

ARTICLE

Sindh Journal of Linguistics (2025)
Volume 5, Issue 1-pp. 1-7
eISSN: 2957-9309
Published Annually By:
© Sindh Madressatul Islam University Karachi
License Type: CC-BY-SA

Acquisition of Allophonic Aspiration Contrast of English Voiceless Stops

Sami Ullah Khan¹
Ghulam Ali Buriro²
Dolat Khan³

Abstract

This study examines the acquisition of the allophonic aspiration contrast of English voiceless stops by Khwar speakers, whose native language encodes aspiration as a phonemic distinction. Drawing on the Speech Learning Model (SLM), the study investigates how first-language (L1) phonology and linguistic environment influence the acquisition of second-language (L2) aspiration patterns. A production experiment measured the Voice Onset Time (VOT) of English voiceless stops produced by thirty native Khwar speakers, divided into Pakistan-based and UK-based groups. The results show that Pakistan-based speakers tended to equate English aspirated stops with their unaspirated L1 counterparts, indicating strong L1 transfer effects. In contrast, UK-based speakers demonstrated greater overall accuracy in producing English aspiration, though their productions also revealed instances of hypercorrection, suggesting heightened sensitivity to the L2 contrast. These findings reveal the critical role of L2 input in phonetic category formation and support the view that sustained exposure and explicit phonetic instruction may facilitate the restructuring of L1-based phonological representations in L2 acquisition.

Keyword: Second Language Acquisition, VOT, First Language (L1), Second Language (L2)

Introduction

This study carries the process of acquisition of aspiration contrast of English voiceless stops by Khwar adult learners of English. English stops are studied with a focus on aspirated and un-aspirated sounds. Consequently, target stimuli contain 'teacher, store, kitchen, skin' produced by Khwar L1 speakers. Aspirated sound is produced when coronal and velar sounds of English produced at the initial position as mentioned stimuli 'teacher' and 'kitchen' whereas un-aspirated sound is produced when there is a cluster of consonants like 'store' and 'skin'. The level of aspirated

¹ Assistant Professor , Faculty of English language and Literature, Lasbela University, Balochistan

² Professor, Institute of English Language, and Literature, University of Sindh, Jamshoro

³ Registrar, University of Gawadar, Balochistan.

and un-aspirated sound is gauged through voice onset time (VOT) which was coined and invented by Lisker and Abramson (1964) who of the view that VOT is the time period between the burst of the stop closure and the following vowel voicing at onset position (Docherty, 1992). English stops sound creates such articulation where the active and passive articulators come across consequently burst takes place to produce such sound. According to Syed (2012) when a native speaker produces an English voiceless stop, the vocal fold starts vibrating longer after the burst which is measured in Milliseconds (ms). There are three kinds of VOTs such as short-lag VOT which is also called short positive VOT whereas long-lag VOT called long positive VOT. It is called pre-voicing if the vocal fold vibrates before the burst. Short-leg VOT represents un-aspirated and long-leg VOT indicates aspirated sounds. English un-aspirated voice and voiceless stops have short-leg VOT values whereas aspirated voiceless plosives carry long-leg VOT values. According to Syed (2013), the spread of glottis shows the difference between aspirated and un-aspirated stops. Aspiration contrast in English is allophonic and the stops are making aspirated sounds [p^h t^h k^h] whereas un-aspirated sounds are [p t k]. Urdu is the national language of Pakistan in which aspiration contrast is phonemic which makes a minimal pair but in English aspiration contrast does not make a minimal pair. According to Liljegren and Khan (2016), the feature of spread glottis is active in the Khowar phonemic inventory. The current study is focused on the acquisition of English un-aspirated stops /p /t /k/ and aspirated stops [p^h t^h k^h].

Literature Review

The acquisition of L2 phonological contrasts has been extensively studied within the frameworks of the Speech Learning Model (SLM) (Flege, 1995) and the Perceptual Assimilation Model (PAM) (Best, 1995). Both models suggest that L2 learners map unfamiliar sounds onto the closest L1 categories, which can result in difficulty acquiring subtle phonetic distinctions. Previous studies on L2 aspiration contrast (Lisker & Abramson, 1964; Cho & Ladefoged, 1999) indicate that speakers of languages with phonemic aspiration struggle with English allophonic aspiration due to their reliance on phonemic distinctions rather than allophonic variation.

