

Linguistics and Literature Review (LLR)

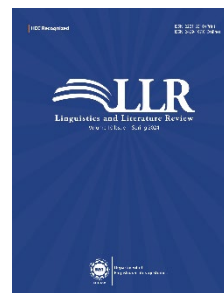
Volume 10 Issue 1, Spring 2024

ISSN(P): 2221-6510 ISSN(E): 2409-109X

Homepage: <http://journals.umt.edu.pk/llr/Home.aspx>



Article QR



Title: Context-Specific Variance in PakE Coronal Stop: An Investigative Study of Pakistani English Speakers

Author (s): Nasir Abbas Syed¹, Shah Bibi¹, and Sehrish Shafi²


Affiliation (s): ¹Lasbela University of Agriculture, Water and Marine Sciences, Uthal, Pakistan.
²Mirpur University of Science and Technology (MUST), Mirpur, Pakistan.

DOI: <https://doi.org/10.32350/llr.101.04>

History: Received: September 14, 2023, Revised: January 22, 2024, Accepted: February 14, 2024, Published: March 31, 2024

Citation: Syed, N. A., Bibi, S., & Shafi, S. (2024). Context-Specific variance in PakE coronal stop: An investigative study of Pakistani English speakers. *Linguistics and Literature Review*, 10(1), 67–84.
<https://doi.org/10.32350/llr.101.04>

Copyright: © The Authors

Licensing:  This article is open access and is distributed under the terms of [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)

Conflict of Interest: Author(s) declared no conflict of interest



A publication of

Department of English and Literary Studies, School of Liberal Arts,
University of Management and Technology, Lahore, Pakistan

Context-Specific Variance in PakE Coronal Stop: An Investigative Study of Pakistani English Speakers

Nasir Abbas Syed¹, Shah Bibi^{1*} and Sehrish Shafi²

¹Faculty of Languages and Literature, Lasbela University of Agriculture, Water and Marine Sciences, Uthal, Pakistan.

²Department of English, Mirpur University of Science and Technology (MUST), Pakistan.

Abstract

The current study shows a new allophonic split in Pakistani English, challenging the widely accepted assumption that Pakistani English speakers produce /t/ as a retroflex [ɖ], claiming instead that they produce alveolar /t/ in [st] cluster. Speech sound utterances of 15 undergrad PakE students were recorded to test this hypothesis. The participants involved in the experiments had spent more than ten years mastering PakE. From the utterances of participants, VOT of coronal stops and F3 of the following vowels were obtained. The findings showed that PakE speakers produced English /t/ with a roughly 8–10 ms longer VOT in the words beginning with ‘st’ cluster. A significant F3 raising was also seen in the vowels after the coronal stops in ‘st’ clusters on word-initial margins. F3 raising is an indicator of the absence of retroflex gesture from the coronal stops. Based on these results, it is argued that a new allophonic split has emerged in the PakE, that is, it is produced as alveolar [t] in word-initial /st/ clusters but a retroflex [ɖ] elsewhere.

Keywords: allophone, F3, plosives, retroflex, VOT

Introduction

It is already established in the literature that L1 has a strong influence on L2/FL (Best & Tyler, 2007; Brown, 1998; Flege, 1995). English, which is spoken worldwide as a lingua franca, has developed into different varieties because of the influence of the L1s of non-native speakers. Each variety has some specific features of its own. Pakistani English is a non-native variety of English which is different from the British or American English (Irfan, 2017). Pakistani English is based on certain factors. These factors play their role in shaping Pakistani English distinct from British and American English in the domains of morphology, syntax, and phonology. For

*Corresponding Author: bibishah52@gmail.com

example, retroflex coronals are a prominent feature of Pakistani English and Indian English because Indo-Aryan languages spoken in Pakistan and India have this feature (Masica, [1993](#)). Therefore, speakers of English in India (Sailaja, [2009](#)) and Pakistan (Rahman, [1991](#)) substitute English alveolar /t/ with a retroflex [ʈ] although it is pronounced alveolar in British and American English.

Pakistani English is an independent variety that enjoys the status of the official language and is used as a medium of instruction in higher education institutions in Pakistan (Ahmed & Ali, [2014](#)). The current study highlights how the L1s and some phonetic articulatory constraints have given birth to a new feature in Pakistani English (PakE). We argue that, contrary to the general perception, established viewpoint and strong claims by previous authors (Mahboob & Ahmar, [2004](#); Rahman, [1990](#), [1991](#), [2020](#)), British English (BrE) voiceless alveolar coronal plosive /t/ is not always pronounced as a retroflex in the PakE. It has its context-specific distribution, for it is pronounced as an alveolar in word-initial clusters in words like ‘steal, stool, star’ and so on. However, in other contexts, it is pronounced as retroflex. Due to this, two different Voice Onset Time (VOT) ranges have been developed for /t/ in the PakE as when it is pronounced as an alveolar unaspirated voiceless stop, it yields slightly longer VOT than in other contexts where it is normally pronounced as a retroflex.

