Galore International Journal of Health Sciences and Research
[Indexed, Peer-Reviewed Journal]
Vol. 11; Issue: 3; July-Sept. 2026
Website: www.gijhsr.com
P-ISSN: 2456-9321
Original Research Article

Voice Onset Time Transfer Patterns in Assamese-English Bilinguals: Evidence for Asymmetric Cross-Linguistic Influence

Debarshi Das[1], Dr Anne Varghese[2]

[1]Assistant Professor, Department of Speech-Language Pathology, NAISH College, Bengaluru, India.
[2]Professor, Department of Speech-Language Pathology, NAISH College, Bengaluru, India.

Corresponding Author: Debarshi Das

DOI: https://doi.org/10.52403/gijhsr.20260331

ABSTRACT

The Voice Onset Time (VOT) was examined in word-initial stops produced by thirty Assamese–English bilinguals (15 male, 15 females; aged 18–25 years) to characterise the pattern of cross-linguistic influence (CLI) between a four-way laryngeal system (Assamese)(L1) and a two-way system (English)(L2). Each speaker produced 96 tokens, covering voiced unaspirated, voice aspirated, voiceless unaspirated, and voiceless aspirated stops in word and non-word contexts in both languages, embedded in controlled carrier phrases. VOT was measured in Praat software from synchronised waveform and spectrogram displays, following current conventions for defining release and voicing onset. Paired comparisons of by-speaker VOT means revealed a category-specific, asymmetric transfer pattern. Voiced unaspirated stops showed a marginal cross-language difference (Assamese M = 80.0 ms vs. English M = 88.4 ms; d = −0.48, p = .064), with English /b d g/ carrying substantial pre-voicing that is rarely reported for monolingual English speakers. English /p t k/ were produced with short-lag VOT (M = 20.7 ms) rather than the long-lag values typical of native English, indicating systematic L1→L2 transfer of the Assamese voiceless unaspirated category. The aspirated categories (breathy /dʱ/ and voiceless aspirated /tʰ/) did not differ across languages, consistent with wholesale transfer of L1 categories onto the English dental fricatives /θ ð/. Lexicality (word vs. non-word) had no effect on VOT in any category, confirming the phonetic robustness of the contrasts. Significant Gender vs Language interactions emerged, most notably for voiced unaspirated stops in English (males 103.5 ms vs. females 73.3 ms, p < .001). The findings support the Revised Speech Learning Model in predicting that bilingual categories share a common phonetic space, and show that CLI in this population is not uniform but is shaped by the structural mismatch between the two inventories.

Keywords: voice onset time; cross-linguistic influence; Assamese–English bilinguals; L1–L2 transfer; SLM-r; Indian English phonetics; laryngeal contrast

1. INTRODUCTION

Voice Onset Time (VOT), the interval separating the release of a stop closure from the onset of vocal-fold vibration, remains, sixty years after Lisker and Abramson (1964) gave it its name, the most widely used acoustic measure for describing stop categories across languages and for diagnosing bilingual phonetic transfer [1,2]. Two properties explain its continued usefulness. First, VOT is a articulatory parameter that speakers cannot readily monitor [3], it offers a sensitive window onto the underlying phonological categories a speaker has established. Second, VOT distributions differ in predictable, quantifiable ways across languages, which makes it possible to compare monolingual norms against bilingual productions with some precision [4,5].

When a coordinate bilingual carries two languages whose stop inventories are unevenly matched, the categories of the two systems may interact, shift, or reorganise. This is what is commonly called cross-linguistic influence (CLI), and it has been documented for many language pairings over several decades [6-8]. Recent work has moved beyond asking whether transfer occurs and has begun to ask, more usefully, what shape it takes: which categories prove most vulnerable, whether influence flows in one direction or both, and which structural properties of the two inventories predict the pattern observed [9]. The Revised Speech Learning Model [9] proposes that bilingual categories occupy a shared phonetic space and that the perceived similarity between L1 and L2 sounds governs whether new categories are formed or existing ones are merged. Similar predictions, formulated for perceptual assimilation rather than production, are made by PAM-L2 [10].

1.1 The Assamese–English laryngeal mismatch

Assamese and English offer a theoretically informative contrast. Assamese, an Indo-Aryan language spoken by over fifteen million people in North-East India, maintains a four-way laryngeal contrast in its stop series: voiceless unaspirated, voiceless aspirated, voiced unaspirated, and voiced aspirated [11]. This contrast is shared across most Indo-Aryan languages and has been acoustically documented in Hindi, Urdu, and Marathi [12-14], among others. English, by contrast, works with only a two-way contrast, realised chiefly through short-lag versus long-lag VOT in word-initial position; it has no phonological aspiration contrast and no breathy-voiced series [15,16].

The mismatch produces uneven mapping pressures. Some Assamese categories have close English analogues: voiceless aspirated /pʰ tʰ kʰ/ maps reasonably well onto English /p t k/ in stressed onsets, and the voiced series /b d g/ shares voicing as a shared feature across both languages. Other Assamese categories have no direct counterpart: voiceless unaspirated /p t k/ has no equivalent in English, and aspirated-voiced /bʱ dʱ gʱ/ is altogether absent. SLM-r predicts that this uneven mapping will produce asymmetric transfer rather than uniform drift [9]. Categories that lack an L2 analogue should either retain their L1 specification wholesale, because there is no L2 pressure to modify them, or be redeployed in L2 contexts in ways that reflect their L1 origin. Categories with close analogues, in contrast, should interact more strongly and may show merging tendencies [17].

