A Description of the Sound System in Pakistani Punjabi

Le Xuan Chan, Yukki Baldoria, Masaki Fukuda, Yuka Omino

International Christian University

1 Introduction

The word for the Indo-Aryan language of Punjabi, also written as Panjabi, comes from the Persian roots *panj* 'five' and *a:b* 'waters.' As the roots suggest, the Punjab region denominates the area of land that is found between the five tributaries of the Indus river, which are presently located between India and Pakistan (Shackle, 2003). Due to this, the different dialects spoken within each territory are categorised as Western or Eastern Punjabi, referring to the dialects spoken in the Pakistani and Indian regions respectively (Hussain et al., 2019). Punjabi can also be written with two different orthographies: Shahmukhi and Gurmukhi, with the Shahmukhi script being the one used in Pakistan and Gurmukhi being the variety used in India. This paper reports on the sound system of a native Modern Standard Punjabi (MSP) variety of Pakistan speaker and compares the data set to Hussain et al. (2019), who investigated on the Punjabi variety of *Lyallpuri* spoken by a native speaker from Faisalabad, Pakistan, and Shackle (2003), who described the modern standard Punjabi (MSP) variety of India.

Shackle (2003) outlined in his chapter on Panjabi 10 vowels, 29 consonants, and 2 semivowels for a total of 41 phonemes. All of the phonemes are written with a one-on-one correspondence between the phoneme and segment, with the exceptions on the five voiceless aspirates sounds, which he represents them as /kh/, /ch/, /th/, /th/, and /ph/. Additionally, the voiceless fricatives /x/ and /f/ and the voiced fricatives / χ / /z/ are only present in the production of Panjabi in speakers who are also highly educated in Urdu. Regarding the 10 vowels, the three centralised vowels /1/ / υ / / ϑ / are seen as the 'short' vowels whereas the other seven peripheral vowels /i/ / ε / / α / / ϑ / / ω / / ϑ / / ϑ / / ω / / ϑ / / ϑ / / ϑ / / ω / / ϑ /

On the other hand, Hussain et al. (2019) also states that there are 10 vowels but argues instead that there are 30 consonants and 2 approximants for a total of 42 phonemes instead of 41, as reported by Shackle (2003). The difference between the two lies on the /v/, which Hussain et al. (2019) argues is only used in loanwords in addition to /f//z//x//y/, which is also what Shackle (2003) reports. Similarly, Hussain et al. (2019) states that these five fricatives are absent in rural speakers of Lyallpuri Punjabi but were present in their consultant's speech. Concerning the vowels, while they both report 10, Hussain et al. (2019) argues that the only low vowel is /a/, whereas Shackle (2003) only reports the /a/. Moreover, while Hussain et al. (2019) reports the /j/ as a palatal, Shackle (2003) argues that it is a velar sound.

Between this paper and the other studies, the main differences lie in the number of vowels found, as we report 12 instead of 10, both of which are peripheral vowels. Intrusion vowels are also reported in our paper. Furthermore, we also found regarding consonants the presence of 32 consonants, which is one more than what Hussain et al. (2019) reports. However, several consonants, which are: $/d^{h}//\eta//t^{h}//t/r$ are reported in this paper that was not mentioned in prior research. Conversely, while they report on the presence of /f//v//t/r/t/r, we were unable to find them in our data. Finally, differences in geminates were also found, as they argue that gemination is limited only if it follows central vowels, but we also found evidence of peripheral vowels, namely: /a//u//i/r preceding geminates.

This paper is structured as follows. First, methods of data collection and analysation will be presented (Section 2). Second, the vowels will be analysed and compared with the two previous studies (Section 3). Third, the consonants will be examined and differences with the two studies will be debated (Section 4). This paper will conclude on a discussion of the results (Section 5).

2 Data Collection

Data was collected remotely except for one session from a male native Western Punjabi speaker. The consultant was born in 1985 and was from the Islamabad district (Shezad town). However, his medium of

instruction in school was Urdu until the 10th grade and English from 10th grade and onwards. As such, while he does speak Punjabi in a casual setting with friends and family, he formally used Urdu or English during the school setting. Nonetheless, he is literate in the Shahmukhi script of Punjabi and uses it with his friends, relatives, neighbours, and colleagues.

