INTERSPEECH 2016
September 8–12, 2016, San Francisco, USA

The Role of Pitch in Punjabi Word Identification

Jasmeen Kanwal[1], *Amanda Ritchart*[2]

[1]University of Edinburgh
[2]UC San Diego
jasmeen.kanwal@ed.ac.uk, aritchart@ucsd.edu

Abstract

Previous work has argued that one class of consonants in Punjabi—those thought to be historically voiced aspirated—have now lost aspiration in all contexts and voicing in certain contexts. Word initially, these consonants are realized as voiceless unaspirated and are differentiated from other voiceless unaspirated consonants by a falling pitch on the following vowel. In this study, we investigate, using a two-alternative forced choice task, whether listeners make use of a falling pitch word-initially to distinguish between these two types of consonants that are otherwise phonetically identical. Our results show that, regardless of talker or listener, differences in falling pitch on the vowel following an unaspirated voiceless consonant are indeed sufficient for listeners to distinguish between words beginning with these consonants. These results provide further evidence that, in word-initial contexts, pitch may be in the process of phonologization in at least some dialects of Punjabi.
Index Terms: Punjabi, F0, pitch, perception, tonogenesis, tone, sound change

1. Introduction

Punjabi is one of the few Indo-European languages reported to have lexical tone. It has approximately 130 million native speakers, concentrated in Northern India and Pakistan, but spread worldwide by a large diaspora population [1]. The tone patterns described in Punjabi are particularly interesting because they are linked to only one class of consonants: those that are believed to have once been voiced aspirated but that are now reportedly realized without aspiration (and sometimes without voicing) and with an added tone.

The set of Punjabi consonants which are relevant for our discussion are the 20 oral stops and affricates, shown in Table 1. In the table, 'T' symbolizes voiceless unaspirated manner, 'TH' voiceless aspirated, 'D' voiced unaspirated, and 'DH' are the consonants which were historically voiced aspirated, but whose current phonetic properties have been contested. The evidence for the historic realization of these consonants as voiced aspirated can be seen in the Gurmukhi script, used by speakers in the Indian state of Punjab as the primary means of writing Punjabi, as well as in cognates found in closely related languages such as Hindi. Further evidence may come from other languages in northern India, such as Mandeali and Northern Haryanvi, that are at intermediate stages of losing the voiced aspirated consonants [2].

Previous descriptive sources conflicted with one another on the phonetic realisation of the DH consonants. First, [3, 4, 5] claim that aspiration no longer occurs, but [6] claims that some remaining traces of breathiness can be observed in vowels preceding DH consonants. Second, it is thought that voicing is also lost, but only word-initially [3, 4, 5]. Finally, a three-tone system is commonly proposed, in which DH consonants induce low, low-rising, or falling tone in the following vowel; high, rising-falling, or rising tone in the preceding vowel; and mid tone elsewhere [3, 4, 5, 6, 7, 8]. As these previous claims are either impressionistic or based on limited data, we conducted a careful production study using a comprehensive data set and modern phonetic analysis tools with the aim of providing a definitive phonetic characterization of these consonants [9].

Table 1: Punjabi oral stops and affricates. Each cell is represented orthographically by a single unique symbol in the Gurmukhi script. For the DH words, the first transcription represents the consonant's realization word-initially, while the second transcription represents the consonant's realization word-medially and word-finally, as found in [9].

	'T'	'TH'	'D'	'DH'
bilabial	p	p^h	b	p/b
dental	t	t^h	d	t/d
retroflex	ʈ	$ʈ^h$	ɖ	ʈ/ɖ
palatal	ʧ	$ʧ^h$	ʤ	ʧ/ʤ
velar	k	k^h	g	k/g

Table 2: A near-minimal set containing word-initial consonants with the same place of articulation but varying in manner. The 'DH' example's phonetic realization is based on the findings of [9].