1.2 The Indian English context

Indian English is now broadly recognised as a nativised variety with its own phonological profile, many features of which can be traced to speakers' L1 (Wells, 1982; Wiltshire, 2005; Sailaja, 2012). Two features stand out for a study of VOT. First, English /p t k/ in Indian English is typically produced as voiceless unaspirated rather than with the long-lag aspiration of General American or Received Pronunciation [18,19]. Second, the English interdental fricatives /θ ð/ are generally replaced with dental plosives, often realised as /tʰ/ and /dʱ/ respectively, reflecting the L1 Indo-Aryan inventory [18,20]. Despite Assamese being a major Indo-Aryan language, its speakers have received less attention in the phonetic literature than speakers of Hindi, Bengali, or Tamil. The earliest systematic description of Assamese phonetics is Mahanta's (2012) IPA illustration [11]; a more recent study by Sarmah and Wiltshire (2024) [21] examined vowel production in Assamese Indian English and found that L1 properties are retained even in English-medium-educated speakers, with some convergence toward a broader Indian English norm. No prior study, to our knowledge, has systematically documented VOT across the full four-way Assamese stop contrast in bilinguals who speak both Assamese and English.

1.3 Research questions and hypotheses

The present study addresses four questions:

RQ1. How do Assamese–English bilinguals realise VOT for stops in each of their two languages?

RQ2. Is cross-linguistic influence uniform across all stop categories, or does it follow the pattern of match and mismatch between the two inventories?

RQ3. Does lexicality (real word vs. non-word) modulate VOT, or is VOT determined purely by the phonetic category?

RQ4. Do speaker-level factors such as gender interact with language and category?

Guided by SLM-r (Flege & Bohn, 2021) and prior findings on VOT in Indian English (Wiltshire, 2005; Sailaja, 2012; Sarmah & Wiltshire, 2024), we propose the following:

- **H1 (bidirectional influence).** Bilingual VOT will deviate from expected monolingual norms in both languages, indicating a shared phonetic space.
- **H2 (categorical asymmetry).** Categories without close cross-linguistic analogues — Assamese voiceless unaspirated and breathy-voiced — will be transferred wholesale into English, producing English realisations close to their Assamese counterparts.
- **H3 (directional asymmetry).** L1→L2 transfer will dominate, showing up as Assamese-like VOT in English productions, especially for the voiceless series.
- **H4 (lexicality invariance).** VOT will not differ between words and non-words, consistent with its status as a category-level phonetic property rather than a lexically stored cue (Klatt, 1975).

2. MATERIALS & METHODS

2.1 Participants

Thirty native Assamese speakers (15 female, 15 male; age range 18–25 years) took part. All had completed their schooling up to the 12th standard in English-medium institutions in Assam and self-reported proficiency in both Assamese (L1) and English (L2). Participants were selected by purposive sampling on the following criteria:

Inclusion criteria:

- Age between 18 and 25 years
- Completion of schooling up to the 12th standard in an English-medium institution
- Self-reported proficiency in both L1 and L2

Exclusion criteria:

- Knowledge of any additional language beyond L1 and L2
- Parents with different L1s
- Any history of formal vocal training
- Any documented neurological, vascular, or motor condition that could affect articulation, phonation, or respiration

All participants were non-smokers. Written informed consent and demographic information (age, gender, educational status, occupation, medical and mental history, and languages known) were obtained prior to recording.

2.2 Stimuli

Each participant produced 96 tokens in total: four cells (voiced × word, voiced × non-word, voiceless × word, voiceless × non-word) for each of the two languages. Each cell contained twelve CV-initial items obtained by crossing four stop categories with three vowel contexts (/a i u/): ***(Appendix 1-Word List**, is attached in the Abstract HTML page**)**. [***Click here for Appendix 1-Word List***]

Category	Assamese stops tested	English stops tested
Voiced unaspirated	/ba, bi, bu, da, di, du, ga, gi, gu/	/ba, bi, bu, da, di, du, ga, gi, gu/
Breathy-voiced	/dʱa, dʱi, dʱu/	/ðV/ items realised as /dʱ/ in Indian English
Voiceless unaspirated	/pa, pi, pu, ta, ti, tu, ka, ki, ku/	/pa, pi, pu, ta, ti, tu, ka, ki, ku/
Voiceless aspirated	/tʰa, tʰi, tʰu/	/θV/ items realised as /tʰ/ in Indian English

The English bisyllabic words and non-words were constructed in combination with CV Structure. For Assamese, bisyllabic words and non-words were constructed to match the English items on initial-CV identity, with the rest of the segmental material varied. Non-words were matched to real words only on their initial consonant and vowel; they differed in the remaining material to ensure that no semantic or frequency information could attach to them. This follows the standard use of nonce controls in phonetic production research [22,23].

2.3 Recording procedure

All recording took place in an acoustically treated chamber. Ambient noise levels were verified before each session using the Decibel X: dB Sound Level Meter mobile application (version 9.8.1; Thanh Dinh). An external lavalier microphone (Boya BY-M1) was positioned approximately 10–12 cm from the speaker's mouth at a 45° angle, following ASHA (2023) recording guidelines. Recordings were captured directly onto a MacBook Air (M1) using PRAAT (version 6.0.43; Boersma & Weenink, 2010) at a sampling rate of 44.1 kHz.