A Tascam DR-100 MK-III recorder, set at 44.1 kHz with 16-bit depth, mono, was used to record the elicitation sessions from the consultant. Also, a head-worn SHURE WH30 unidirectional microphone with an XLR connector was used during the elicitations. The distance between the microphone and the mouth of the consultant was at around 10 cm, and words from Bowern (2015)'s Swadesh list in English were first translated into Punjabi by the consultant, then repeated in Punjabi three times. The recordings were then processed and visualised in Praat (Boersma & Weenink, 2020). Figures used throughout this study are also created with this program, and all figures have the spectrogram as well as a textgrid with the following four tiers in descending order: segment, IPA, orthography, and gloss.

3 Vowels

In this study, a total of 11 vowels were found, one more than the 10 phonological vowels found in both Shackle (2003) and Hussain et al. (2019), which is shown below in Table 1.

Table 1: Vowels

	Front	Central	Back
High	i	ΙÜ	u
Mid	e E	Э	0 0
Low	a		a

3.1 *Central Vowels* This study found 3 central vowels, which are the same ones found in previous research: $|\partial/i$ is mid-central unrounded, |i/i is near-high, near-front unrounded, and |v/i is near-high, near-back unrounded.

3.2 *Peripheral Vowels* This study differs from previous research with regards to peripheral vowels. The peripheral vowels found in this study are /i e ε a a u o o/. The vowels that differ from that found in previous studies are /a/ and /a/. In Shackle (2003), /a/ is the only low vowel, whereas in Hussain et al. (2019), /a/ is the only low vowel. In this study, we found both /a/ and /a/. These vowels will be further elaborated on in the discussion section below.

3.3 *Vowel Placement* With regards to word-positions, previous research suggest that the placement of central vowels are restricted in Punjabi. Shackle (2003) states that only peripheral vowels can be in the word-final position. This is supported by Hussain et al. (2019), which shows central vowels only existing in word-initial and word-medial position. This is consistent with what we found in this study, where examples are illustrated in Table 2.

Table 2: Examples of words in Punjabi with the vowel in different positions.

Vowel	Initial		Medial		Final	
	IPA	gloss	IPA	gloss	IPA	gloss

I Ə	ıkk	one	pījo: pətla	father thin		
υ	udəna	fly	turna	walk		
i			pit	back	dhərti	earth
е			khedra	play	othe kətte	there
ε			рега	bad	tətta	where
а	ak	eye	pari	heavy	sз:na	warm
a			paŋ	five	tu:	sleep
u			pura	full	do:	уои
0			kʰoːbṟa	stab	bo	two
э	o:rət	woman	kə.m	who		many

3.4 Long Vowels Hussain et al. (2019) states that peripheral vowels tend to be longer than central vowels. Similarly, it was found in this study that long counterparts exist for all peripheral vowels, but long central vowels were not found. However, no minimal pairs were found between long and short versions of the peripheral vowels, hence it is possible to postulate that vowel length is not contrastive in Punjabi. This is supported by Shackle (2003), which stated that vowel length is less significant than vowel quality in Punjabi.

3.5 *Nasalized Vowels* Previous research states that contrastive nasalization exists for all peripheral vowels in Punjabi, but not for central vowels. In this study, nasal counterparts were found for several peripheral vowels, namely $/\tilde{a} \tilde{i} \tilde{o} \tilde{u}$. However, no minimal pairs between these nasal and oral counterparts were found.

3.6 *Rhotacized Vowels* Previous research does not make any mention of the existence of rhotacized vowels in Punjabi. However, in this study, a rhotacized schwa was elicited in certain words. Figure 1 is an example of this rhotacized schwa, and Table 3 lists words with this vowel found in the present study. This vowel will be elaborated on in the discussion below.



