogenesis in a given language triggered by onsets is unusual and is reported primarily in the context of areal influence. Kurtöp is likely under influence from Dzongkha, a tonal language, and thus could also be considered as an example of contact-induced tonogenesis.

That tone enters languages following the sonorants first has also been reported for other languages of Asia. Tshangla, a Tibeto-Burman language of eastern Bhutan and western Arunachal Pradesh in India (Andvik, 1999, 2003), exhibits tone following only the sonorants but in some dialects the obstruents are also conditioning tonogenesis. Andvik (personal communication) further reports that the palatal fricative is the only segment amongst the obstruents to have triggered tonogenesis. Mazaudon (1977) claims that tone first phonologized in Tibetan following the historically prefixed nasals and resonants. She states that this has also happened in Tawang, a Tibeto-Burman language of Nepal (personal communication). Finally, in Tai languages it has been suggested that the shift from voiceless nasals to tone on the following vowel preceded the shift from a voiced contrast in obstruents to a tonal contrast on vowels (L-Thongkum, 1997). The findings of this study, in light of tonogenesis reported for Tibetan, Tshangla and Tai, suggest that sonority may play a role in tonogenesis, or to be more explicit we suggest that sonorants tend to phonologize tone on their following vowels before a contrast in voice is neutralized in favor of tone. Indeed, further research examining the role of sonority in tonogenesis may prove fruitful.

6. Conclusion

This paper has described the tonogenetic properties of Kurtöp, a Tibeto-Burman language of Bhutan, based on data from two native speakers. We have demonstrated that Kurtöp has phonologized pitch following the sonorants and palatal fricative onsets first. Evidence supporting a relatively recent tonogenetic sound change for the latter have been found in the form of a voiced palatal fricative reported for a variety of Kurtöp studied in the 1970s (Michailovksy and Mazaudon, 1994) and in the related and geographically proximate language Tshangla (Andvik, 2003). Results from the acoustic study have shown that tone is in position to replace a contrast in voice following the remainder of the stops, a trend we predict would extend to all obstruents.

Kurtöp offers itself as another example of a language undergoing tonogenesis triggered by onsets, rather than codas, probably in light of areal influence with Dzongkha, a tonal language. The fact that historically attested sonorant clusters, rather than obstruents, have been the initial genesis of tone is at first blush an unusual find. However, given the evidence that tonogenesis has begun in the context of nasals in other languages, further research considering the role of sonority in tonogenesis could demonstrate that Kurtöp tonogenesis is not as unusual as it may appear.

Acknowledgements

I would like to express my gratitude first and foremost to Pema Chhophyel and Kezang Wangchuk for patiently sharing their language with me. This work has also benefited from discussion with Scott DeLancey, Spike Gildea, Lisa Redford and Susan Guion. This paper was presented in different incarnations at the 2006 summer meeting of the Linguistic Society of America and the second conference on the Typology of Tone and Intonation (TIE 2). Discussions with participants of these conferences have been particularly fruitful. I am grateful especially for the comments of Mark Aronoff, Marc Brunelle, Bob Ladd and Tomas Riad, which helped form the ideas in this article. John Kingston read a previous draft of this paper and provided valuable input. The comments and advice of three anonymous reviewers and the editors of this volume also contributed greatly to this article. Finally, I am grateful to the Linguistic Society of America for financial assistance and the Graduate School at the University of Oregon for a Gary E. Smith Summer Grant, both of which made travel to the LSA summer meeting and TIE 2 possible. Any errors or misrepresentation of facts are the sole responsibility of the author.

References

Andvik, E., 1999. Tshangla grammar. Ph.D. Dissertation. University of Oregon, Eugene.

Andvik, E., 2003. Tshangla. In: Thurgood, G., LaPolla, R. (Eds.), The Sino-Tibetan Languages. Routledge, London/New York, pp. 439–455.

Gordon, M., Ladefoged, P., 2001. Phonation types: a cross-linguistic overview. Journal of Phonetics 29 (4), 383–406.

Haudricourt, A., 1954. De l'origine des tons en Vietnamien. Journal Asiatique 242, 69–82.

Hombert, J.-M., 1978. Consonant types, vowel quality and tone. In: Fromkin, V. (Ed.), Tone: A Linguistic Survey. Academic Press, New York, pp. 77–111.

Hyslop, G., 2008a. Kurtöp phonology in the context of Northeast India. In: Morey, S., Post, M. (Eds.), North East Indian Linguistics 1: Papers from the First International Conference of the North East Indian Linguistic Society. Cambridge University Press, Cambridge, pp. 3–25.

Hyslop, G., 2008b. Kurtöp and the classification of the languages of Bhutan. In: Proceedings from the 42nd Annual Meeting of the Chicago Linguistic Society 42, vol. 2, South Asian Linguistics, Case, Voice, and Language Coexistence. University of Chicago Press, Chicago.

Kingston, J., 2004. Mechanisms of tone reversal. In: Kaji, S. (Ed.), Proceedings of the Third Symposium on Cross-linguistic Studies of Tone, Institute of the Study of the Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies, Tokyo, pp. 57–120.

Lisker, L., Abramson, A., 1964. A cross-language study of voicing in initial stops: acoustical measurements. Word 20, 384–422.

L-Thongkum, T., 1997. In: Edmondson, J., Solnit, D. (Eds.), Implications of the Retention of Proto-Voiced Plosives and Fricatives in the Dai Tho Language of Yunnan Province for a Theory of Tonal Development and Tai Language Classification. SIL Publishers, Dallas, pp. 191–219.

Lowes [Hyslop], G., 2006. Kurtoep phonetics and phonology. Master's Thesis. University of Oregon, Eugene.

Maspero, H., 1912. Études sur la phonétique historique de la langue Annamite: les initiales. Bulletin de l'Ecole Francaise d'Extreme Orient 12, 1–126.

Matisoff, J., 1970. Glottal dissimilation and the Lahu high-rising tone: a tonogenetic study. Journal of the American Oriental Society 90.1, 13–44.

Matisoff, J., 1999. Tibeto-Burman tonology in an areal context. In: Shigeki, K. (Ed.), Proceedings of the Symposium 'Cross-Linguistic Studies of Tonal Phenomena: Tonogenesis, Typology, and Related Topics', Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies, Tokyo, pp. 3–32.

Mazaudon, M., 1977. Tibeto-Burman tonogenetics. Linguistics of the Tibeto-Burman Area 3.2, 1–123.

Michailovksy, B., Mazaudon, M., 1994. Preliminary notes on the languages of the Bumthang group. In: Kvaerne, P. (Ed.), Tibetan Studies: Proceedings from the Sixth Seminar of the International Association for Tibetan Studies, vol. 2. The Institute for Comparative Research in Human Culture, Oslo, pp. 545–557.

Thurgood, G., 2002. Vietnamese tonogenesis: revising the model and the analysis. Diachronica 19.2, 333–363.

van Driem, G., 1995. Een eerste grammaticale verkenning van het Bumthang, een tal van Midden-Bhutan, met een overzicht van der talen en volkeren van Bhutan. Centrum voor Niet-Westerse Studiën, Leiden.

van Driem, G., 1998. Dzongkha. Research School CNWS, School of Asia, African, and Amerindian Studies, Leiden.

van Driem, G., 2001. Languages of the Himalayas. Brill Publishers, Leiden.