

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

Word count: 6944

Abstract Existing research demonstrates that mixed-language productions of bilingual speakers differ from productions in a single language. Studies have largely focused on temporal properties of consonants, and reported varied, and often contradictory, results. The present study examines the effect of mixed-language processing on two L2 vowel categories in a group of highly proficient Bengali-English bilingual speakers in India. Following from a discussion in (Olson 2013), we compare productions elicited in two paradigms– cued picture-naming and code-switching. Results indicate a dynamic increase in cross-language phonetic transfer during mixed-language use leading to a shift towards L1 norms, asymmetrical shifts in the two categories, and no difference in transfer patterns between the two paradigms.

Keywords:

1 Phonetic transfer

Bi/multilingual speakers can distinguish between the phonetic norms of their languages Caramazza et al. (1973); MacLeod and Stoel-Gammon (2010); Bosch and Sebastián-Gallés (2003) and therefore maintain separate sound categories for each language. However, these categories are not autonomous, and show cross-language influence in both perception and production Flege (1995); Fowler et al. (2008); Flege et al. (2002). 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 (Guion 2003; Caramazza et al. 1973; Flege 1987); (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. The present study is concerned with the latter.

Existing research on the phonetic effects of mixed-language production has largely focused on a few pairs of phonologically related languages, and used temporal properties of consonants (in particular, voice onset time or VOT) to measure transfer. However, great variation in results (elaborated in sec.1.4) 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 short-term 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

be particularly valuable to understanding the nature of such cross-language interactions, as they are more likely to reflect real-world experience with mixed-language processing.

The present study examines phonetic transfer between Bengali and English in a group of highly proficient bilingual speakers in India. We measure spectral properties (F1 and F2) of two English vowels to ascertain if L1 influence on L2 increases during mixed-language use, relative to a participant's unilingual baseline production of L2. Mixed-language data is elicited in two switching paradigms: code-switching and cued picture-naming. The results demonstrate a dynamic effect of transfer on L2 vowels, but no difference in the extent of transfer between the two paradigms. Findings are discussed in light of recent proposals about asymmetries in transfer effects and the role of connected speech in introducing discursive factors to transfer studies.

The rest of this section is structured as follows: 1.1 provides background information on Bengali and Indian English focusing on vowel systems, and suggests on the basis of socio-linguistic facts that for the present population, short-term phonetic interaction during mixed-language use is an ecologically valid paradigm that is likely to be a part of everyday language-use experience. 1.2 discusses two potential sources of transfer during language processing— global co-activation of multiple languages vs the local act of switching between languages. Given the highly multilingual setting that characterizes this population, we note that the former is expected to be a constant feature and therefore any observed phonetic interaction is better understood as resulting from the latter. The design of the present study is motivated on the basis of this discussion. A logical consequence of this model is that such phonetic interaction must be highly localized. 1.3 elaborates on this by reviewing literature on the duration and ‘reversibility’ of transfer effects. A recurring pattern that emerges from existing research on short-term phonetic transfer is that the phonetic influence does not affect both languages of the bilingual speaker equally, although the nature of these reported differences varies greatly across studies. In 1.4, we review some of these findings, focusing on proposed sources of asymmetries, in order to highlight the factors that seem to mediate such phonetic interaction. We argue that these results point to an urgent need for expanding the range of language pairs and sound categories examined in order to make meaningful generalizations about the nature of short-term phonetic transfer. Following from this, 1.5 discusses reported asymmetries between different sound categories of a single language, and develops the hypotheses of the present study on the basis of these. In 1.6, we motivate a secondary aim of this study: to verify the postulated differences between two paradigms for eliciting short-term phonetic interaction in an experimental setup, namely code-switching and cued picture-naming. 1.7 summarizes the research questions and hypotheses.

1.1 Bengali and English in India

1.1.1 Demography

Bengali (also, Bangla) is an Indo-Aryan language primarily spoken in India 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 sub-continent. 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 pres-

tige value (Tollefson and Tsui 2014). 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 Sirsa and Redford (2013) for a review). Thus, regional variations are seen as the result of L1-influence on a common underlying target. Excepting a small minority of specific communities (Pandey 2015), there are very few L1 speakers of IE.

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

- (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 in this study) consists of [i, e, æ, a, ɔ, o, u], and their nasalized counterparts (Rubino and Garry 2001). Note that there is no mid-central vowel category.

The vowel system of IE contains the monophthongs [ɪ, i, ɛ, e, æ, ə/ʌ, aɪ, ɔ, o, ʊ, u], represented by the lexical set KIT, FLEECE, DRESS, FACE, and TRAP, STRUT, PALM, LOT/CLOTH, GOAT, FOOT, GOOSE (Wells and Wells 1982; Masica 1972). There is a single mid-central vowel corresponding to the categories [ʌ, ə, ɜ:], which are treated as distinct in many 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?

Existing research has distinguished changes in category REPRESENTATIONS due to the acquisition of multiple sound systems (cf Speech Learning Model (Flege 1995; 2007); Perceptual Assimilation Model-L2 (Best and Tyler 2007)), from interaction during processing (transfer vs interference (Grosjean 2012); competence vs performance interference (Paradis 1993)). Given their transient nature (cf sec.1.3), dynamic changes in production during mixed-language use are generally thought to involve the latter (Elias et al. 2017; Simonet 2014). What triggers this interaction? Olson (2016) argues that while cross-language phonetic effects are largely measured at the point of language switch, it could have two potential sources:

- i. The local point of switch itself
- ii. Global co-activation of two languages – bilingual language mode (Grosjean 1998)

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 (Simonet and Amengual 2020; Simonet 2014); (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) (Olson 2016; Tsui et al. 2019) or spontaneous conversation (Piccinini and Arvaniti 2015), 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—Olson (2016) found no additive effects, Olson (2013) 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 a separate monolingual test block, separated by a few hours to days (Bullock and Toribio (2009); Antoniou et al. (2011); Elias et al. (2017); Šimáčková and Podlipský (2015; 2018)).