The current study is the outcome of a decade-long series of experiments with several PakE speaker’s groups. In these studies, the voiceless coronal plosive in the speech of PakE speakers from different L1s was acoustically analyzed. The study focuses on the nature of English voiceless coronal stop /t/ in the PakE and its variants in different contexts so, for the analysis it is needed to recapitulate some previous findings related to this topic. PakE has been accepted as a variety of English like other standard varieties (Gargesh, [2019](#); Jenkins, [2009](#); Kachru & Nelson, [2006](#)).

The common characteristics that differentiate between PakE and native BrE speakers are based on the facts that the PakE speakers

- a) do not differentiate between English /v/ and /w/ and assimilate these consonants with a single labio-dental approximant of their L1s,
- b) produce dental fricatives as stops,
- c) do not maintain context-specific allophonic variance between dark and clear lateral,

- d) assimilate post-alveolar fricative /ʒ/ with the approximant /j/,
- e) produce all voiceless plosives as unaspirated without maintaining allophonic variance (Syed & Bibi, [2024](#)), and
- f) pronounce alveolar stops as retroflex (as claimed in the previous research).

In an empirical study, nearly all these features of PakE have been attested. The last two assertions, however, require more explanation. The previous experimentation further illustrates some facts about these aspects of the PakE. Voiced stops in the indigenous languages of Pakistan are pre-voiced therefore, the voiced stops of English are also produced with negative VOT (pre-voicing) by Pakistani speakers (Syed et al., [2023](#)). Along with this, the previous research also shows that although all voiceless stops are produced unaspirated with short-lag VOT there is a variation between the voiceless coronal stop produced on word-initial position, for instance in words like ‘teach, tool, tart’, and in ‘st’ clusters with words like ‘steal, stool, start’ and so on. The difference between these two contexts is that the Pakistani English ‘t’ in the former is produced with slightly smaller VOT and lowered F3 which are acoustic correlates of a retroflex consonant but in the latter context, it is produced with relatively longer VOT and raised F3 which are the indicators of loss of retroflex gesture.

Based on these facts, it is hypothesized that there is an allophonic split in the PakE between the alveolar and retroflex coronal stops in that the coronal /t/ which is alveolar in the BrE. It is also produced as alveolar in words like ‘stop, steal, star’ and, in the context where it follows /s/ on word-initial position although it is produced as retroflex in the PakE (as claimed by the previous researchers) in all other contexts. The hypothesis tested in this study may be contextualized in the following theoretical background.

Theoretical Background

The languages spoken in different regions of the world have different laryngeal settings for stops. These settings are defined through phonological features like voicing or aspiration. Both are acoustically defined using VOT along with some other paradigms, as standard correlates for the classification of plosives (Abramson & Whalen, [2017](#)). VOT is a time interval between the burst of a stop and the onset of a vowel (Docherty, [1992](#)). VOT is an acoustic correlate of aspiration for plosives (Foulkes et al., [2010](#)). Based on VOT, voiceless plosives are divided into two

categories, namely unaspirated and aspirated stops. If the post-burst voicing for the vowel happens immediately after a plosive takes a long time to start, it is considered as an aspirated stop. On the other hand, if the post-burst voicing for the immediately following vowel starts soon after the burst, it is considered as an unaspirated stop. In this way, plosives are classed into aspirated and unaspirated stops based on the aspiration.

In the previous literature, it is already established that there is a significant relationship between the place of articulation and the VOT of stops. For example, it is claimed that a bigger contact zone between active and passive articulators yields relatively longer VOT and the distance between the vocal folds of a speaker and the place of articulation of stops are in inverse correlation. Therefore, it conveys that a bigger distance between vocal folds and place of articulation causes a decrease in VOT and vice versa (Lisker & Abramson, 1967; Stevens et al., 1986). It is also known that a retroflex has a smaller VOT than a coronal produced without retroflex gesture (Ladefoged & Maddieson, 1996). This fact is already established in the study of several world languages including Marathi, Hindi (Lisker & Abramson, 1967), and Tiwi (Anderson and Maddieson, 1994) and Malyalam (Dart, 1991). Significant lowering of the third formant (F3) occurs in the vowels that immediately follow retroflex consonants and on the other hand, significant raising occurs in F3 of the vowels which follow alveolar stops (Steriade, 2001).

This theoretical context provides grounds to test the hypothesis raised in this study. As pointed out in the previous section, the previous researchers have developed a consensus on the claim that the BrE coronal /t/ is produced as a retroflex in the PakE in all the contexts. On the other hand, the study claims that the BrE coronal /r/ is produced as an alveolar in PakE by Pakistani speakers of English when it occurs in words starting with the 'st' clusters although it is produced as retroflex in all other contexts (as other researchers have already claimed). This study further claims on the basis of empirical evidence that voiceless alveolar stop has become a permanent feature of PakE thus, giving birth to a new allophonic split in the PakE which is not yet identified in the literature by other researchers.