Studies on South Asian language speakers, including Urdu and Hindi speakers, have demonstrated a tendency to treat English aspiration contrast as phonemic rather than allophonic (Syed, 2012; Rafat, 2016). This aligns with research on Khowar, where aspiration is contrastive, leading to potential misperception of English aspirated stops as distinct phonemes rather than allophones (Khan, 2020). The influence of orthographic representations has also been noted in L2 phonological acquisition (Bassetti, Escudero & Hayes-Harb, 2015), with non-Roman scripts reinforcing phonemic categories over allophonic variation.

Methodology

A production test was employed to measure the VOT values of English voiceless stops. Thirty Khowar L1 speakers (Pakistan-based and UK-based) participated in the study. Each participant was asked to produce words containing the target sounds. Each word was repeated three times, and a paired sample test was conducted to compare the VOT values of Khowar learners with those of native English speakers.

Research Questions

1. How do Khowar speakers acquire the English allophonic aspiration contrast?
2. What role does exposure to native English input play in the acquisition process?
3. How do Pakistan-based and UK-based Khowar speakers differ in their production of English aspirated and unaspirated stops?

Objectives

1. To examine the acquisition of English allophonic aspiration contrast by Khowar speakers.
2. To analyze the influence of exposure to native English input on the acquisition of aspiration contrast.
3. To compare the production accuracy of Pakistan-based and UK-based Khowar speakers.

Analysis and Discussion

To find out the acquisition of aspiration contrast of English voiceless plosives by two groups based in Pakistan and UK of Khowar participants as aspiration contrast is allophonic in English but phonemic in Khowar. If aspiration already exists in the L1 of Khowar which means that Khowar L1 speakers are familiar with the use of aspiration so, this study looked into their development of the acquisition process of phonemic contrast could be converted into allophonic contrast but the result shows that Pakistan-based did not show the desired results as the VOT value of Pakistan-based group is different from UK-based Khowar participants but in contrast, we can see better performance by UK-based group which is confirmed by paired-test which is significant. While looking into the averaged VOT values of the UK-based group are not up to the VOT values of English which means that they did not produce native-like aspiration, but they produced according to their L1 range which also endorsed the notion of SLM that the learners will equate the sound of L2 with the sound of L1 which is phonemic which is also called positive transfer of L1 sound which is occurred with the group Khowar UK-based only. Contrary, the Pakistan-based group did not transfer L1 phonemic contrast which is obvious from the result. The UK-based group has come to know about the aspiration of sound which is due to the input they received from the native, but this acquisition of aspiration can be considered as a new separate phonetic category for English voiceless stops allophonic variant which is not like-native sound rather the transfer of VOT values from L1. The new phonetic category is not native-like but existing literature named it deflected away from native sound (Syed, 2012).

It is noted that the Pakistan-based group did not produce English allophonic aspiration rather they substituted the aspirated coronal and dorsal sounds with the un-aspirated sound of L1. Let's discuss each aspirated and un-aspirated sound production of each group. The average VOT value for the un-aspirated Pakistan-based group for the sound [t] is 15.73 and for the sound [k] is 20.87 which is lesser than the UK-based group which means that UK-based participants have shown better results in the production of the un-aspirated sounds. The VOT values of UK-based Khowar participants for un-aspirated coronal sound [t] is 51.22 ms (millisecond) and for dorsal sound [k] is 20.87 ms. The VOT values for un-aspirated coronal sound [t] of UK-based is higher than native English which indicates the idea that foreign-based participants produced English un-aspirated

stop as aspirated which further gives a notion that the foreign group has learned aspiration contrast at a phonemic level rather than allophonic level that shows they did not learn the allophonic variation of English aspiration contrast. Therefore, even the un-aspirated sound of English coronal stop is taken as aspirated. UK group performed better at dorsal un-aspirated with an average VOT is 29.15 ms and a similar sound VOT of the Pakistan-based group is 20.87 ms. Thus, the foreign-based group showed a positive transfer of VOT for dorsal which is equal to the VOT values of English which is 29.42 ms. Thus, it is confirmed that both groups of Khowar learners produced dorsal un-aspirated stops without aspiration. However, it was expected that UK-based participants would produce dorsal sound [k] with aspiration as they did in the production of un-aspirated coronal sound which has greater VOT than native English but in the result of un-aspirated dorsal sound foreign group did not create such hyper-correction rather produced like native which is un-aspirated. In addition, the Pakistan-based group VOT is lower than the foreign group which means that the foreign group performed better but it is confirmed that both groups produced dorsal un-aspirated sound [k] with aspiration, and the positive transfer of L1 to L2 is taken place.