Stimuli were presented in written form and embedded in language-appropriate carrier phrases to stabilise prosody and rate across tokens [1]:

- English: "This is _____" /ðɪs ɪz .../
- Assamese: এইটো হয় _____ /ˌaitu hɔɪ .../

Each participant was seated in an ergonomic posture and read the randomised stimulus list at a comfortable conversational rate. The two languages were recorded in separate blocks, with block order counterbalanced across participants to minimise order effects [7].

2.4 VOT measurement

VOT was measured offline in PRAAT from synchronised waveform and wide-band spectrogram displays, following the conventions outlined by Abramson and Whalen (2017) [1]. For each token, two landmarks were identified: (a) the burst release of the initial stop, marked at the onset of aperiodic energy following closure, and (b) the onset of periodicity, marked at the first glottal pulse associated with the following vowel. VOT was computed as the signed interval between these two landmarks. Positive values (lag VOT) indicate that voicing began after release, while negative values (lead VOT) indicate voicing during closure. For the statistical analyses and graphical display that follow, pre-voicing magnitudes are reported as absolute values, as is standard in cross-linguistic VOT comparisons involving languages that differ in the range of voicing lead they admit [2,4].

Statistical Analysis

2.5 Data preparation and statistical analysis

The full dataset comprised 2,880 observations (30 participants × 96 tokens). By-speaker means were computed by averaging VOT across place of articulation, vowel, and lexicality within each language × category cell, yielding a 30 × 8 matrix for statistical testing.

The primary analysis was a paired-samples t-test (Assamese vs. English, within speakers) for each of the four stop categories, with Cohen's d reported as a standardised effect-size measure [24]. Gender effects were tested using Welch's independent-samples t-tests to accommodate unequal variances. Lexicality was tested using paired t-tests within each language × category cell. All analyses were conducted in Python 3 using the scipy.stats library (v. 1.13). An alpha level of .05 was adopted throughout, with $p < 10$ treated as a trend. Because the Gender × Language interaction analyses were exploratory, uncorrected p-values are reported, but their interpretation is qualified by the number of comparisons involved.

3. RESULT

3.1 VOT by language and category

Table 1 shows mean VOT values, standard deviations, and sample sizes for each stop category in each language. Figure 1 displays the same information graphically.

Table 1. Mean VOT (ms) by language and stop category (by-speaker means, N = 30).

Category	**Assamese *M* (SD)**	**English *M* (SD)**	**Δ (Eng − Ass)**	***t* (29)**	***p***	**Cohen's *d***
Voiced unaspirated /b d g/	80.02 (8.91)	88.39 (22.31)	8.37	-1.93	0.064	-0.48
voiced aspirated /d^{ɦ}/	66.52 (21.27)	68.84 (18.55)	2.32	-0.54	0.591	-0.12
Voiceless unaspirated /p t k/	21.93 (3.05)	20.72 (3.07)	-1.21	1.69	0.102	0.39
Voiceless aspirated /t^{h}/	65.03 (9.64)	66.12 (9.12)	1.09	-0.58	0.566	-0.12

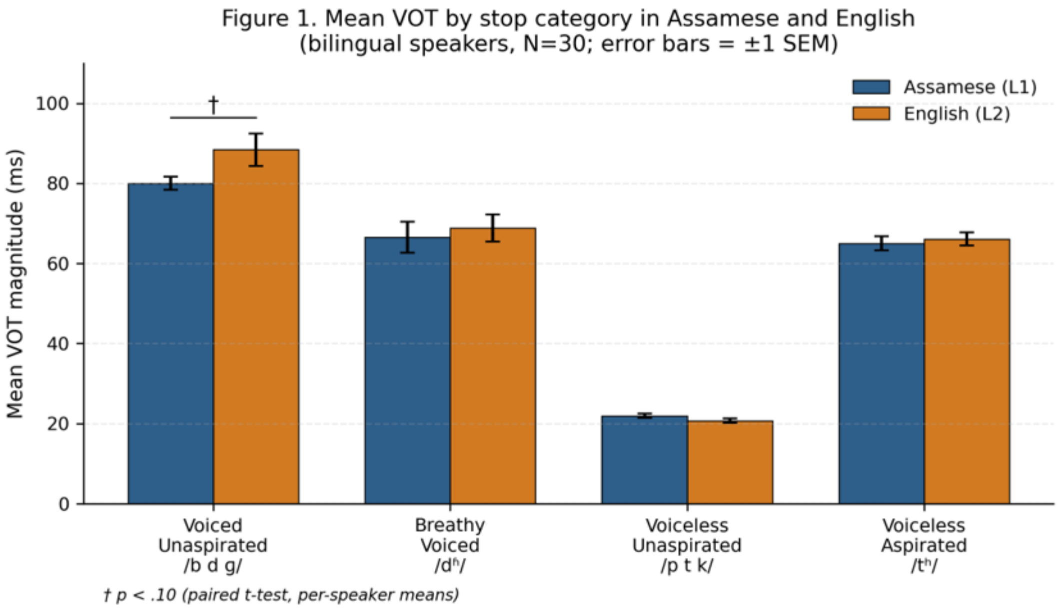

Figure1. Mean VOT by stop category in Assamese and English (bilingual speakers, N=30; error bars = ±1SEM) Note. $p < .10$. Voiced-category values are reported as pre-voicing-lead magnitudes; voiceless-category values represent post-release lag.