Objectives of the Study

The goal of the current investigation is to verify the above hypothesis. For this purpose, the following objectives are addressed in the current study.

- a) To study the nature of BrE coronal stop /t/ in PakE.
- b) To confirm whether /t/ is produced as alveolar or retroflex in the words starting with ‘st’ clusters in PakE.

The researchers on PakE have not studied PakE stops in different contexts. Therefore, the results of the current study will contribute to the literature significantly. It will investigate a specific phonological characteristic of the PakE that has never been examined before. Additionally, it will show how various factors, such as non-native speakers’ L1 and phonetic constraints influence the evolution of a language.

Literature Review

The pronunciation of various linguistic groups of PakE speakers was repeatedly studied in the past. In the first study, 30 male students of MA English who were native speakers of Central Saraiki, were asked to produce a list of words, which along with other words, also included the words ‘teeth, steal’ (Syed, [2013a](#)). Each token was produced six times, three times as an exclusive word and three times in a carrier sentence. The carrier sentence was “I say _____ again”. VOTs of /t/ in these utterances were elicited for the comparison. The VOTs in the tokens produced as exclusive words were not significantly different from those produced in the sentences. Thus, VOT values obtained in all of the six repetitions were averaged. The results showed that the participants produced English voiceless coronal stops with an average of 14.73 ms VOT in the word ‘teeth’ but that in the word ‘steal’ was produced with an average of 20.50 ms VOT. A difference of approximately six milliseconds of VOT between the coronal plosives in two different contexts (initial position with cluster and without cluster) was statistically significant ($p < .05$).

The same experiment was repeated with a group of ten male English language teachers in Pakistan. These teachers were also native speakers of Central Saraiki. The data obtained from the isolated words and carrier sentences was not significantly different therefore, the repetitions were averaged. VOTs of voiceless coronal stops obtained from the utterances of the teachers show that they had produced the coronal stop in the word ‘teeth’ with an average of 19.34 ms VOT and that in the word ‘steal’ with an average of 26.63 ms VOT. Thus, there was a significant ($p < .05$) difference of approximately 7 ms VOT between the two variants of the voiceless coronal plosives produced in two different contexts (Syed, [2015](#)).

Syed and Saleem (2019) arranged another experiment with two groups of L1 Saraiki speakers of English. The study involved 29 Pakistani PakE learners who migrated to England and lived there for five years and 30 educated L1 Saraiki speakers of PakE living in Pakistan. VOTs of /t/ in the words 'teach' and 'steal' were obtained in the recordings from these participants in two experiments. The method of data collection was the same as that used in the previous studies. The first experiment was conducted in England and the second in Pakistan. The results showed that Pakistan-based PakE speakers produced English voiceless coronal stops with 19.31 ms and 25.02 ms VOT on average in the words 'teach' and 'steal', respectively. The same sounds were pronounced by the UK-based participants with 25.60 ms and 30.07 ms VOT on average in these words 'teach' and 'steal', respectively. A difference of 5-6 ms VOT between the target plosives in the two different contexts was found statistically significant ($p < .05$) in the pronunciation of both groups of participants.

Syed and Bibi (2024) carried out the same experiment with two groups of participants using six words as stimuli, namely 'teach-steal, tart-start, tool-stool'. The stimuli were produced in a word reading task thrice in sentences and three times as exclusive words. In this study, one group of participants consisted of ten female native speakers of L1-Sindhi and the other group consisted of 20 male speakers of L1-Eastern Balochi. They all had learnt PakE for more than ten years on average at the time of the experiment. A significant ($p < .05$) difference of 3-5 ms VOT on average was found between English voiceless coronal stops produced in the words starting with st clusters and those starting with the word-initial coronal stops. The mean VOTs of Eastern Balochi speakers were 20.00 ms and 23.50 ms on average in the word-initial coronal plosives and those in /st/ clusters, respectively. The mean VOT of the two variants in the speech of Sindhi speakers were 17.00 ms and 22.50 ms in these contexts and the difference between means was statistically significant ($p < .05$) in both the groups.