	'T'	'TH'	'D'	'DH'
Punjabi (IPA)	koɽa	k^hora	goɽa	kôɽa
English gloss	*whip*	*rough*	*white*	*horse*

Our findings confirmed that these DH consonants are now realized as unaspirated, and, word-initially, also as voiceless (Figure 1A). We also found that these consonants induce a falling F0 on the following vowel, but only in word-initial position. All other stop manners induced no distinctive F0 contour in the following vowel (Figure 1B). Furthermore, DH consonants in word-medial position also induced no distinctive F0 contour in following vowels. Noting that Punjabi words are usually stressed word-initially [7], this pattern of restricted F0 modulation closely resembles languages with emerging tone or "pitch accent" [2]. See Table 2 for an example minimal set contrasting the phonetic realization of the DH consonants with other consonants word-initially.

The emergence of tone, or tonogenesis, is thought to be initiated in some cases by F0 perturbations caused by articulatory constraints during production of a sound [10]. These perturbations can be reproduced and exaggerated by hearer after hearer, eventually leading to the reanalysis of the original sound as con-

http://dx.doi.org/10.21437/Interspeech.2016-1445

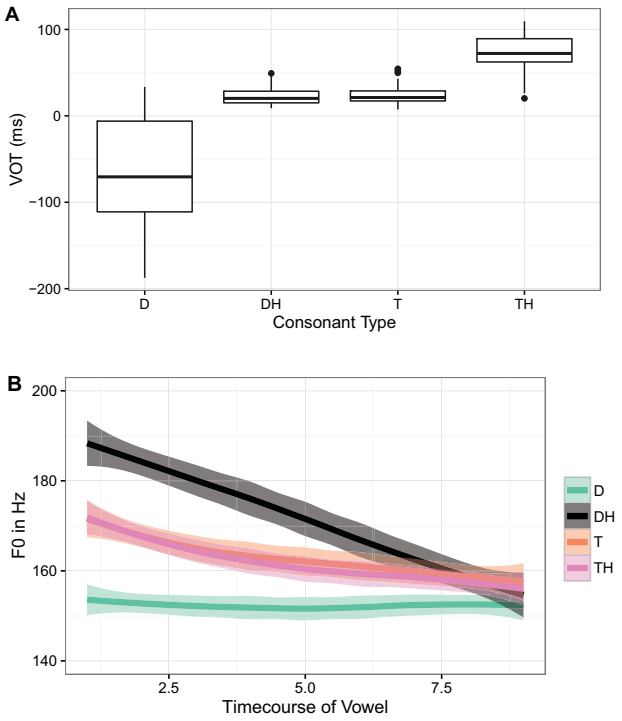


Figure 1: Results from [9]. (A) Voice Onset Time for consonants in word-initial position. DH words have the same VOT as T words (both are voiceless unaspirated). (B) Average F0 of vowels in disyllabic words (with 50% CIs due to large between-sex differences) following word-initial consonants. DH words have a falling F0 contour.

taining an F0 change as part of its phonological profile. In [9], we discuss possible articulatory mechanisms which lead to a lowering of F0 following voiced aspirated consonants in prominent syllables. If the observed F0 change were due to coarticulation, we should expect it to die out towards the end of the vowel. This is true for any changes occurring after T, TH, and D, but not generally after DH. Furthermore, coarticulatory considerations lead us to expect a low onset F0 after DH, with the F0 rising to meet the endpoint of the other consonants. Instead we see a high onset F0 after DH that lowers throughout the duration of the vowel, suggesting that there is an F0 specification on the vowel independent of coarticulatory effects (Figure 1B).

Our findings thus show that clear pitch contrasts are being produced by speakers across the relevant contexts. However, we cannot determine from production data alone whether these pitch contrasts have reached the next stage of becoming perceptually salient to speakers, aiding in word discrimination and thus perhaps eventually achieving phonological status. Therefore, to determine whether listeners use F0 to identify stop manner in Punjabi, we present a perception study which investigates the ability of speakers to discriminate between minimal pairs of words which differ primarily in terms of their F0 contours.

