

# Mixed language processing increases cross-language phonetic transfer in Bengali-English Bilinguals

Auromita Mitra, Indranil Dutta

## 1 Phonetic transfer

Research on the phonological systems of bi/multilingual speakers suggests that speakers maintain separate, but phonologically linked sound categories for both their languages. These show cross-language influence in both perception and production. In production, this kind of influence (variously termed transfer, drift, accommodation, interference– used interchangeably here to indicate any interaction between two sets of phonetic norms) has been studied in two conditions: (i) While a bilingual speaker is operating in any one of their languages. These studies often compare the speech of bilinguals to monolingual norms, and view transfer as changes to long-term memory representations as a result of acquiring an L2 (? ,? ,? ,? ); (ii) When both languages of a bilingual speaker are co-activated. These compare productions during mixed-language processing to participants’ own productions while using a single language. The resulting transfer is variously thought to involve short-term memory, online processing costs, language mode, and context-awareness. This is the focus of the present study.

Studies of the phonetic effects of mixed-language production are relatively fewer, and much more recent. Existing work has largely focused on a few pairs of phonologically related languages, and used temporal properties of consonants (in particular, VOT) to measure transfer. However, great variation in results (cf sec.?? ) suggests that data from a wider variety of populations, language pairs, and phonetic features is required in order to make meaningful generalizations. There is no work yet on phonetic transfer in the Indian subcontinent, or in any Indo-Aryan language. Nevertheless, widespread multilingualism in this part of the world suggests that phonetic behavior in these populations can improve our understanding of cross-language interactions at the phonetic level, as they are more likely to reflect real-world experience with mixed-language processing.

The present study looks at phonetic transfer between Bengali and English in a group of highly proficient bilingual speakers in India. We measure the spectral properties (F1 and F2) of two English vowels to check if L1 influence on L2 increases during mixed-language use relative to a participant’s unilingual baseline production. Mixed-language data is elicited in two switching paradigms– code-switching and cued picture-naming. The results show an effect of dynamic transfer on L2 vowels for the first time, and a difference in the extent of transfer between the two paradigms. The findings are discussed in light of recent proposals about asymmetries in transfer effects, and debates about the role of connected speech in enhancing cross-language interaction.

The next section gives background information on Bengali and English in India, focusing on vowel systems. The rest of the introduction discusses causes, duration, and possible sources of asymmetries in transfer reported in previous studies, paradigms used, and the design of the present study in light of these.

## 1.1 Bengali and English in India

### 1.1.1 Demography

Bengali (also, Bangla) is an Indo-Aryan language spoken in Indian and Bangladesh. In India, more than 97 million people speak Bengali as a first language (Census of India, 2011), mostly in the state of West Bengal. A majority of this population also speaks other additional languages.

Indian English (IE) refers to the variety of English that has developed in the Indian subcontinent. In India, it is spoken as an L2 by 129 million people (Census of India, 2011). English is one of the two official languages, used in education, law, media, as a lingua franca mainly for an educated elite in most metropolitan regions, and carries a high prestige value (?). In spite of the different L1s of its speakers, recent literature suggests that IE has a target phonology that is distinct from any of these, as well as from other native varieties of English (see ? (?) for a review). Thus, regional variations are seen as the result of L1-influence on a common underlying target. Early attempts to describe IE phonologically (General Indian English or GIE, (?,?)) were primarily attempts at standardization for pedagogical purposes. Excepting a small minority of specific communities (?), there are hardly any L1 speakers of IE.

Considering these facts, the present study focuses on short-term phonetic interaction during mixed-language use, because:

- (i) Since the population is multilingual, we expect long-term representations to be affected by multiple languages.
- (ii) Given that examples of an IE phonology without L1 ‘influence’ are very rare, it is more meaningful to think of cross-language transfer in L2 as relative to a speaker’s own production in a given ‘baseline’ condition.

### 1.1.2 Vowel systems

The vowel inventory of Western Bangla (the variety spoken by the participants of the present study) consists of [i, e, æ, a, ɔ, o, u], and their nasalized counterparts (?). Note that there is no mid-central vowel. Of relevance to this study, the category [æ] in Bengali is lower than the corresponding English category (?).