In the above discussed experiments, it was arguably clear that PakE speakers who speak Saraiki, Sindhi, and Eastern Balochi as their mother tongues produce voiceless coronal stop 't' with 3-7 ms longer VOT when it occurs in the word-initial 'st' clusters. It is already established that the PakE speakers produce English voiceless coronal plosive as a retroflex /t/ (Rahman, 1990, 1991, 2020) whereas the same is pronounced as an alveolar

/t/ by the native speakers of British English (BrE). An alveolar stop is produced with slightly longer VOT than a retroflex because a retroflex gesture minimizes the VOT of plosives (Ladefoged & Maddieson, 1996; Steriade, 2001). F3 lowering is another acoustic correlate of the retroflex gesture (ibid). Considering these, the utterances of 59 participants from the two groups of speakers in the study by Syed and Saleem (2019) were re-analyzed. In the re-analysis, it was found that all participants had produced English voiceless coronal with third formant (F3), significantly ($p < .05$) raised when produced in words like ‘steal’ compared with that in the word ‘teach’. All these experiments motivated the researchers to hypothesize that a new allophonic split has emerged in the PakE voiceless coronal plosive, in that, it is pronounced as alveolar in the word-initial ‘st’ clusters but a retroflex in other contexts. The current study aims to test this hypothesis by acoustically analyzing the coronal stops of English produced by Pakistani speakers who speak Urdu as their L1.

Before moving on to the details of the research methods used for data collection in the current study, it is relevant to reproduce the results of another study that is relevant to the current topic but has yielded different results. Syed (2013b) also studied VOTs of plosives by L1-Pashto speakers of PakE. For this purpose, 12 L1-Pashto speakers of PakE were recorded using the same methods of data collection as in the previously discussed study. The participants were living in the UK for 2-3 years to acquire PhD degrees at the time of the experiment. They had learnt PakE in Pakistan for more than sixteen years. English is taught as a compulsory subject at every stage of education in the province where these participants had lived in Pakistan (namely KPK). In this way, each of these participants had learnt PakE for a minimum period of sixteen years because all the participants, before moving to England for higher studies, had obtained a master’s degree from Pakistan. These participants produced English voiceless coronal stops on the word-initial margin and in ‘st’ clusters with a VOT of 35.64 ms and 33.17 ms on average respectively. There was no significant difference between these mean values. However, these results are different from those obtained from the participants of L1-Saraiki, Sindhi, and Balochi speaker participants, in that, the Pashto speaker participants do not pronounce /t/ in two different contexts with significantly different VOTs. A comment on this study is added in the analysis section below.

Method

In the current study, data was collected in two phases and the participants were divided into two groups, namely the monolingual group and the L2 speakers group. They were all native speakers of Urdu. In the first phase, Urdu words were recorded to confirm the VOT of Urdu plosives. Subsequently, in the second phase English words were recorded.

Monolingual L1 Speakers

Fifteen monolingual female native speakers of L1-Urdu living in the suburbs of Karachi were selected as participants. Their ages were between 17 and 26 years (mean: 22.73, standard deviation: 2.40). The selection of these participants was based on the convenient sampling. They were asked to produce some Urdu words which were written on a piece of paper. Each word was written in the list three times along with some distractors. For analysis, only target words were selected. Later, the recordings of the target words were acoustically analyzed. The list of stimuli contained the following target words: [ʈaɪ ʈʰaɪ, ʈaɪ ʈʰaɪʰ, paɪ, pʰaɪ, kaɪ kʰaɪ]. The list carries words starting with eight voiceless (aspirated and unaspirated) plosives of Urdu. Urdu has aspiration contrast at phonemic level at four places of articulation namely labial [p pʰ, b bʰ], dental [ʈ ʈʰ, ɖ ɖʰ], post-alveolar (retroflex) [ʈ ʈʰ, ɖ ɖʰ], and velar [k kʰ, g gʰ] (these obstruents also occur in Urdu on word-medial and word-final position). The stimuli that were recorded consisted of words with voiceless aspirated and unaspirated stops only on the word-initial position because the current study focuses on voiceless stops on the word-initial position.

The participants (who were monolingual speakers of Urdu) were asked to read the list. The utterances of the participants were recorded at the place of their choice. Before recording, the researchers confirmed that the native speakers were familiar with the listed words. However, the distractors among the list of stimuli were not acoustically analyzed.

Utterances were recorded only as exclusive words but not in the running speech because all the studies summarized in section 2 confirmed that there is no significant difference between the VOTs of plosives produced in sentences and those produced as exclusive words.

PakE Speaker Participants

Fifteen native female speakers of Urdu were selected as participants in this group. These participants were exclusively students. 5 of them were students at a college (3rd year and 4th year) and 10 were university students (registered in 3rd, 2nd and 1st semester of an undergraduate programme). Their ages ranged between 18 and 28 years ($M = 22.27$, $SD = 2.69$). The selection of participants was based on purposive, convenience, and availability sampling. All of them had studied PakE for a minimum of twelve years in academic institutions in Pakistan where mostly English is taught by non-native Pakistani teachers who themselves speak the PakE (Syed, 2015). The participants were asked to produce six English words (three with aspirated stops and three with [st] cluster) on a computer screen. The English target words were ‘teach, steal, tart, star, tool, stool’.