It was found from the above analysis that Khowar participants-based Pakistan produced English coronal un-aspirated [t] and dorsal un-aspirated [k] without aspiration which is because of the influence of L1. These un-aspirated sounds were given in the word 'store' and 'skin' and these sounds in such positions of word are produced without aspiration by native English.

Table

Average VOT values of English voiceless plosives

Sound		Pakistan-based	UK-based	English
store	St	15.73	51.22	24.00
teacher	t ^h	23.36	62.80	80.00
skin	sk	20.87	29.15	29.42
kitchen	k ^h	29.10	39.56	82.00

The above table shows the performance of both groups is significantly different in terms of place of articulation and aspiration contrast which shows the different performance of both groups. The results also show the increase of VOT values from coronal to dorsal which shows the place of articulation (Syed, 2013). While keeping in view the range of VOT values in English native which are mentioned in Table 6.1, the Pakistan and UK-based Khowar L1 participants' VOT values are measured to be able to see the closer range to the VOTs of native speakers. Larson-Hall (2010) also endorsed the effective criteria of VOT value through VOT range measurement. The results show that the Pakistan-based group's VOT values are below the range from the native English VOT at both coronal and dorsal positions, but UK-group manifests a higher VOT range for un-aspirated coronal stop and closer to the range of aspirated coronal English native.

Table shows that the Pakistan-based group VOTs for English coronal aspirated sound is 23.3693 which indicates that they have produced the given sound without aspiration whereas the UK-based group VOT is 62.8081 which is nearer to the range of native English which means that foreign-based produced English coronal aspirated sound with aspiration. As aspirated sound is present in the L1 of Khowar speakers therefore, there positive transfer of L1 into L2 and this could be taking place due to the input they received from the native environment. Moreover, the Pakistan-based group's VOTs for the aspirated dorsal stop is 29.10 which is lesser than the VOT of the UK-based

group which is 39.56 which means that the foreign group increased VOT. Thus, both UK groups have improved in learning aspirated dorsal English sounds. The Pakistan-based group has equated English aspirated sound with un-aspirated dorsal sound. According to the findings of the result, the foreign-based have produced both sounds at the coronal and dorsal position due to the input they received abroad which made them conscious about the aspirated sound in their L1 in the form of phoneme therefore, they produced aspirated both English sound at word initial position like in the word 'teacher' as well as in the consonant cluster which is not aspirated like in the word 'store'. These findings show that Khawar UK-based did not learn English allophonic variations rather they are considered as phonemic. Therefore, they produced with aspiration at both target positions and such correction is considered as hyper-correction which is done by UK-based foreign participants. As far as Pakistan-based participants are concerned, their VOT values show that they have produced English coronal and dorsal aspirated sounds without aspiration thus equivalence classification is taken place where both English allophonic aspirations equated with the phoneme of L1, and this happened due to the poor understanding of English allophonic variations. Therefore, the VOT values of both coronal and dorsal stops are equal to the VOT values of English un-aspirated stops even less than the minimum VOT values of English aspirated stops (Syed, 2013). This could also be the reason that Pakistan-based learners did not have the opportunity of foreign input which is why they had not become conscious of English allophonic variations.

Another reason for equivalence classification is the orthography, as Urdu is the national language of Pakistan, and Urdu grapheme for aspirated and un-aspirated is represented differently. The same different grapheme representation for both aspirated and un-aspirated sounds in Khawar but speakers mostly correspond in Urdu therefore they are more conscious about the difference of sounds for aspiration and un-aspiration. Secondly, the teachers are also unaware of English allophonic differences in Pakistan, and they teach English the way they have learned Urdu therefore, when they see English sounds with allophones which is orthographically not aspirated for them therefore, they took them as un-aspirated sounds. Hence, regional language orthographies also affect the acquisition of English allophones. Orthography established influence languages ascertained by existing literature (LaCharite & Paradis, 2005; Syed, 2012).

Moreover, the reason for poor learning of aspiration contrast by the Pakistan-based group is that the aspiration contrast is allophonic in the L2 but phonemic in the L1 which is why allophonic aspiration contrast is equated with L1 phoneme.

It is noticed from the result that English un-aspirated sounds on VOT which is performed better by Pakistan-based Khawar learners but poor performance on VOT of English aspirated sounds which is because of positive transfer of L1 for un-aspirated sounds and negative transfer for aspirated sounds which is because of considering English allophones of L2 taken as one sound of L1 that is un-aspirated therefore, they merged the two allophones of English into closest one of L1 and the performance for aspiration sound of English is also poor which is the reason of less VOTs because the aspirated sounds produced un-aspirated so, to acquire aspirated stops of English Khawar learners of Pakistan-based participants need to increase VOT values of aspirated stops.