The vowel system of IE contains the monophthongs [ɪ, i, ɛ, e, æ, ə/ʌ, ʌ, ɔ, o, ʊ, u], represented by the lexical set KIT, FLEECE, DRESS, FACE, and TRAP, STRUT, PALM, LOT/CLOTH, GOAT, FOOT, GOOSE (?). Their distribution in the F1XF2 space is shown in (insert table 1.7, Pandey). There is a single mid-central vowel corresponding to the categories [ʌ, ə, ɜ:], which are treated as distinct in most native varieties of English. Since the English items used in this study are traditionally transcribed with [ʌ], we use this symbol to indicate the mid-central vowel throughout.

## 1.2 What causes transfer and what does it affect?

Many studies have distinguished between changes in memory REPRESENTATIONS due to the acquisition of multiple sound systems (cf SLM ?; PAM-L2 ?), and interaction during PROCESSING (transfer vs interference; ?, competence vs performance interference; ?). Given their transient nature (cf sec.??), dynamic changes in production during mixed-language use are generally thought to involve the latter (?). What triggers this interaction? ? (?) argues that while cross-language phonetic effects are largely measured at the point of the switch, it could have two potential sources:

1. The local point of switch itself
2. Global co-activation of two languages – bilingual language mode (?,?)

A number of studies have specifically manipulated language mode, both in the presence and absence of switching. Overall, results suggest that (i) In the absence of other manipulations, productions in a bilingual language mode show increased cross-language influence compared to a monolingual mode (?,?,?); (ii) However, language mode is not the sole source of influence during mixed language use – studies comparing switched and non-switched tokens produced in the same test block (identical language mode) (?,?,?) or spontaneous conversation (?,?), have still reported a difference, suggesting that independently of mode, switching between languages triggers a local increase in cross-language transfer. (iii) How the two sources interact to influence the final outcome of transfer is not fully understood– (?,?) found no additive effects, (?,?) found the balanced language context to inhibit transfer compared to unbalanced contexts. Other studies have not analyzed the two separately, eliciting switched tokens in a bilingual test block and non-switched tokens in separate monolingual test block, separated by a few hours to days (?,?,?,?,?,?,?).

? (?) suggests that various aspects of the communicative setting, including exposure to (spoken or written) stimuli in multiple languages, and awareness of the interlocutor's being bilingual, could trigger a bilingual mode. Since the participants in this study are in an environment which largely contains mixed-language input, multilingual interlocutors, and no stable ambient language, we expect that phonetic transfer during daily language use, if any, takes place in a bilingual mode. Thus to preserve ecological validity, we elicited both unilingual and switched utterances in a bilingual language mode (cf section Methodology). Any observed differences in this paradigm would result from interaction during online processing.

In a bilingual mode, both language systems are expected to be (nearly) equally accessible throughout the test block. Thus, a consistent difference between switched and non-switched tokens in such a paradigm would be possible only if the effects were highly localized (if not, we should expect a gradual convergence over the course of the experiment). This is discussed in the next section.

### 1.3 Duration: How long do effects last?

Studies which compare bilinguals' phonological systems with monolinguals see them as relatively stable over time. The difference from monolingual norms is interpreted as the cumulative result of cross-language influence over long periods, (e.g. ?,?,?). However, longitudinal studies (citations) and between-subject comparisons of bilingual speakers with different durations of L2 exposure suggest that the interaction between the sound systems is dynamic. Over time, increasing exposure to an L2 can reduce effects of L1 influence, leading to a more fine-grained separation and native-like production of foreign contrasts– e.g. ? (?)

Moreover, changes due to transfer are not unidirectional or irreversible. ? (?) first demonstrated that spending 2-5 months in an L1 or L2 environment causes productions in both languages to “drift” towards the ambient language. This not only showed that cross-language interaction can be triggered in an order of months, but also that the effects can be reversed within a similar time range. A recent study by ? (?) found comparable effects in an even shorter duration (2-4 weeks). A short-term longitudinal study by ? (?) showed that over the course of the first five weeks of learning an L2, there was a gradual convergence of L1 towards L2. For some sounds and features (though not others, cf sec.??), transfer was additive over time. How long these effects last in the absence of regular L2 input was not tested. (previous two paragraphs needed?)

Studies looking at mixed language use within an utterance mostly analyze only the switched (target) token. Thus, very few direct measurements of the duration of transfer effects are available. One study which measured this (?,?) did not find any residual effects on the matrix language following the switch, suggesting that transfer during code-switching is localized.