The whole procedure of recording was explained to the participants but they were not informed about the specific purpose of the study. None of the participants reported any speaking or hearing loss or difficulty. All of the participants participated in the study voluntarily.

Instrument

For data collection, M-Audio digital recorder was used and before recording all the participants were informed about the procedure. The data was collected in two recording sessions with a gap of a month in between. In the first session, the monolingual Urdu speaker participants were recorded. The list of stimuli is crafted in a special attentive manner that the first and last words are not the target words because it is expected that in such recording sessions speakers are slightly nervous in the beginning and at the end thus, produce words with abnormal/unexpected haste, ultimately changing their intonation. Therefore, in order to avoid such anomalies, the first and last words of the list were not the target words.

The second session was arranged one month later. During this session, the PakE speaker participants read a list of words of English carrying six target words and some distractors each listed three times. The first and last words in the list were not the target words and the order of stimuli was randomized for each participant. All participants of this group were also the native speakers of Urdu.

Acoustic Analysis

For the analysis of data, VOT of target stops and F3 of the following vowel were required. For this purpose, the recordings were shifted to a laptop computer. The list of Urdu words had eight coronal stops, while the English stimuli had six English words both consisting of aspirated and unaspirated allophones of coronal stops with each target stimulus having three repetitions. Using Praat software (Boersma & Weenink, 2019), VOT of stops of English and Urdu words were taken. The selection of VOT on the spectrogram was done manually. The area from the burst of the stop to the onset of the first complete cycle of vocal fold vibration for the next vowel was selected for VOT analysis. Later, a computer software ProsodyPro (Xu, 2013) was used to get VOT values of the selections computationally. The output files yielded by the computer software were imported to SPSS for quantitative data analysis. To get F3 of the adjacent vowel the same procedure was followed, except for the selected area on the spectrogram, for this elicitation spanned from onset of the vowel to the middle of the formant. The reason for selecting the first half of the formant length was to determine the impact of the preceding stop on the following vowel which is expected to be more visible on the initial part of the vowel. The results are presented in the next section.

Data Analysis

In this section, VOTs of Urdu plosives by monolingual speakers of Urdu, F3 of the following vowels, and VOTs of L2-English plosives produced by the learners of English are presented.

VOT of L1 Plosives

The participants of the monolingual group produced words of Urdu transcribed in column 2 of Table 1. Mean VOTs are given in column 4 with standard deviations in parentheses.

Table 1

VOT of Urdu Plosives

Stimuli	Phonetic Transcription	Target Sound	Mean (Std. Deviation)
taal	$[\text{t}]_{\text{al}}$	$/\text{t}/$	10.98 (04.26)
thaal	$[\text{t}^{\text{h}}]_{\text{al}}$	$/\text{t}^{\text{h}}/$	62.37 (12.30)
taat	$[\text{t}^{\text{h}}]_{\text{at}}$	$/\text{t}^{\text{h}}/$	08.76 (02.90)

Stimuli	Phonetic Transcription	Target Sound	Mean (Std. Deviation)
thath	t ^h aʈ ^h	/t ^h /	71.08 (21.84)
paal	pal	/p/	11.57 (05.93)
phar	p ^h aʈ	/p ^h /	55.17 (15.11)
kaal	kal	/k/	26.55 (09.64)
khaal	k ^h al	/k ^h /	84.45 (16.12)

A repeated measure ANOVA applied on the VOT values confirms that the aspiration contrast ($F=512.79, p = .0001$), place of articulation effect ($F = 19.22, p = .0001$), and the interaction of place of articulation and aspiration are all significant ($F = 4.23, p = .011$) in these data values. The main concern of this study is to know the VOT of coronal plosives of Urdu only. However, VOTs of labial and velar plosives were obtained to confirm the actual VOT ranges of stops in Urdu because, though it is already known that Urdu is a language that has aspiration contrast at the phonemic level (Schmidt, 2007), the actual ranges of the VOTs of Urdu stops are not recorded in the literature before.

VOT of PakE Coronal Plosives

VOT of coronal stops and F3 values of the vowels adjacent to the stops produced by 15 PakE speakers (second group of participants) were taken using Praat in Prosodypro software. Average values of VOT and F3 are given in the table below. Standard deviations are in parentheses.