It is observed from the findings that the VOT values of both groups are observed more VOT at dorsal than coronal positions which poses two reasons. First is the reason for the large contact area of the tongue between the active and passive parts of articulators which is prone to give bigger VOT. Secondly, vocal folds reside closer to the dorsal part of the tongue than the coronal therefore,

they give more VOT than the coronal. Structurally, the coronal part of the tongue carries a shorter contacting area than the dorsal which is the reason that coronal sounds have lesser VOT than the dorsal in this study (Kent & Read, 2002; Syed, 2012).

The impact of markedness is also crucial in the acquisition of L2 sounds. If the effect of markedness of L1 is stronger then there is a possibility of weaker learning on coronal and dorsal position in learning L2 sounds so, the impact of markedness is observed in this study and such similar cases are also attested in other studies (Syed, 2012).

If the poor learning of English voiceless stops is studied in terms of features therefore the features of English coronal and dorsal sounds are compared with Khowar's similar features to decipher the production style of the learners. As discussed earlier the amount of range of aspiration between English dorsal stops and Khowar speakers provided the difference. However, English coronal stops are concerned, they are also called alveolar sounds but contrary to alveolar Khowar has retroflex stops. The features of English alveolar is [-distributed +anterior] whereas Khowar L1 is either retroflex is [-distributed -anterior] or dental stop [-distributed -anterior] thus, Khowar speaker produced English alveolar stop either retroflex or dental stop but according to the findings of the previous part of this study which stated that Khowar L1 speakers produced English fricative sounds as a dental stop because the dental place is filled by dental fricative therefore, English coronal sound is replaced by the retroflex stop of L1. Thus, Khowar L1 speakers produce English coronal with drifted place of articulation causing not native-like VOT which is the reason that Khowar speakers produce English coronal stop differently from English speakers.

Conclusion

It is found that when there is an absence of native input of English then Khowar learners relate English aspirated allophones with the un-aspirated sounds of L1 and such finding came to the surface with the Pakistan-based Khowar participants. This phenomenon is also found in other languages of Pakistan (Rahman, 1991; Mahboob & Ahmar, 2004; Syed, 2012, 13). However, contrary to the results of Pakistan-based participants, and UK-based participants came with different findings as they received native input therefore, they acquired aspiration of English stops but they did not learn the allophonic variation of English which is aspirated at the initial position of the word (teacher) and without aspirated in the consonant cluster (store) but Khowar participants took the un-aspirated sound as aspirated and which has created hyper-correction (Syed, 2013) and due to such hyper-correction they have increased the amount of VOT for English un-aspirated stops to such extend which is bigger than the VOT of native English stops. They have also created a separate phonetic category for L2 allophones with the substitution of L1 phonemes.

References

- Bassetti, B., Escudero, P., & Hayes-Harb, R. (2015). Second language phonology at the interface between acoustic and orthographic input. *Phonology and Second Language Acquisition*, 191–210.
- Best, C. T. (1995). A direct realist view of cross-language speech perception. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 171–204). York Press.

Cho, T., & Ladefoged, P. (1999). Variation and universals in VOT: Evidence from 18 languages. *Journal of Phonetics*, 27(2), 207–229. <https://doi.org/10.1006/jpho.1999.0094>

Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. In W. Strange (Ed.), *Speech perception and linguistic experience: Issues in cross-language research* (pp. 233–277). York Press.

Khan, S. U. (2020). Phonemic aspiration contrast in Khovar: An acoustic analysis. *Journal of South Asian Linguistics*, 12(1), 45–62.

Kent, R. D., & Read, C. (2002). *The acoustic analysis of speech* (2nd ed.). Singular Publishing Group.

Lisker, L., & Abramson, A. S. (1964). A cross-language study of voicing in initial stops: Acoustical measurements. *Word*, 20(3), 384–422. <https://doi.org/10.1080/00437956.1964.11659830>

Rafat, Y. (2016). The acquisition of English stop aspiration by native speakers of Urdu: The influence of orthography. *Journal of Phonetics*, 55, 19–37. <https://doi.org/10.1016/j.wocn.2016.01.003>

Syed, H. (2012). Perception and production of English stops by Urdu speakers. *Second Language Research*, 28(2), 131–153. <https://doi.org/10.1177/0267658311431741>