2. Methods

2.1. Participants

A total of 20 participants participated in the study (11 males, 9 females). All participants were recruited through the Sikh Gurdwara of San Diego and were at least 18 years old. The Sikh Gurdwara is comprised of primarily Indian Punjabi speakers. Each participant self-identified as being a native speaker of Punjabi and being able to read Gurmukhi script. Participants had varying levels of English proficiency, but most were fluent L2 speakers of English who had been settled in the USA for many years. However, the majority arrived in the USA during adulthood, and kept up frequent use of Punjabi within the household as well as in the wider diaspora community; thus, they would not fall under the class of heritage speakers of Punjabi. All participants received monetary compensation for their participation.

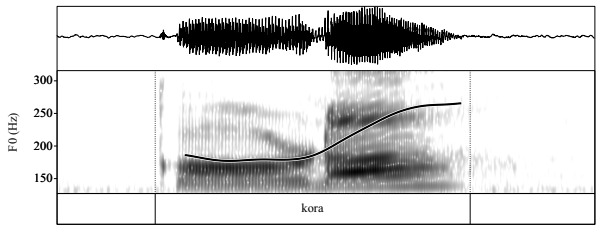
2.2. Stimuli

A two-alternative forced choice (2AFC) task was created using production data of two females from our previous study [9]. Semitone alterations were made to two different DH/T minimal pair words: *bhari* 'full' vs. *pari* 'fairy' [pəri]; *ghora* 'horse' vs. *kora* 'bitter' [koɽa]). According to our previous results [9], the initial consonants in each pair are both realized as voiceless unaspirated; the words in each pair differ only in terms of their F0 (see Figure 1), though there are slight differences in VOT and CPP (Cepstral Peak Prominence, a measure of noise) for each original token, as shown in Table 3.

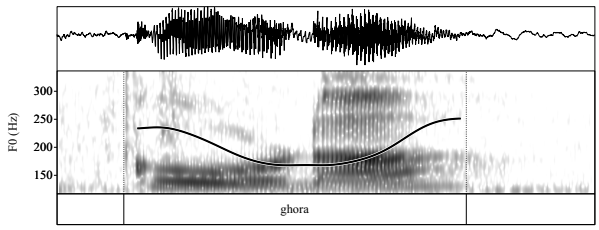
Table 3: Acoustic characteristics of original stimuli. Lower CPP values indicate greater noise.

Word	Speaker	VOT (ms)	CPP
bhari	F1	15.99	19.83
bhari	F2	18.69	19.92
pari	F1	14.27	25.69
pari	F2	13.24	23.28
ghora	F1	38.55	18.79
ghora	F2	39.35	19.25
kora	F1	23.17	17.22
kora	F2	20.35	16.75

First, each word was resynthesized in Praat [11]. In the resynthesis, only the pitch endpoints of the first syllable were kept. For the DH words, which had a falling F0 on the first syllable for each speaker, the beginning pitch point was lowered to match the ending pitch point, resulting in a level pitch/F0. For the T words, the F0 for both speakers was already level during the first syllable. Examples of original tokens are given in Figure 2. For each stimulus, the beginning endpoint was increased by 1 semitone at a time until reaching +6 semitones from the lowest starting point. This resulted in a range from 0-6 semitones per word, for a total of 7 stimuli per word. A semitone of '0' indicated a completely level pitch/F0, while a semitone of '6' indicated the greatest starting pitch value, and hence the greatest F0 falling slope, as illustrated in Figure 3. An equal number of filler stimuli were created matching the previously described procedure, though these minimal pairs did not involve DH consonants. Instead, the filler pairs involved a TH/T aspiration contrast (*pakkhaa*/*pakkaa*) or a D/T voicing (*saj*/*sachch*) contrast. A total of 56 target stimuli were created (2 speakers * 2 word pairs * 2 words * 7 stimuli per word) and 56 fillers were



(a) 'T' token: *kora* 'bitter'



(b) 'DH' token: *ghora* 'horse'

Figure 2: Examples of original tokens. (a) 'T' words have an original level F0 on the first syllable. (b) 'DH' words have an original falling F0 on the first syllable. The rising F0 at the end of each word is attributed to sentential prosody.