More indirect evidence for this comes from studies which have elicited switched and non-switched tokens in the same test block (?,?,?), and still found differences between the two, suggesting that changes due to transfer are quickly ‘reset’ – in an order of seconds. However, note that these studies measure VOT, which is a temporal feature. We have no apriori reason to assume that these durations generalize to vowel quality. However, findings from sub-categorical phonetic shifts triggered by other factors (such as convergence towards an interlocutor) do evidence rapid shifts in vowel quality within similar time-frames (?,?,?,?). Thus in the current study, we present switched and non-switched tokens in the same test block to induce a bilingual mode. We expect the intervening words between two subsequent targets to undo any residual effects of transfer.

## 1.4 Asymmetries between languages in extent and direction of transfer

Previous research has largely reported asymmetrical shifts between languages. Two sources of such differences are discussed below:

### 1.4.1 L1 vs L2 status

Flege’s Speech Learning Model (SLM) (?,?,) posits that the sound categories of a bilingual speaker exist in a common phonological space, and thus in principle, both L1 and L2 sounds can affect one another. Thus, the patterns of cross-language transfer depend on the mapping of L2 phonemes in relation to existing L1 categories. In non-switched production, this is evidenced through pervasive L1 influence on contrasts which are perceptually linked to an existing L1 contrast (equivalence classification– ?,?,?), and the observation that both L1 and L2 sound systems of bilinguals differ from corresponding monolingual speakers (?,?,).

In mixed-language production, the role of language status is less clear– studies have reported unidirectional influence of L1 on L2 (?,?,?,?,?), L2 on L1 (?,?,?,?), bidirectional convergence (?,?,?), divergence (?,?,?), and no influence (?,?,?). Since most existing studies measure shifts in VOT, the observed asymmetries between languages have been variously explained either in terms of their L1 vs L2 status, or language-specific differences between long- and short-lag VOT languages (cf sec. ??).

? (?) first found unidirectional VOT shift of L1 towards L2 in two different groups – native speakers of Spanish (short-lag) and English (long-lag), matched for proficiency and age of L2 acquisition. This suggests that beyond language-specific differences, the L1 vs L2 status of the language does mediate transfer. They explain the asymmetry in terms of the Inhibitory Control Model (ICM) (?,?,) – to select a phonetic realization from one language, the other must be inhibited. Being the ‘stronger’ language, more inhibition is required on the L1 while speaking in an L2, than vice-versa. Thus, switching into L1 incurs a greater switch-cost (more cross-language influence) than switching into L2. The lesser switch cost results in a lack of visible cross-language transfer effects on L2. ? (?) found comparable results, but only in the groups which were non-balanced in language dominance.

Since VOT is independently affected by short- vs long-lag differences, studies of other sound contrasts/features can avoid this conflation. However, given the implications of the ICM, we would first have to establish that these L2 categories can indeed be affected by dynamic interference. The few existing studies which have measured vowel quality (?,?,?,?) and phonological processes (?,?,?) have examined transfer effects on L1. Thus, this study builds on the existing work by checking if the L2 vowel quality of proficient bilinguals is affected by mixed-language processing.

### 1.4.2 Sound systems of the languages:

In contrast to the cognitive factors discussed above, many studies have attributed asymmetries to language-internal factors. Specifically, a pattern that emerges from the body of work on VOT is that shift in long-lag VOTs towards the short-lag norms of the other language is much more consistent and systematic than the reverse (?,?,?,?/?). In general, short-lag VOTs seem to resist accommodative shift. ? (?) suggests that because long-lag languages offer a greater range of acceptable VOTs, there is more ‘room’ for movement. In contrast, shift in short-lag languages would risk the loss of a phonological contrast (VOT is the primary cue for voicing contrast in most of these languages). This is a linguistic/phonological constraint of transfer. In a similar vein, ? (?) found that while English stops shifted towards Greek (a short-lag language), the reverse was seen only in a subset of sounds— word-medial non-nasalized stops. They attribute this to the low frequency of these sounds in Greek, which makes them less ‘stable’, and thus more amenable to shift. This suggests that the distribution of sounds in a language can affect the kind of transfer patterns observed.

These results highlight that beyond interaction in cognitive processing, phonetic transfer is ultimately a linguistic phenomenon, and thus subject to language-specific constraints. Therefore, to make meaningful generalizations, it is imperative to consider data from a variety of language pairs. There are very few studies on phonologically(?) unrelated languages. This is the first transfer study on an Indo-Aryan language.