Table 2
F3 and VOT for English Plosives

Variables Stimuli	F3	VOT
	Mean (Standard Dev.)	Mean (Standard Dev.)
teach	3156 (217.87)	12.69 (2.67)
steal	3257 (192.15)	22.66 (9.73)
tart	2744 (274.76)	10.73 (4.26)
star	2907 (255.70)	19.91 (6.49)
tool	2832 (111.58)	10.82 (3.48)
stool	2930 (102.67)	23.53 (9.97)

A repeated measure ANOVA with context (coronal stop on word-initial position and in 'st' cluster) and vowels as dependent variables was run on the VOT data. The results confirmed that effect of the context is significant ($F = 28.88, p = .0001$), but that of vowel ($F = 3.09, p = .061$) and interaction

between aspiration and vowel are non-significant ($F=1.48$, $p=.24$). A similar test was also applied on F3 data and the results confirm that the context of occurrence ($F=11.45$, $p=0.004$) and effect of the vowel are significant ($F=23.40$, $p=.000$) but the interaction between them is non-significant ($F=1.18$, $p=.32$). Figure 1 reflects the formant raising in the vowels after the coronal stops of different vocalic contexts.

Figure 1
Formant Values

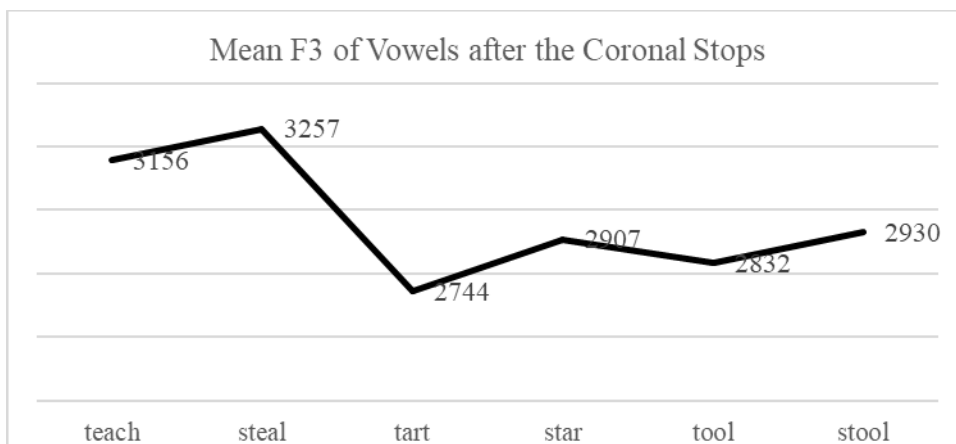


Figure 1 indicates that the values of the third formant are significantly lowered when it follows a coronal which is preceded by /s/ in an /st/ cluster, while it is relatively raised in the vowels which follow a word-initial coronal stop.

These results confirm that there is not only a significant increase of 9-12 ms VOT in voiceless coronal stops produced by the participants in 'st' clusters than in the word-initial position, a 100-200 Hz lowering is also noticed in F3 of the vowels adjacent to the stop on the word-initial margin as compared to the ones adjacent to the coronal stop in 'st' cluster. Since F3 lowering and minimized VOT duration are the indications of retroflex gestures in stops. Hence, the results confirm that these PakE speakers produce word-initial voiceless stops without retroflex gestures in 'st' clusters.

In a further analysis, VOT ranges of L1 and L2 stops were found to be significantly different from each other. In this analysis, variants of Pak-English coronal plosive were compared with those of retroflex ($F=65.68$,

$p=.000$) and dental ($F=111.91, p=.000$) stops of Urdu which confirmed that both L1 and L2 VOT ranges were significantly different from each other. These comparisons confirmed that the learners were not transferring L1 VOT stops rather they had developed different phonetic categories for PakE stops. These issues are discussed in detail in the following section.

Discussion

The results show that although Urdu-L1 speakers of PakE produce both variants of English coronal stops without aspiration, they produce it in ‘st’ clusters with 9-12 ms longer VOT. They also lower their F3 while producing it on word-initial position but such lowering is not observed when it occurs in ‘st’ clusters word-initially. Both these acoustic correlates (smaller VOT and F3 lowering) confirm the claims of the previous researchers that coronal plosive is produced as retroflex in PakE on word-initial margins. Contrary to this, when the coronal stop ‘t’ occurs in words starting with ‘st’ clusters it becomes articulatorily different. It loses its retroflexion which is reflected in the increase in VOT and raising of F3. This is due to the fact that it becomes alveolar in this context. The consistent results of the experimentation summarized in the section 2 as well as the findings of this study confirmed that an allophonic split has occurred in PakE coronal plosives for they are produced as alveolar in word-initial clusters but retroflex in all other contexts. The previous researchers could not identify this feature of the PakE as they all unanimously claimed that PakE coronal plosives are produced retroflex across the board.

The emergence of this split in the PakE is triggered because of some articulatory constraints. As is already known, the PakE typically produces coronal stops as retroflexes, but because /s/ is alveolar and requires a forward movement of the tongue tip in its articulation whereas the PakE retroflex /t/ demands a backward movement of the tongue tip simultaneously thus, it becomes articulatorily difficult to produce a retroflex immediately after /s/. A retroflex /t/ is [-anterior] in feature geometry, whereas an alveolar /s/ is [+anterior] (Clements & Hume, 1995). The co-occurrence of two consonants (/s/ and /t/), one with [+anterior] and the other with [-anterior] feature simultaneously demands two opposite gestures from the speakers which are difficult to perform. Resultantly, the speakers after producing an anterior /s/ cannot produce a posterior retroflex. They would rather assimilate the place of active articulator of the two phonemes and

produce both as alveolar in this context. That is why, the VOT of the stop immediately following /s/ increases and F3 of the adjacent vowels also rises.