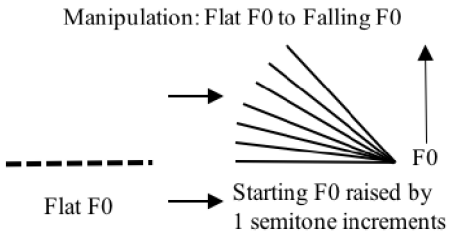


Figure 3: Schematic F0 contours for the manipulations performed on the original DH and T stimuli.

created, for a total of 112 tokens. The intensity of each stimulus was normalized across tokens.

2.3. Procedure

Each participant completed the 2AFC task using Praat with headphones in a quiet room. Each participant was told that this was a Punjabi listening experiment and to click on the word that was heard (out of two choices shown on the screen). No practice rounds were given. Each stimulus could be replayed once, and listeners were able to change their response until they pressed the 'OK' button to continue to the next token. No feedback was given. Presentation of stimuli was randomized as well as the order of the buttons, which displayed one of the two possible word choices orthographically in Gurmukhi script.

2.4. Statistical Analyses

To test for significant effects, a logistic mixed effects regression model was run. The dependent variable was word choice (e.g., *kora* vs. *ghora*; *pari* vs. *bhari*). The fixed effects were semitone (0-6, where 0 indicated no fall and 6 indicates the greatest fall in pitch), original word type ('DH' vs. 'T'), word pair type (*kora*/*ghora* vs. *pari*/*bhari*), gender of participant (M/F), and speaker of stimulus (one of two possible females). Speaker of stimulus was used as a control variable to ensure that participants did not choose tokens based on aspects not related to our hypotheses. A random slope of original speaker by participant was included in the model; this was the maximal random effects structure that converged.

Separate logistic regressions were run with the same fixed and random effect structure on each set of filler words (voicing or aspiration contrast) to ensure that the above result was tied specifically to the 'DH'/'T' F0 distinction, and not just to perceived pitch differences that could extend to non-'DH' words. However, the fixed effect 'word pair type' was not included in these models since there was only one word pair type per model. The 'original word type' variable was different for the fillers: one model contained the voicing difference between *saj* and *sachch* and the other contained the aspiration difference between *pakkaa* and *pakkhaa*.

3. Results

As expected, word choice was significantly affected by semitone ($z = -14.84$, $p < 0.001$) such that a greater semitone value (greater falling pitch) resulted in more 'DH' word choices, while a lower semitone value (more level pitch) resulted in more 'T' word choices. Original word type was also significant ($z = 4.53$, $p < 0.001$), as well as word pair type ($z = -2.23$, $p = 0.026$). These results were not expected and may be attributable to the slight phonetic differences for each original token, as shown in Table 3. Gender of participant and speaker of stimulus were not significant. Results by word pair are shown in Figure 4.

For the fillers, only original word type (voiced vs. voiceless; aspirated vs. unaspirated) was significant in each model (*saj* vs. *sachch* model: $z = -13.05$, $p < 0.001$; *pakkaa* vs. *pakkhaa* model: $z = 13.5$, $p < 0.001$). As predicted, no other fixed effects were found to be significant in either case. As shown in Figure 5, each word was chosen about 50% of the time—regardless of F0—since each filler occurred 50% of the time in the experiment.

4. Discussion

The results of our perception study are in line with the findings of our production experiment. Listeners rely on F0 to distinguish words beginning with DH vs. T, whereby greater pitch falls over the initial vowels are associated with DH words. Our result is robust across changes in speaker of stimulus, as well as in gender or identity of hearer. This suggests that listeners rely on the F0 distinction between 'DH' and 'T' consonants in at least some dialects of Punjabi.