Three observations are immediate. First, three of the four stop categories showed essentially identical VOT across Assamese and English (all $|\Delta| \leq 2.4$ ms, all $|d| \leq 0.12$). Second, the one category showing a cross-language difference, voiced unaspirated, did so in a direction informative for the transfer hypothesis: bilinguals produced English /b d g/ with longer pre-voicing (88.4 ms) than their Assamese counterparts (80.0 ms). Prevoicing of this magnitude is unusual in native English [5,15,16], which generally admits only variable short-lag realisations of /b d g/. Third, the English /p t k/ values of only 20.7 ms fall well below the long-lag values of 60–80 ms reported for General American and Received Pronunciation [2,15]. This gives direct acoustic evidence of L1-category transfer and converges with earlier descriptions of Indian English voicing [18-20].

3.2 Place-of-articulation effects

To check whether the cross-linguistic patterns held uniformly across places of articulation, VOT was tabulated by place × language (Table 2 and Figure 2).

Table 2. Mean VOT (ms) by place of articulation and language.

Place	Assamese *M* (SD)	English *M* (SD)
Bilabial /p/	15.78 (5.79)	12.24 (3.97)
Dental /t/	17.12 (5.68)	16.17 (4.81)
Velar /k/	32.90 (8.67)	33.76 (9.01)
Aspirated /tʰ/	65.03 (16.95)	66.12 (14.01)
Bilabial /b/	114.78 (23.81)	106.67 (32.75)
Dental /d/	70.72 (37.95)	77.97 (35.37)
Velar /g/	54.58 (26.15)	80.53 (36.64)
Breathy /dʱ/	66.52 (25.32)	68.84 (25.44)

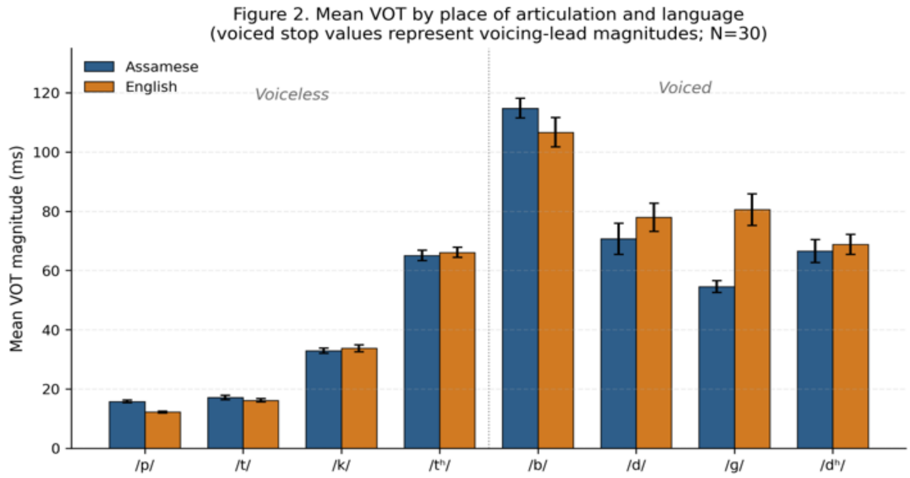

Figure 2. Mean VOT by place of articulation across Assamese and English. The voiceless series (/p, t, k, tʰ/) is shown on the left and the voiced series (/b, d, g, dʱ/) on the right. Cross-language differences are largest at the velar /g/ position.

The most striking place-related asymmetry emerged at the velar /g/, where English productions showed markedly longer pre-voicing (80.5 ms) than Assamese productions (54.6 ms). The pattern runs counter to the aerodynamic expectation. In monolingual production across many languages, velars typically show the shortest pre-voicing of the three places because a more posterior closure leaves less supra-glottal cavity in which voicing can be sustained [4,25]. One possible interpretation is that, in English production, these bilinguals may be compensating by extending pre-voicing at the velar position to keep it distinct from a nearby English voiceless competitor. This reading is speculative and would benefit from a targeted follow-up.

The voiceless unaspirated series followed an orderly place gradient in both languages: /p/ < /t/ < /k/. This replicates the well-attested aerodynamic constraint that more posterior places yield longer short-lag VOT [4,5]. Even the most posterior voiceless token (/k/: ~33 ms in both languages) sits comfortably within the short-lag range and does not approach English long-lag norms.

3.3 Lexicality effects

To address RQ3, VOT was compared between real words and non-words within each language × category cell. No significant lexicality effect was found for any category (all $p > .17$; Table 3). The largest numerical difference, 4.8 ms between English aspirated words and non-words, did not reach significance ($p = .167$).

The absence of any lexicality effect across eight independent comparisons supports the view that VOT is determined at the category level rather than being lexically indexed [1,22]. This null result also validates the use of non-word controls in CLI research on this population, since novel items behave identically to real items.

3.4 Gender × Language interactions

Gender effects differed markedly across categories (Table 4 and Figure 3). Within Assamese, males produced significantly shorter pre-voicing than females in the breathy-voiced category (54.5 vs. 78.5 ms, p = .001) and longer short-lag VOT in the voiceless unaspirated category (23.4 vs. 20.5 ms, p = .008). Within English, the most pronounced effect appeared in the voiced unaspirated category, where male pre-voicing reached 103.5 ms compared to 73.3 ms for females (p < .001). The direction of the gender effect for voiceless unaspirated stops reversed between the two languages — males were longer than females in Assamese but shorter in English (19.5 vs. 21.9 ms, p = .027).

Table 3. Effect of lexicality (word vs. non-word) on VOT, by language and category.