Figure 1: An example of the rhotacized schwa /o/ in [so:fjona]

IPA	gloss
kʰrut͡∫æna	scratch
potəna	dig
ળવેરુમ્	fly (v)
le:tona	lie
dīgəna	fall
ບອຽອາຖຸa	flow
υε:kʰæηa	see

Table 3: Examples of words in Punjabi with the rhotacized schwa /ə/:

4 Consonants

In this study, a total of 32 consonants were found, which is one more than the number of consonants Hussain et al. (2019) found. Table 4 below lists the consonants and Table 5 lists examples of consonants in different word positions.

Table 4: Consonants.

	Bilabial	Labiodental	Alveolar	Retroflex	Palatal	Velar	Glottal
Plosive	p p ^h b		t t ^h d d ^h	td	$\widehat{\mathfrak{t}}\widehat{\mathfrak{f}}\widehat{\mathfrak{t}}^{h}\widehat{\mathfrak{d}}\widehat{\mathfrak{z}}$	k k ^h g	
Nasal	m		n	η		ŋ	
Tap or flap			ſ	t t ^h			
Fricative			s z		ſ	x	h
Approximant		υ		ŀ	j		
Lateral approximant			1	l			

Table 5: Examples of words in Punjabi with the consonants in different positions.

	Initial		Medial		Final	
Conso- nant	IPA	gloss	IPA	gloss	IPA	gloss
р	paŋ	five	t∫uːpra	suck	sap	snake
\mathbf{p}^{h}	pʰuːkղa	blow				
b	bi:	seed	khabe	left		
t	taŋg	narrow	tətta	warm	o:rət	woman

t ^h	t ^h ənda	cold			rə:t ^h	sand
1		C	1 1	man (adult		7
d	duːr	far	bənda	male)	dand	tooth
d^{h}	d ^h ərti	earth				
t			moţa	thick	pit	back
d	thənda	cold	darana	fear	tīd	belly
€Ĵ	t͡ʃaŋga	good	bətt∫a	child	biːt͡∫	seed
$\widehat{t}\widehat{f}^{\rm h}$	t͡ʃʰoṟa	wide	matJhi	fish		
dz	dza:r	root	sud͡ʒəɽŋɐ	swell		
k	ku:la	smooth	nika	small	ıkk	one
k ^h	kʰuŋḍa	dull	akʰã	eye		
g	gə.am	warm	ragna	rub	siːng	horn
m	mandra	short	ləma	long	gə.am	warm
n	nika	short	dzanver	animal	xu:n	blood
η			gɔːŋa	sing		
ŋ			taŋg	narrow	deŋ	three
ſ			pari	heavy	sməndər	sea
t			t͡ʃʰoṟa	wide	d͡ʒa:ṟ	root
$\mathfrak{l}^{\mathrm{h}}$	rʰεηa	live				
s	so:ŗa	narrow	həssəna	laugh	ma:s	meat
z	zəmi:n	earth			tɛ:z	sharp
ſ	∫ika:rmarna	hunt			barı∫	rain
x	xuːn	blood	naxin	fingernail		
h	hara:	green	lahu:	blood	muh	mouth
υ	vədda	big	həva	wind		
ŀ			mə.[ŋa	die	tõ:.t	neck
j			gəlja	rotten		
1	lat	leg	pətla	thin	go:l	round
l			ku:la	smooth	sa:L	year

4.1.1 /p/ No difference was found between the present study and prior research. According to Hussain et al (2019), it occurs at all the positions and our data shows no difference.

4.1.2 $/p^{h/}$ According to Hussain et al. (2019) and Shackle (2003), the unaspirated [p] and the aspirated [p^h] are different phonemes. However, since this study did not find any minimal pairs that have these two sounds, this claim could not be verified. Additionally, Hussain et al. (2019) argues that it occurs at all the positions, but we could not find examples at word-medial and word-final positions.

4.2 /b/ No difference was found between the present study and prior research. According to Hussain et al. (2019), it occurs at all the positions, but we did not find it in word-final positions.