Interestingly, even as this tonal distinction seems to have become cemented in the phonology of some dialects of Punjabi, the environments in which it appears are restricted, and there is evidence that even this limited appearance of tone may be already being lost in certain other populations. As mentioned above, our previous study [9] showed a falling F0 only after

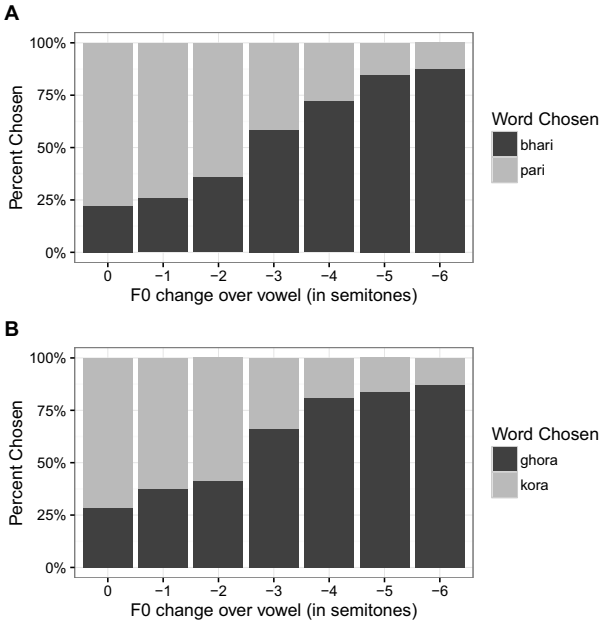


Figure 4: Results by word pair. (A) Word choice between *bhari* ('DH' word) and *pari* ('T' word) by semitone. (B) Word choice between *ghora* ('DH' word) and *kora* ('T' word) by semitone. A higher value of semitone indicates a greater fall in pitch, while a lower value of semitone indicates a more level pitch.

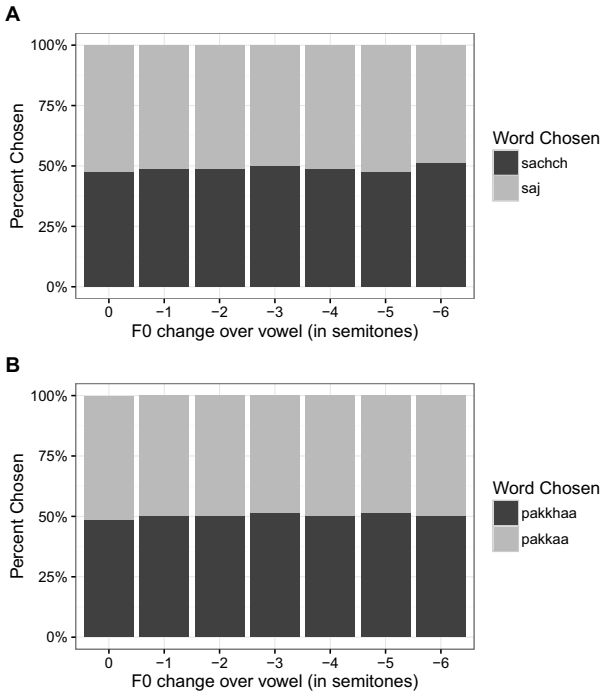


Figure 5: Word choice of filler words that differed in (A) a stop voicing contrast and (B) an aspiration contrast by semitone. A higher value of semitone indicates a greater fall in pitch, while a lower value of semitone indicates a more level pitch.

word-initial 'DH' consonants: word-medial 'DH' consonants induced no distinctive F0 contour in the following vowel. Initial syllables usually bear stress [7], so that might explain why these positions show F0 differences. In so-called "pitch-accented" languages, tonal distinctions are often only found on stressed syllables [12, 13].