Language	Category	Word *M*	Non-word *M*	Δ	*t*(29)	*p*
Assamese	Voiced unaspirated	79.48	80.57	-1.09	-0.45	0.659
Assamese	Breathy-voiced	66.98	66.07	0.91	0.36	0.722
Assamese	Voiceless unaspirated	21.96	21.91	0.05	0.16	0.872
Assamese	Voiceless aspirated	64.50	65.56	-1.06	-0.60	0.553
English	Voiced unaspirated	88.68	88.10	0.57	0.26	0.796
English	Breathy-voiced	71.24	66.44	4.80	1.42	0.167
English	Voiceless unaspirated	20.68	20.76	-0.08	-0.17	0.869
English	Voiceless aspirated	65.69	66.56	-0.87	-0.51	0.613

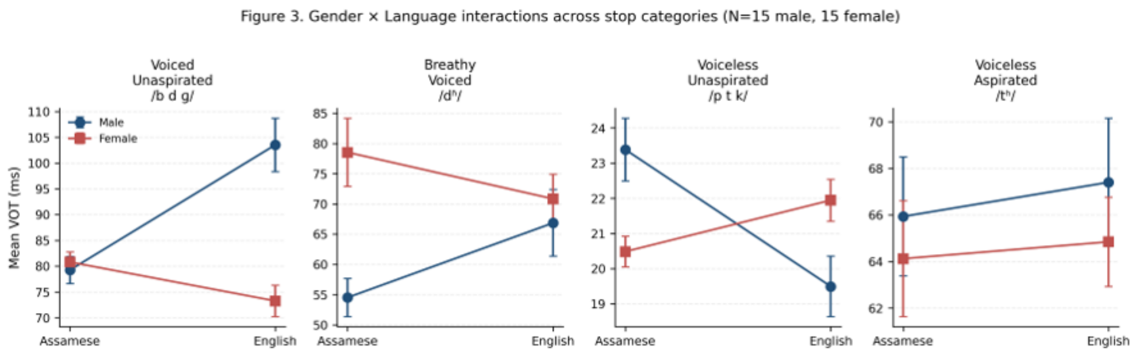

Figure 3. Gender × Language interactions across the four stop categories. Notable crossover interactions appear in the voiced unaspirated and voiceless unaspirated categories.

3.5 Individual speaker patterns

Figure 4 plots VOT for each speaker in Assamese against VOT in English for each category. The dashed diagonal marks perfect cross-language equivalence. Three category-specific patterns are visible.

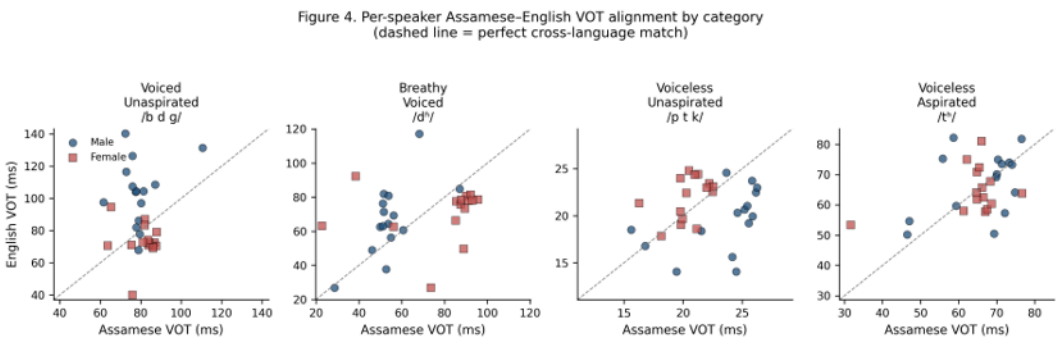

Figure 4. Assamese–English VOT alignment across the four stop categories, one point per speaker. The dashed line marks perfect equivalence. Male speakers (blue circles) cluster above the diagonal for voiced unaspirated stops, indicating longer English pre-voicing than Assamese pre-voicing.

For voiced unaspirated stops, male speakers cluster well above the diagonal (English > Assamese), while female speakers sit on or just below it. The overall cross-language difference reported in Table 1 is therefore driven primarily by male productions, a point we return to in Section 4.4. For breathy-voiced stops, speakers scatter broadly around the diagonal with no systematic directional bias, consistent with the absence of any cross-language mean difference. For voiceless unaspirated stops, most speakers fall close to the diagonal in the short-lag range (15–25 ms), confirming the stability of this category across languages. For voiceless aspirated stops, speakers again cluster close to the diagonal, though with wider variance. As expected, since measurement uncertainty tends to grow with longer VOT values [1].

Table 4. Gender comparisons by language and category.

Language	Category	Male *M* (SD)	Female *M* (SD)	*t*	*p*
Assamese	Voiced unaspirated	79.24 (10.27)	80.80 (7.59)	-0.47	0.640
Assamese	Voiced aspirated	54.51 (12.20)	78.53 (21.89)	-3.71	**0.001**
Assamese	Voiceless unaspirated	23.38 (3.45)	20.49 (1.69)	2.92	**0.008**
Assamese	Voiceless aspirated	65.93 (9.91)	64.12 (9.63)	0.51	0.616
English	Voiced unaspirated	103.52 (20.07)	73.26 (11.75)	5.04	**<.001**
English	Voiced aspirated	66.84 (21.32)	70.84 (15.80)	-0.58	0.564
English	Voiceless unaspirated	19.50 (3.32)	21.94 (2.30)	-2.35	**0.027**
English	Voiceless aspirated	67.40 (10.66)	64.84 (7.43)	0.76	0.453

4. DISCUSSION

The present study set out to describe the shape of cross-linguistic influence in Assamese–English bilinguals using VOT as an acoustic index. Our findings converge on a consistent picture: CLI in this population is not uniform but is patterned by the structural match or mismatch between the two laryngeal inventories.