4.3.1 /t/ No difference was found between the present study and prior research. According to Hussain et al. (2019), it occurs at all the positions and our data shows no difference.

4.3.2 $/t^{h/}$ No difference was found between the present study and prior research. According to Hussain et al. (2019) and Shackle (2003), the unaspirated [t] and the aspirated [t^h] are different phonemes. However, since this study did not find any minimal pairs that have these two sounds, we could not confirm this. Hussain et al. (2019) also shows that it occurs at all the positions, but we were unable to find it in word-medial positions.

4.3.3 /t/ According to both prior research, [t] and the retroflexed [t] are different phonemes. However, this study did not find any minimal pairs that have these two sounds. Hussain et al. (2019) states it occurs at all the positions, but we did not find examples at word-initial positions.

4.4.1 /d/ No difference was found between the present study and prior research. According to Hussain, it occurs at all the positions, and our data shows no difference.

4.4.2 $/d^{h}$ According to Hussain et al. (2019) and Shackle (2003), Punjabi does not have the sound [d^h], but our data showed otherwise. We were only able to find it at the word-initial position. Further research is needed to see whether there are restrictions in the places it can occur or not.

4.4.3 /d/ According to Hussain et al. (2019) and Shackle (2003), both /d/ and the retroflex /d/ are different phonemes. Yet, the present study did not find any minimal pairs that have these two sounds. Hussain et al. (2019) states that it occurs at all the positions, but we did not find examples at the word-final position.

The differences among /d d^h d/ are seen in the spectrograms below:



Figure 2: /d/ in the word [du:r], $/d^{h}/$ in the word [d^hərti], and /d/ in the word [darana].

4.5.1 /tf/ No difference was found between the present study and prior research. According to Hussain et al. (2019), it occurs at all the positions, and our data shows that no difference.

4.5.2 /t/h/ No difference was found between the present study and prior research. According to Hussain et al. (2019), it occurs at all the positions, and our data shows that no difference.

4.6 /dz/ No difference was found between the present study and prior research. According to Hussain et al. (2019), it occurs at all the positions, and our data shows that no difference.

4.7.1 /k/ No difference was found between the present study and prior research. According to Hussain et al. (2019), it occurs at all the positions, and our data shows that no difference.

4.7.2 $/k^{h/}$ No difference was found between the present study and prior research. According to Hussain et al. (2019), it occurs at all the positions, which is also supported by our data.

4.8 /g/ No difference was found between the present study and prior research. According to Hussain et al. (2019), it occurs at all the positions, and this is also seen in our data.

4.9 /m/ No difference was found between the present study and prior research. According to Hussain et al. (2019), it occurs at all the word positions, and our data shows no difference.

4.10.1 /n/ No difference was found between the present study and prior research. Hussain et al. (2019) argues that it occurs in all word positions, which is supported in our present study.

4.10.2 $/\eta$ / This study found this consonant, which is not seen in Hussain et al. (2019) nor Shackle (2003). We found it occurring at the word-medial and word-final positions.

4.10.3 $/\eta$ / While Hussain et al. (2019) states that it occurs at word-medial and word-final positions, this study did not find it occuring at the word-final position. Differences between /n η n/ are shown in the spectrograms in Figure 3:



Figure 3: /n/ in the word [d₃anver], /n/ in the word [tang], and /n/ in the word [go:na].

4.11.1 /c/ No difference was found between the present study and prior research. According to Hussain et al. (2019), it occurs at all the positions, and our data shows no difference.

4.11.2 /t No difference was found between the present study and prior research. Hussain et al. (2019) argues that it occurs only at the word-medial and word-final positions, which is supported by our data.

4.11.3 $/t^{h}$ The present study found this consonant, which was not found in previous studies. We found it occurring at the word-initial position.

4.11.5 $/I_{\rm c}$ This study also found this consonant, which was not discussed in the two other prior studies. We found it occurring at the word-medial and word-final positions.

The differences between $/r \downarrow l^h \downarrow$ are shown in the spectrograms below.