Lastly, there is an issue related to the results of the study (Syed, [2013b](#)) in which L1-Pashto speakers did not face this difficulty in producing /st/ since no significant F3 raising or VOT increase was seen in their coronal stops in this context. The reason for this is that, Pashto is a language that has a complex system of stop-stop and fricative-stop clusters with a difference of zero sonority between the members of the clusters (Khan & Bukhari, [2014](#)). In this way, native speakers of Pashto are already familiar with and quite competent to produce clusters of consonants with more complex combinations of gestures. Therefore, they do not face as much difficulty in the production of PakE /st/ clusters as speakers of other L1s in Pakistan do. Other languages spoken in Pakistan, namely Saraiki, Balochi, Sindhi, and Urdu do not allow any fricative-stop clusters. Thus, adult speakers of these languages, unlike Pashto speakers, are not competent in producing such difficult clusters and they repair this difficulty with progressive assimilation of the anterior place of /s/ with the following /t/ in this context. The result is the emergence of an allophonic split in English coronal stops in these two contexts.

Conclusion

Using recordings of the L1 Urdu speakers of PakE, this study empirically demonstrated that PakE speakers pronounce voiceless coronal stops in word-initial “st” clusters as alveolar and not as retroflex, as is asserted in PakE literature. However, it is produced as retroflex in other contexts. The previous researches by the authors during the past ten years, has examined Saraiki, Sindhi, and Balochi speakers’ PakE speech through several experiments. Finally, in this experiment, fifteen PakE speakers whose first language was Urdu were recorded for computational analysis. Acoustic examinations of their coronal stops showed exactly when they were created after /s/ on word-initial margin, with VOT of the stop, and F3 of the following vowel raised by 9–12 ms and 100–200 Hz, respectively. Based on these results, it is claimed that a new allophonic split has emerged in the PakE, in that, the voiceless coronal plosive which is produced as alveolar /t/ in native English, is, contrary to the claims of the indigenous researchers, also produced as alveolar when it occurs after word-initial /s/. However, in other contexts, it is produced as retroflex. A major difference in the articulation of coronal stop produced by L1-Pashto speakers of PakE

and other PakE speakers also implies the probability of other grammatical differences between various groups of PakE speakers in Pakistan. To comprehend the real characteristics of various kinds of PakE, the future researchers need to study these differences. Therefore, it is hypothesized that PakE is not a single variety instead, it is a combination of many dialects of English that are spoken by L1 speakers. This research opens the door for researchers working on PakE grammar to test this hypothesis.

Conflict of Interest

The author of the manuscript has no financial or non-financial conflict of interest in the subject matter or materials discussed in this manuscript.

Data Availability Statement

The data associated with this study will be provided by the corresponding author upon request.

References

- Ahmed, S. R., & Ali, S. (2014). Impact of urduised English on Pakistani fiction. *Journal of Research*, 50, 61–75.
- Abramson, A. S., & Whalen, D. H. (2017). Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. *Journal of Phonetics*, 63, 75–86. <https://doi.org/10.1016/j.wocn.2017.05.002>
- Anderson, V., & Maddieson, I. (1994). *Acoustic characteristics of Tiwi coronal stops* (Working Paper No. 87). University of California, Los Angeles. <https://escholarship.org/content/qt0942x2jv/qt0942x2jv.pdf>
- Best, C. T., & Tyler, M. D. (2007). Nonnative and second-language speech perception: Commonalities and complementarities. In O. -S. Bohn & M. J. Munro (Eds.), *Language experience in second language speech learning: In honor of James Emile Flege* (pp. 13–34). J. Benjamins. <https://doi.org/10.1075/llt.17.07bes>
- Boersma, P., & Weenink, D. (2019). *Praat: Doing phonetics by computer*. University of Amsterdam. <https://www.fon.hum.uva.nl/praat/>
- Brown, C. A. (1998). The role of the L1 grammar in the L2 acquisition of segmental structure. *Second Language Research*, 14, 136–193. <https://doi.org/10.1191/026765898669508401>