4.1 Evidence for L1→L2 transfer

The clearest evidence of L1→L2 transfer lies in the English voiceless unaspirated category. Bilinguals produced English /p t k/ at a mean VOT of 20.7 ms, almost identical to their Assamese /p t k/ at 21.9 ms and well below the 60–80 ms long-lag values typical of native English in stressed initial position [2,15,16]. The near-identity of the two distributions, with a cross-language difference of only 1.2 ms (well within measurement noise), suggests that these speakers have not developed a distinct English aspirated category but instead deploy their native Assamese voiceless unaspirated category when speaking English. The pattern confirms acoustically what has long been noted descriptively for Indian English [18-20] and aligns with Sarmah and Wiltshire's (2024) [21] finding that Assamese speakers retain L1 phonetic properties even when producing English vowels.

The voiced-aspirated and voiceless-aspirated categories provide complementary evidence. English has no phonemic voiced aspirated or aspirated series [16], so English productions in these categories necessarily reflect L1 transfer: the dental fricative /θ/ is rendered as /tʰ/ and /ð/ as /dʱ/ [18,20]. The statistical identity of Assamese and English values in these two categories, mean differences of only 1.1 ms and 2.3 ms respectively, confirms that speakers are producing essentially the same phonetic object in both contexts. This supports **H2 (categorical asymmetry)**: categories without close L2 analogues are transferred wholesale rather than modified. The pattern is consistent with SLM-r's prediction that when no corresponding L2 category exists, learners do not establish a new one [9].

4.2 Evidence for L2-to-L1 influence, or its absence

SLM-r predicts that for categories with close cross-language analogues, the two systems should interact and produce bidirectional influence [9]. The voiced unaspirated category, where both Assamese and English possess /b d g/ and where both admit pre-voicing under some conditions [12,15], is the natural place to test this prediction. The marginal effect observed here (Assamese 80.0 vs. English 88.4 ms, $d = -0.48$, $p = .064$) points in the direction SLM-r would predict: the two categories appear to have assimilated to a common phonetic space, with the English realisation carrying the substantial pre-voicing characteristic of Assamese. Values of 85–105 ms, particularly those produced by the male speakers, fall well outside the usual native-English range reported in the literature [5,16].

The fact that the effect achieved only marginal significance in the pooled analysis but became highly robust once gender was considered, suggests that the population-level pattern is speaker-dependent. This interpretation is consistent with the general emphasis of SLM-r on individual variation in bilingual category formation [9], and with earlier findings that dominance and usage modulate the degree of CLI in bilingual VOT [7,17].

4.3 Lexicality invariance

None of the eight word-vs-non-word comparisons reached significance (all $p > .17$). This null result directly supports **H4**. Had VOT been lexically indexed. For example, if frequently used words were produced with reduced or exaggerated VOT [23]. We would expect consistent word-vs-non-word differences. Their absence confirms that the cross-linguistic patterns reported here reflect category-level phonetic implementation, not item-specific memory traces [1,22]. This is methodologically important: it validates the inclusion of non-word controls in future CLI research on Assamese–English bilinguals.

4.4 Gender × Language interactions

Perhaps the most unexpected result is the Gender × Language interaction for voiced unaspirated stops. Male speakers produced English pre-voicing of 103.5 ms compared with 79.2 ms in Assamese, while female speakers produced nearly equal values in the two languages (73.3 vs. 80.8 ms). A plausible interpretation is that male and female speakers adopt different phonetic strategies in L2 production. Male speakers may be hyper-distinguishing English /b d g/ from English /p t k/ by exaggerating pre-voicing, since their short-lag /p t k/ (around 20 ms) and a native-English-style short-lag /b d g/ would otherwise be acoustically crowded. Female speakers, by contrast, seem to maintain a narrower voicing-lead range across both languages, possibly reflecting a more compressed bilingual phonetic space. Similar speaker-level dominance effects on bilingual VOT have been reported by Antoniou et al. (2011)[17] for L2-dominant bilinguals in other language pairings.

A second interaction worth noting emerged for the breathy-voiced category in Assamese, where female speakers produced longer pre-voicing (78.5 ms) than male speakers (54.5 ms). This is consistent with well-documented anatomical and acoustic differences in glottal configuration between male and female speakers: breathiness is reported to be a more prominent feature of female speech across several languages [26,27]. The Gender × Language crossover for voiceless unaspirated stops (males longer in Assamese, shorter in English) is harder to interpret on physiological grounds alone and may reflect different usage or dominance patterns between the two groups.

4.5 Theoretical implications

Taken together, our findings align closely with the SLM-r prediction that bilingual phonetic categories are governed by the perceived similarity between L1 and L2 sounds, and that this similarity structure produces category-specific rather than uniform transfer [9]. In this population, transfer can best be characterised as asymmetric along three dimensions:

1. **Categorical asymmetry.** The four stop categories differ in the degree of CLI they show. Categories without a close L2

analogue are transferred wholesale (breathy /dʱ/, aspirated /tʰ/). The voiceless unaspirated category is also transferred wholesale because English /p t k/ offers only an imperfect match (long-lag aspiration is absent from the Assamese inventory). Only the voiced unaspirated category, where both languages share /b d g/ and where pre-voicing is a language-specific option. Shows a genuine cross-language difference, and that difference points in the direction SLM-r would predict.