Figure 4: /r/ in the word [pari], /r/ in the word [$\hat{t}f^{h}$ ora], /r/ in the word [r^{h} εηa], and /r/ in [mərna].

4.12 /s/ No differences were found between the present study and prior research. According to Hussain et al. (2019), it occurs at all the positions, which is seen supported by our data.

4.13 /z/ While Hussain et al. (2019) states that it occurs at all the word positions, this study did not find it an example of it occurring at the word-medial position.

4.14 /// While Hussain et al. (2019) argues that it occurs at all the positions, this study did not find it occuring at the word-medial position.

4.15 /x/ Hussain et al. (2019) argues that it can occur at all word positions, but this study did not find it occurring at the word-final position.

4.16 /h/ According to Hussain et al. (2019), they state that it only occurs at the word-initial position, but this study found it occurring at all positions.





Figure 5: /h/ in [hara:], [lahu:] and [muh].

4.17 /v/ No difference was found between the present study and prior research. According to Hussain et al. (2019), it occurs at the word-initial and the word-medial positions, which is also reflected in our data.

4.18 /j/ While Hussain et al. (2019) states that it occurred at the word-medial and the word-final positions, we did not find an example of it occurring at the word-final position.

4.19.1 /l/ No difference was found between the present study and prior research. According to Hussain et al. (2019), it occurs at all the positions, and our data shows no difference.

4.19.2 /// No difference was found between the present study and prior research, as Hussain et al. (2019) argue it can only occur at the word-medial and the word-final position, which was also seen in our data.

5 Discussion

5.1 *Vowels* As mentioned in Section 3, both /a/ and /a/ were found in this study. Mean F2 for /a/ was 1195.2 Hz, lower than /a/, which is 1431.91 Hz. This affirms that /a/ is more back than /a/. However, no minimal pairs could be found between the two sounds, so it is possible to suggest that [a] is an allophone of /a/, or vice versa. More data would be required to affirm this.

As mentioned in Section 3.6, prior research makes no reports of any rhotacised vowels in Punjabi. In this present study, a rhotacized schwa was elicited in certain words, as seen in Figure 1. There are two main questions regarding this vowel: 1) Whether the vowel is rhotacized or coarticulated with the following consonant, and 2) Whether it is intrusive. As shown in Table 3, this vowel appears only before the /-na/ suffix for verbs. This presents the possibility that the vowel is coarticulated with the following retroflex /n/, rather than being truly rhotacized. However, additional research is necessary to answer this question.

Regarding intrusion, Punjabi orthography does not reflect the existence of this vowel, but acoustic evidence shows that a vowel is present. This vowel appears where consonant clusters appear in orthography, suggesting that it is an intrusive vowel. Hall (2004) states that "intrusive vowels occur within consonant clusters that include a sonorant" (p.3), which is seen in our study, as the rhotacized schwa appears only before the retroflex /n/. Additionally, Dutch speakers do not recognise intrusive vowels as syllabic (Hall, 2004). Our consultant too,

mentioned that he did not produce any vowels between these consonant clusters, indicating that the produced vowel is also non-syllabic. Additional data is still necessary before any conclusions can be drawn regarding intrusive vowels in Punjabi.

5.2 Consonants As we mentioned in Section 4, we found some consonants which are not mentioned in previous research, namely: $\frac{d^{h}}{\eta} \frac{1}{t^{h}}$. In this section, we will discuss how these consonants are different from the ones mentioned in previous research.

First, we focus on /d^h/, which we mentioned as being a different sound from /d/ or /d/. A comparison of the spectrograms of these consonants are in Figure 2. While all three spectrograms show a negative VOT, characteristic of the /d/ sounds, we also see a closure which immediately follows, indicating the aspiration. This pattern is not seen in the other spectrograms, thus allowing us to conclude that /d^h/ exists as a separate phoneme when compared to /d/ and /d/.

Second, let us explain the characteristics of $/\eta$ /, which is different from /n/ or $/\eta$ /. The spectrograms of these sounds are in Figure 3. Note that F4 of $/\eta$ / is higher than the others, and also F2 is slightly lower. This fact leads us to hypothesize that Punjabi has the nasal $/\eta$ / velar sound.