In a pilot study we conducted in conjunction with our previous study [9], we found that F0 distinctions were not produced by heritage speakers of Punjabi raised in the US, and that 'DH' consonants were often produced with a flat pitch and retained their voicing and aspiration to distinguish them from 'T' consonants. A study of Punjabi spoken in Lahore [14] shows another interesting pattern: both 'DH' and 'T' words are produced with a falling pitch by these speakers, with only context now being used to differentiate between them. [14] suggests that this loss of tonal distinction may be attributed to language contact with English, Urdu, and Hindi, none of which use tone phonemically. As Ratliff notes, "once born, a tone is not necessarily stable" [2]. As a rare case of an Indo-European language with relatively recently emerging tonal distinctions, perhaps Punjabi is also a ripe target for tone endangerment.

5. Conclusion

In a previous study [9], we showed that 'DH' consonants in Punjabi are produced word-initially as voiceless unaspirated, with a falling tone induced in the following vowel. Here, we follow up this result with a perception study, to determine whether this tone is robustly perceived and used to distinguish words. Our results show that, regardless of speaker or hearer, differences in pitch gradient in the vowel following a 'DH' or 'T' consonant (whose realizations are otherwise phonetically similar) are indeed sufficient for speakers to distinguish between words beginning with these consonants. We therefore suggest that this restricted pattern of tone may be in the process of phonologization in at least some dialects of Punjabi.

6. Acknowledgements

Both authors contributed equally to this work. We are grateful to our friends at Sikh Foundation Gurdwara San Diego for helping us with our study and to everyone at the UC San Diego Linguistics Department for feedback on this work. We are especially grateful to Marc Garellek for his guidance and advice on this project.

7. References

[1] "Ethnologue: Languages of the world, nineteenth edition," Dallas, Texas, 2016, online version: http://www.ethnologue.com.

[2] M. Ratliff, "Tonoexodus, Tonogenesis, and Tone Change," in *The Oxford handbook of historical phonology*, P. Honeybone and J. Salmons, Eds. Oxford: Oxford University Press, 2015.

[3] E. Purcell, G. Villegas, and S. Young, "The acoustics of tonogenesis," *BLS 1*, 1975.

[4] E. Purcell, S. Young, G. Villegas, J. Cromshaw, and R. Smith, "A before and after for tonogenesis," *Phonetica*, vol. 35, pp. 284–293, 1978.

[5] N. I. Tolstaya, *The Punjabi Language*, R. S. McGregor, Ed. Cambridge: Cambridge University Press, 1981.

[6] E. J. A. Henderson, "Tonogenesis: some recent speculations on the development of tone," *Transactions of the philological society*, vol. 80, no. 1, pp. 1–24, 1982. [Online]. Available: http://onlinelibrary.wiley.com/doi/10.1111/j.1467-968X.1982.tb01192.x/abstract

[7] T. K. Bhatia, *Punjabi: A Cognitive-descriptive Grammar*. London: Routledge, 1993.

[8] A. L. Bowden, "Punjabi tonemics and the Gurmukhi script: A preliminary study," 2012.

[9] J. Kanwal and A. Ritchart, "An experimental investigation of tonogenesis in Punjabi," *Proceedings of the 18th International Congress of Phonetic Sciences*, 2015.

[10] J.-M. Hombert, J. J. Ohala, and W. G. Ewan, "Phonetic explanations for the development of tones," *Language*, pp. 37–58, 1979. [Online]. Available: http://www.jstor.org/stable/10.2307/412518

[11] P. Boersma and D. Weenink, "Praat: doing phonetics by computer [computer program] ver. 5.3. 76," 2014.

[12] G. Bruce, "Intonational prominence in varieties of Swedish revisited," in *Prosodic typology*, S.-A. Jun, Ed. Oxford: Oxford University Press, 2005.

[13] S. Godjevac, "Transcribing Serbo-Croatian intonation," in *Prosodic typology*, S.-A. Jun, Ed. Oxford: Oxford University Press, 2005.

[14] M. S. Rafi, "Semantic variations of Punjabi toneme," *Language in India*, vol. 10, no. 8, 2010.