2. **Directional asymmetry.** L1→L2 influence dominates the picture. English productions carry clearly Assamese phonetic properties (short-lag /p t k/, breathy /ð/, aspirated /θ/), while Assamese productions show no compensatory drift toward English norms. For a bilingual population educated in English-medium schools but presumably Assamese-dominant in daily life, L1 phonetic categories retain priority in L2 production. The result is consistent with reports from other Indian English varieties [20,21].
3. **Featural asymmetry.** Aspiration and short-lag voicelessness are transferred more completely than pre-voicing. This pattern aligns with the SLM-r claim that perceptually salient features survive cross-linguistic transfer more robustly. Aspiration and short-lag VOT are categorical distinctions in Assamese, whereas pre-voicing is gradient and sociolinguistically variable in English [5,16], which leaves more room for cross-language variation.

The findings contribute to the broader literature on Indian English phonetics by providing, for the first time to our knowledge, quantitative VOT data from Assamese–English bilinguals across the full four-way laryngeal system. They also add a new case study to the growing body of CLI research that tests SLM-r predictions in language pairings with large inventory mismatches [13,17].

4.6 Limitations and directions for future research

Several limitations should be acknowledged. First, we did not include a monolingual English control group. The comparison of bilingual English VOT with published monolingual norms is standard practice, but is inherently less precise than a matched-group comparison. Second, our participants were purposively sampled from a single geographic and educational background: 12th-standard English-medium graduates aged 18–25 years, so the generalisability of the findings to older bilinguals, to those educated in Assamese-medium schools, or to bilinguals with a different exposure profile remains to be established. Third, dominance and proficiency were classified on educational criteria rather than through formal instruments such as the Bilingual Language Profile [28]. Future studies should incorporate such instruments to enable a more direct test of dominance × category interactions.

Several directions suggest themselves. First, adding Assamese near-monolingual controls would allow a direct quantification of L2-to-L1 influence on Assamese /b d g/, a gap that the current within-speaker design can only address indirectly. Second, measurement of secondary cues: closure duration, f0 at vowel onset, and H1–H2 spectral tilt for aspirated stops [12,13], would provide a multi-dimensional characterisation of CLI that goes beyond VOT alone. Third, the gender effects reported here, particularly the crossover interactions, deserve targeted investigation in a design that can separate physiological from sociolinguistic sources of variation [27]. Finally, longitudinal study of English-medium students transitioning into English-dominant working environments would clarify whether the patterns documented here remain stable or shift with increasing L2 immersion.

CONCLUSION

Assamese–English bilinguals produce stop consonants with VOT patterns that reveal a systematic, category-specific pattern of cross-linguistic influence. Categories without close analogues in the two languages: voiced aspirated, voiceless aspirated, and voiceless unaspirated, are transferred wholesale from Assamese into English, producing the characteristically Indian-English phonetic profile long described in the descriptive literature. The one category with a close cross-language analogue, voiced unaspirated, shows a marginal cross-language difference driven largely by male speakers, consistent with the SLM-r prediction that shared categories interact in a common phonetic space. Lexicality has no effect on VOT, supporting its status as a category-level phonetic property. The findings offer a first systematic VOT characterisation of Assamese–English bilingual speech, provide a new test case for SLM-r in a four-way-to-two-way laryngeal mapping, and show that CLI in this population is best understood not as uniform drift but as structured, asymmetric transfer.

Declaration by Authors
Ethical Approval: Approved
Acknowledgement: None
Source of Funding: None
Conflict of Interest: The authors declare no conflict of interest.

REFERENCES

1. Abramson AS, Whalen DH. Voice Onset Time (VOT) at 50: Theoretical and practical issues in measuring voicing distinctions. Journal of Phonetics [Internet]. 2017 May 23;63:75–86. Available from: https://doi.org/10.1016/j.wocn.2017.05.002
2. Cho T, Whalen DH, Docherty G. Voice onset time and beyond: Exploring laryngeal contrast in 19 languages. Journal of Phonetics [Internet]. 2018 Dec 11;72:52–65. Available from: https://doi.org/10.1016/j.wocn.2018.11.002
3. Flege JE. Age of learning affects the authenticity of voice-onset time (VOT) in stop consonants produced in a second language. The Journal of the Acoustical Society of America [Internet]. 1991 Jan 1;89(1):395–411. Available from: https://doi.org/10.1121/1.400473
4. Cho T, Ladefoged P. Variation and universals in VOT: evidence from 18 languages. Journal of Phonetics [Internet]. 1999 Apr 1;27(2):207–29. Available from: https://doi.org/10.1006/jpho.1999.0094
5. Chodroff E, Golden A, Wilson C. Covariation of stop voice onset time across languages: Evidence for a universal constraint on phonetic realization. The Journal of the Acoustical Society of America [Internet]. 2019 Jan 1;145(1):EL109–15. Available from: https://doi.org/10.1121/1.5088035
6. Flege JE. The production of “new” and “similar” phones in a foreign language: evidence for the effect of equivalence classification. Journal of Phonetics [Internet]. 1987 Jan 1;15(1):47–65. Available from: https://doi.org/10.1016/s0095-4470(19)30537-6
7. Simonet M. Phonetic consequences of dynamic cross-linguistic interference in proficient bilinguals. Journal of Phonetics [Internet]. 2014 Feb 13;43:26–37. Available from: https://doi.org/10.1016/j.wocn.2014.01.004
8. Chang CB. Rapid and multifaceted effects of second-language learning on first-language. speech production. Journal of Phonetics [Internet]. 2011 Dec 31;40(2):249–68. Available from: https://doi.org/10.1016/j.wocn.2011.10.007
9. Flege JE, Bohn OS. The Revised Speech Learning Model (SLM-R). In: Cambridge University Press eBooks [Internet]. 2021. p. 3–83. Available from: https://doi.org/10.1017/9781108886901.002
10. Best CT, Tyler MD. Nonnative and second-language speech perception. In: Language learning and language teaching [Internet]. 2007. p. 13–34. Available from: https://doi.org/10.1075/lllt.17.07bes
11. Mahanta S. Assamese. Journal of the International Phonetic Association [Internet]. 2012 Aug 1;42(2):217–24. Available from: https://doi.org/10.1017/s0025100312000096
12. Dutta, Indranil. Four-way Stop Contrasts in Hindi: an acoustic study of voicing, fundamental frequency and spectral tilt. [Internet] [Doctoral Dissertation]. University of Illinois; 2007. Available from: https://www.researchgate.net/publication/36711132_Four-