Finally, we will talk about $/t^{h}$ and $/t^{l}$ sounds, which are not reported by either Hussain et al. (2019) or Shackle (2003). Spectrograms of these rhotics, in addition to $/t^{l}/t^{l}$ are seen in Figure 4. As for the spectrogram of $/t^{h/}$, there is a clear difference between this phoneme and the others in terms of closure. Just like $/d^{h/}$, this point indicates the presence of aspiration in Punjabi. From this, we can conclude the existence of the $/t^{h/}$. As for $/t^{l}$, we can see that the F4 is much higher than the other rhotics. Ergo, it seems that the $/t^{l}$ sound is present and is a different sound.

From all of this, we conclude that the $/dh//\eta//t^{h}//t$ sounds are recognized in the MSP of Pakistan of our consultant. However, we were unable to find minimal pairs, so we are unable to state whether they are phonemes or allophones. Additionally, since we only had one consultant, we are unable to generalise as to whether this phenomenon occurs in the MSP of Pakistan or just in our consultant.

In regard to the phones $/f//v//\chi//t^{h/}$, which are present in Hussain et al. (2019)'s paper, a possible reason why the first three were not found in our study is that they are only found in loanwords, which were not elicited from our consultant. About $/t^{h/}$, one possible reason is that the sample of words we used did not include a word with this phone.

5.2 *Geminates* We also found data regarding geminates from our elicitation sessions. Below is a table of examples comparing singletons and germinates of each phoneme category.

	Singleton		Geminate		
Consonant	IPA	gloss	IPA	gloss	
b	k ^h abε	left	ab:a	father	
t	kətəna	cut	tət:a	warm	
d	goda	knee	vəd:a	big	
k	nika	small	suk:a	dry	
s	asi:	we	həsːəŋa	laugh	

Table 6: A table comparing singletons and geminates.

According to Hussain et al. (2019), a total of 19 consonants have contrastive geminate forms, which include the consonants we found in our study listed above. Below are figures that illustrate a singleton and a geminate:



Figure 6: A singleton /t/ in [kətəna] and a geminate /t:/ in [tət:a].

As can be seen, there is a clear difference between the length of the /t/ in both, as the singleton hovers at around 62 ms while the geminate is at 138 ms. However, while Hussain et al. (2019) argues that geminates can only occur if they follow the central vowels, we found in our case that geminates can also follow peripheral vowels, seen in words such as [ab:a] [suk:a] and [həs:əŋa].

Additionally, we also found that the tense /i/ in [itte] 'here' can also precede a geminate. Minimal pairs of [itte] that differ in the word-initial vowel were also found in the elicited data, namely [ət:e] 'and', [ut:e] 'at', and [it:e] 'here'. These are shown in the figure below:



Figure 7: Geminate /t:/ in the words [ət:e], [ut:e], and [it:e]

The closure duration of the geminate [t] is 158 ms in [ət:e], which is consistent with Hussain et al. (2019) as it is a central vowel. However, the closure durations of the stop preceded by [i] in [u] were similar as well, at 138 ms in [it:e] and 180ms in [ut:e]. This is also comparable to the examples seen in Figure 6, leading us to believe that peripheral vowels can also precede geminates in Punjabi.

7 References

Boersma, P., Weenink, D. (2020). Praat: doing phonetics by computer (Version 6.1.03) [Computer software]. https://www.fon.hum.uva.nl/praat/

Bowern, Claire. (2015). Linguistic fieldwork: A practical guide. Springer.

Shackle, C. (2003). Panjabi. In G. Cardona, & D. Jain (Eds.), The Indo-Aryan Languages (pp. 581-621). Routledge.

Hall, N. (2004). Implications of vowel intrusion for a gestural grammar. Ms. University of Haifa.

Hussain, Q., Proctor, M., Harvey, M., & Demuth, K. (2019). Punjabi (Lyallpuri variety). J. Int. Phon. Assoc.(published online). https://doi.org/10.1017 S.