- Clements, G. N., & Hume, E. V. (1995). The internal organization of speech sounds. In J. Goldsmith (Eds.), *A handbook of phonological theory* (pp. 245–306). Blackwell.
- Dart, S. N. (1991). *Articulatory and acoustic properties of apical and laminal articulations* (Publication No. 9122664) [Doctoral dissertation, University of California]. ProQuest Dissertation & Theses.
- Docherty, G. J. (1992). *The timing of voicing in British English obstruents*. Foris Publications.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange, W. (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). York Press.
- Foulkes, P., Docherty, G. & Jones, M. J. (2010). Analyzing Stop. In M. Di Paolo & M. Yaeger-Dror (Eds.), *Sociophonetics: A student's guide*. Routledge.
- Gargesh, R. (2019). South Asian Englishes. In C. L. Nelson, Z. G. Proshina & D. R. Davis (Eds.), *The handbook of world Englishes* (pp. 105–134). John Wiley and Sons Inc.
- Irfan, H. (2017). The postgraduate students and their teachers' perceptions of the use of Pakistani English (PakE) in Pakistani Universities. *Journal of Research and Reflections in Education*, 11(1), 38–48.
- Jenkins, J. (2009). *World Englishes: A resource book for students*. Routledge.
- Kachru, Y., & Nelson, C. L. (2006). *World Englishes in Asian contexts*. University Press.
- Khan, M. K., & Bukhari, N. H. (2014). Reversed sonority clusters in Pashto: An optimality theoretic justification. *Kashmir Journal of Language Research*, 17(1), 33–50.
- Ladefoged, P., & Maddieson, I. (1996). *The sounds of the worlds languages*. Blackwell.
- Lisker, L., & Abramson, A. (1967). Some effects of context on voice onset time in English stops. *Language and Speech*, 10(1), 1–28.

- Mahboob, A., & Ahmar, N. H. (2004). Pakistani English: Phonology. In E. W. Schneider (Ed.), *A handbook of varieties of English: A multimedia reference tool* (pp. 1003–1016). Mouton de Gruyter.
- Masica (1993). *The Indo-Aryan languages*. Cambridge University Press.
- Rahman, T. (1990). *Pakistani English: The linguistic description of a non-native variety of English* (Vol. 3). National Institute of Pakistan Studies, Quaid-i-Azam University.
- Rahman, T. (1991). Pakistani English: some phonological and phonetic features. *World Englishes*, 10(1), 83–95.
- Rahman, T. (2020). Pakistani English. In K. Bolton, W. Botha. & A. Kirkpatrick (Eds.) *The handbook of Asian Englishes*. Wiley Blackwell.
- Sailaja, P. (2009). *Dialects of English: Indian English*. Edinburgh University Press.
- Schmidt, R. L. (2007). Urdu. In G. Cardona & D. Jain (Eds.), *The Indo-Aryan languages* (pp. 315–385). Routledge.
- Steriade, D. (2001). Directional asymmetries in place assimilation. In E. Hume & K. Johnson (Eds.), *The role of speech perception in phonology* (pp. 219–50). Academic Press.
- Stevens, K. N., Keyser, S. J., & Kawasaki, H. (2014). Toward a phonetic and phonological theory of redundant features. In J. S. Perkell & D. H. Klatt (Eds.), *Invariance and variability in speech processes* (pp. 426–463). Lawrence Erlbaum Associates.
- Syed, N. A. R. (2013a). *The acquisition of English consonants by Pakistani learners* [Doctoral dissertation, University of Essex]. University of Essex Library. <https://tinyurl.com/23ebjhud>
- Syed, N. A. (2013b). Voice onset time (VOT) for voiceless plosives in Pashto (L1) and English (L2). *The Journal of Humanities & Social Sciences*, 21(3), 79–94.
- Syed, N. A. (2015). The role of teacher in second language learning. *NUML Journal of Linguistic Inquiry*, 13(1), 71–91.
- Syed, N. A., Aldaihani, S. M., & Bibi, S. (2023). Acquisition of voice onset time for voiced plosives of English by adult learners of Balochistan.

International Journal of English Linguistics, 13(2), 21–28.
<https://doi.org/10.5539/ijel.v13n2p21>

Syed, N. A., & Bibi, S. (2024). VOT for plosives in indigenous languages of Balochistan: Implications for adult learners of English. *Journal of Second Language Studies*, 7(1), 157–192.
<https://doi.org/10.1075/jsls.23002.sye>

Syed, N. A., & Bibi, S. (2022). The emergence of a new phonological feature in Pakistan English: Focusing on an allophonic split in VOT of coronal stops in PakE. *English Today*, 38(4), 278–283.
<https://doi.org/10.1017/S0266078421000316>

Syed, N. A., & Saleem, A. (2019). A review of the speech learning model in the perspectives of learners of English in Pakistan. *Journal of Independent Studies and Research*, 17(2), 85–110.
<https://doi.org/10.31384/jirmsse/2019.17.2.12>

Xu, Y. (2013). *ProsodyPro—A tool for large-scale systematic prosody analysis*. UCL Discovery.
<https://discovery.ucl.ac.uk/id/eprint/1406070/>