way_stop_contrasts_in_Hindi_An_acoustic_study_of_voicing_fundamental_frequency_and_spectral_tilt
13. Schertz J, Khan S. Acoustic cues in production and perception of the four-way stop laryngeal contrast in Hindi and Urdu. Journal of Phonetics [Internet]. 2020 Jun 5;81:100979. Available from: https://doi.org/10.1016/j.wocn.2020.100979
14. Dmitrieva O, Dutta I. Acoustic correlates of the Four-Way laryngeal contrast in Marathi. Phonetica [Internet]. 2019 Sep 18;77(3):209–37. Available from: https://doi.org/10.1159/000501673
15. Lisker L, Abramson AS. A Cross-Language study of voicing in initial stops: Acoustical measurements. WORD [Internet]. 1964 Jan 1;20(3):384–422. Available from: https://doi.org/10.1080/00437956.1964.11659830
16. Docherty GJ. The timing of voicing in British English obstruents. Walter de Gruyter; 1992.
17. Antoniou M, Best CT, Tyler MD, Kroos C. Inter-language interference in VOT production by L2-dominant bilinguals: Asymmetries in phonetic code-switching. Journal of Phonetics [Internet]. 2011 Apr 15;39(4):558–70. Available from: https://doi.org/10.1016/j.wocn.2011.03.001
18. Knowles G. J. C. Wells, Accents of English. Cambridge University Press, 1982. 3 vols. Pp. xx + 673; cassette. Journal of Linguistics [Internet]. 1983 Sep 1;19(2):502–3. Available from: https://doi.org/10.1017/s0022226700007908
19. Sailaja P. Indian English: features and sociolinguistic aspects. Language and Linguistics Compass [Internet]. 2012 May 28;6(6):359–70. Available from: https://doi.org/10.1002/lnc3.342
20. Wiltshire CR. The “Indian English” of Tibeto-Burman language speakers. English World-Wide a Journal of Varieties of English [Internet]. 2005 Oct 31;26(3):275–300. Available from: https://doi.org/10.1075/eww.26.3.03wil
21. Sarmah P, Wiltshire CR. Acoustic properties of the monophthongs of Assamese Indian English speakers. World Englishes [Internet]. 2023 Oct 19;43(1):161–81. Available from: https://doi.org/10.1111/weng.12637
22. Klatt DH. Voice onset time, frication, and aspiration in Word-Initial consonant clusters. Journal of Speech and Hearing Research [Internet]. 1975 Dec 1;18(4):686–706. Available from: https://doi.org/10.1044/jshr.1804.686
23. Krug M. Phonology and language use . By Joan Bybee. (Cambridge studies in linguistics 94.) Cambridge: Cambridge University Press, 2001. Pp. xviii, 238. ISBN 0521533783. $28. Language [Internet]. 2003 Dec 1;79(4):793–4. Available from: https://doi.org/10.1353/lan.2003.0246
24. Lakens D. Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs. Frontiers in Psychology [Internet]. 2013 Jan 1;4:863. Available from: https://doi.org/10.3389/fpsyg.2013.00863
25. Ohala JJ. The origin of sound patterns in vocal tract constraints. In: The production of speech [Internet]. Springer; 1983. p. 189–216. Available from: https://doi.org/10.1007/978-1-4613-8202-7_9
26. Henton CG, Bladon RAW. Breathiness in normal female speech: Inefficiency versus desirability. Language & Communication [Internet]. 1985 Jan 1;5(3):221–7. Available from: https://doi.org/10.1016/0271-5309(85)90012-6
27. Simpson AP. Phonetic differences between male and female speech. Language and Linguistics Compass [Internet]. 2009 Mar 1;3(2):621–40. Available from: https://doi.org/10.1111/j.1749-818x.2009.00125.x
28. The Bilingual Language Profile - Bilingual Language profile [Internet]. Bilingual Language Profile. 2026. Available from: https://blp.coerll.utexas.edu/

How to cite this article: Debarshi Das, Anne Varghese. Voice onset time transfer patterns in Assamese-English bilinguals: evidence for asymmetric cross-linguistic influence. *Gal Int J Health Sci Res*. 2026; 11(3): 282-293. *DOI: https://doi.org/10.52403/gijhsr.20260331*

******