Since Bengali has a 4-way laryngeal contrast, VOT does not just encode a voiced/voiceless distinction, and moreover is not the only cue for the corresponding phonological categories (?,?). Thus, effects of Bengali on English VOT are likely to be affected by additional phonological factors. To avoid any potential confounds, we chose to look at vowels as a target for transfer.

### 1.5 Asymmetries between sounds:

Existing research reports asymmetrical transfer effects not only between languages, but also among different sounds/features in a language, in both long-term and transient interactions. Studies which have examined multiple sound categories have found that interactions between individual sounds pairs do not necessarily reflect the overall pattern of global (system-wide) shift (?,?), suggesting that ‘extent of transfer’ cannot be seen as an atomic measure. In light of the discussion in section ??, this is not surprising – if a general linguistic principle of ‘room for movement’ constrains transfer, then we would expect it to apply to individual sounds too. Once again, this emphasizes the need to study a wider range of sound contrasts.

In the present study, we examine two vowel contrasts in Indian English– the mid-central vowel [ʌ] and the low front vowel [æ]. As table () shows, the former is in a less crowded part of the vowel space, and thus has more latitude for movement without risking contrast. Thus, considering phonological constraints, we expect a greater degree of shift in [ʌ] compared to [æ].

However, there is another potential source of asymmetry between the two sounds, which could lead to an alternative pattern:

[ʌ] is not present as a phoneme in Bengali, whereas [æ] is a common category across both languages (cf sec.??). Flege’s Speech Learning Model (SLM) (?,?) suggests that sound categories which are common across languages influence each other because they share a common acoustic-phonetic space. Other categories can in principle be produced with native-like accuracy, showing little effect of transfer. Given that our participants are highly proficient bilinguals, and assuming that cross-language correspondences remain the same in representation and processing, this would lead us to expect more transfer in the common category, [æ].

This also seems to follow from the fact that there is no obvious competing L1 category (no specific target for transfer) during the production of [ʌ]. However, it is unlikely that this should altogether preclude a shift

in [ʌ], since at least one existing study has reported transfer effects on a non-common vowel category in a comparable paradigm (?,?). Here, the Catalan [ɔ] moved towards an acoustically close category [o], which is common to Catalan and Spanish, suggesting that shift might be towards a set of phonetic norms rather than a particular target.

Thus, we expect the English [æ] to lower towards the corresponding Bengali category(cf sec.??). The vowel [ʌ] is expected to shift towards an acoustically close category in IE – a probable candidate being the low vowel [a:], as anecdotal observations of heavily-accented English suggest that the sounds are perceived as close by Bengali speakers. The extent of transfer in each vowel could plausibly show either of the patterns outlined above.

## 1.6 A note on paradigms: connected speech or not?

Studies of phonetic transfer during mixed-language use have largely used spontaneous or scripted code-switching (CS, defined as the use of multiple languages within a single utterance, ie in connected speech; ?,?) to co-activate the two languages. Other paradigms include cued picture-naming, delayed repetition, interpreting across languages, reading word-lists etc. ? (?) argues that code-switching introduces discursive factors such as discourse planning and pragmatic considerations, which ultimately show patterns of language practice, rather than cognitive behavior. This finds support from studies which report transfer effects before the switch point (?,?), or other evidence of planning such as hyperarticulation (?,?) and interlocutor-awareness in spontaneous CS (?,?). To avoid such confounds, Olson suggests cued picture-naming (where pictures are named in rapid succession in either language, based on the given language cue) as an alternative paradigm to isolate purely phonetic effects in controlled setups.

Note that while there is evidence for planning in CS, exactly how this affects phonetic transfer cannot be predicted from the existing literature– while (?,?) found tokens immediately preceding the switch point to be phonetically different from the rest, one group showed convergence towards the switch language(explained in terms of proficiency) while the other showed divergence (explained in terms of extra-linguistic factors). The participants in Muldner et al.'s (?) study hyperarticulated switched tokens (suggesting some degree of planning), and showed no significant shift. Moreover, recall that Olson's (?) findings suggest that there is a limit on transfer– multiple factors do not necessarily lead to additive effects. Thus, the existence of additional factors in CS does not automatically mean that this should affect the extent of transfer.

While much more work is necessary to tease out the various factors and their interactions, the focus of the present study is to check whether, in an experimental setting, the postulated discursive elements in a CS paradigm affect the extent of observed transfer, compared to a paradigm which arguably lacks them–cued picture-naming. We are less concerned with the exact nature of differences, than with their comparability. While studies using each of these paradigms to induce transfer exist, comparison across studies is difficult, since they also differ in other factors. This study provides a starting point for such comparison by eliciting productions of the same tokens from the same participants, in both paradigms. A difference in the extent of transfer will suggest that the planning effects in connected speech affect transfer independently of the phonetic results of switching, which has implications that future studies should look into.

## 2 Methodology

Design: 10 participants\*20 target words\*2 utterance types \*2 tasks \*4 iterations = 3200 items

## 2.1 Participants

- 10 Bengali-English bilingual speakers (5 female)
- age range– 19 to 28
- selected on the basis of response to a Language Background Questionnaire, to match for language experience and LSRW skills in L1 and L2

## 2.2 Stimuli

- Target words: 20 monosyllabic English words containing [ʌ] or [æ] (not lexicalized loanwords)
- Words with pre-vocalic voiced plosives (/b/ and /d/) were selected, to minimize coarticulatory effects on the target vowel
- Filler words: 10 monosyllabic English words not containing the target vowels

## 2.3 Procedure

1. Cued picture-naming: Participants were presented with slides in the following sequence-
  - Language cue- a word in either English or Bengali orthography
  - a picture displayed for 50ms; named by the participant in either English or Bengali, depending on the language cue
  - Target word- English word printed in English orthography; read out by the participant
  - Distracter math problem

Since all target words were in English, language cue was used to manipulate the preceding word, giving target words in non-switched and switched utterances. English and Bengali orthographies are visually distinct, so it was used to cue language.

2. Code-switching: Each slide contained one target word embedded in either an English (unilingual) or a Bengali (bilingual; target produced as a code switch from Bengali) carrier sentence; read out by the participant. Carrier sentences:
  - Unilingual sentence: That is a yellow [Target Word].
  - Mixed-language sentence: o-ta æk-ta ka:lo [Target Word].

The sound preceding the target word was uniform across all sentences (the final sound in ‘yellow’ is assimilated to [o] in Indian English), to minimize the possible effects of coarticulation on the target word. A specifier-noun construction was used to ensure identical word order in both languages, giving uniform prosodic context for the target word. For bilingual utterances, mixed orthography was used. Four iterations of each task were recorded. Every target word appeared twice during a test block; once in a switched and once in a non-switched context. Order of items and conditions (switched vs non-switched) was randomized within each trial. Participants alternated between the two tasks.

## 2.4 Acoustic analysis

The first iteration of each task was treated as a trial. An additional 67 tokens (2.79%) were excluded due to errors in presentation/production, giving 2333 target word items for analysis. Recodings were manually segmented and annotated in PRAAT (?), F1 and F2 of the target vowels [ʌ, æ] were measured at 5% and 50% into the vowels.

### 3 Results

- A linear mixed effects model with utterance type (unilingual vs bilingual), vowel, and gender as fixed effects shows a significant effect of utterance type on lobanov normalized F1 at 5% ( $p=0.0096$ ). Normalized F1 is higher in the bilingual utterances.
- No effect of utterance type on F1 at 50% of the vowel, or on normalized F2 values.
- [ʌ] is lowered to a greater extent than [æ].
- Significant effect of task
- /d/-initial target words show more lowering than /b/-initial words

### 4 Discussion

Results are in line with previous studies which find that CS between languages does not lead to a parallel switch in phonology. Instead, gradient, momentary phonetic shifts are observed.

- Both the vowels are lowered (shift towards related L1 category) in the bilingual utterances compared to speakers' own unilingual productions, showing increased cross-language influence during mixed-language processing.
- Lowering at 5% of the vowel, but not at 50%, suggests that only the initial part of the vowel is receptive to dynamic shift, not the steady state
- No change in F2
- More shift in [ʌ] compared to [æ]– suggests that the extent of shift is constrained by the distribution of vowels and room for variability. Since we don't have monolingual Bengali data, can't say anything about the extent of shift relative to the initial distance between the categories.
- CS in connected speech leads to greater cross-language phonetic interaction than unplanned switches through a cued picture naming task, suggesting that at least part of the observed effects in CS paradigms can be attributed to discourse awareness and planning.

### 5 Conclusion

Mixed language processing temporarily increases cross-language phonetic transfer, causing L2 vowels to converge towards L1. The extent of shift is mediated by language-specific distribution of sound categories, and is greater in connected speech.