Grosjean (1998) 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. Any observed differences in this paradigm would arguably 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 speakers' 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, Guion (2003); Caramazza et al. (1973). However, longitudinal studies and between-subject comparisons of bilingual speakers who differ in the duration of L2 exposure suggest that 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. Bohn and Flege (1992) found that an approximately 7 years' difference in L2 exposure between groups having the same age of acquisition led to a significant difference in the amount of cross-language transfer. These have been interpreted as changes to long-term phonological representations.

Moreover, changes due to transfer are not unidirectional or irreversible. Sancier et al. (1997) 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 Tobin et al. (2017) found comparable effects in an even shorter duration (2-4 weeks). A short-term longitudinal study by Chang (2012) 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.1.5), transfer was additive over time. How long these effects last in the absence of regular L2 input was not tested.

Studies examining 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 (Bullock and Toribio 2009) did

not find any residual effects on the matrix language following a 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 (Tsui et al. 2019; Olson 2013), and still reported differences between the two, suggesting that changes due to transfer are quickly ‘reset’– in an order of seconds. Note that all 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 comparable time-frames (Pardo 2010; Babel 2010; 2012). Thus in the current study, we present switched and non-switched tokens randomly within 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 proposed sources of such differences that are relevant to this study are discussed below:

1.4.1 L1 vs L2 status

Flege’s Speech Learning Model (SLM) 1995; 2007 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 a pervasive L1 influence on non-native contrasts that are perceptually linked to an existing L1 contrast (equivalence classification– Flege and Hillenbrand (1984); Flege (1987)), and the observation that both the L1 and L2 sound systems of bilinguals differ from corresponding monolingual speakers (Guion 2003).

In mixed-language production, the role of language status is less clear– studies have reported unidirectional influence of L1 on L2 (Balukas and Koops 2015; Antoniou et al. 2011; Šimáčková and Podlipský 2015; Goldrick et al. 2014), L2 on L1 (Tsui et al. 2019; Elias et al. 2017; Olson 2013), bidirectional convergence (Bullock and Toribio 2009; Olson 2016), divergence (Bullock and Toribio 2009; Šimáčková and Podlipský 2018), and no influence (Muldner et al. 2019; Schwartz et al. 2015). 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. 1.4.2).

Olson (2013) first reported a 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 suggested 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) (Green 1998) – to select a phonetic realization from one language, the other must be inhibited. Being the ‘stronger’ language, more inhibition is required on the L1. 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. Tsui et al. (2019) found comparable results, but only in participants who were not equally dominant in

both languages. Balanced bilinguals did not demonstrate any transfer effects in VOT. They propose that this is because balanced bilinguals have better inhibitory control (low switch costs in both languages) due to greater experience with language switching.

Since VOT is independently affected by short- vs long-lag differences, studies of other sound contrasts/features can avoid this conflation, and give a clearer picture of the effect of language status. 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 focus on vowel quality (Simonet 2014; Muldner et al. 2019; Elias et al. 2017) and phonological processes (Simonet and Amengual 2020; Schwartz et al. 2015) have all examined transfer effects on L1. As far as we are aware, there has been no work yet on vowels in L2. Thus, this study builds on the existing work by examining whether the L2 vowel quality of proficient bilinguals can be affected by mixed-language processing.

1.4.2 Differences between sound systems of the languages

In addition to the cognitive factors discussed above, many studies have attributed the asymmetries to language-internal factors. Specifically, a pattern that emerges from the body of work on VOT is that a shift in long-lag VOT towards the short-lag norms of the other language is much more consistent and systematic than the converse Tobin et al. (2017); Olson (2016); Bullock and Toribio (2009); Antoniou et al. (2011); Chang (2012). In general, short-lag VOTs seem to resist accommodative shift. Bullock and Toribio (2009) suggest 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 could risk the loss of a phonological contrast (VOT being the primary cue for voicing contrast in these languages). This is a language-specific phonological constraint on transfer. In a similar vein, Antoniou et al. (2011) 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 could affect the kind of transfer patterns observed.

These results highlight that beyond interaction during the cognitive processing of sounds, phonetic transfer is ultimately a linguistic phenomenon, and thus subject to language-specific phonological constraints. Therefore, to make meaningful generalizations, it is imperative to consider data from a variety of language pairs. Existing research on phonetic transfer largely revolves around a small set of phonologically related languages. Thus, the present study extends the scope of this research to a new set of languages— Indian English and Bengali (an Indo-Aryan language).

Since Bengali has a four-way laryngeal contrast, VOT does not encode a binary voiced—voiceless distinction, and moreover is not the only cue for the corresponding phonological categories (Dmitrieva and Dutta 2020). Thus, the effects of Bengali on English VOT are likely to be affected by additional phonological factors. Thus, avoiding any potential confounds was another reason we chose to focus on vowels as a target for transfer.

1.5 Asymmetries between sounds

Existing research reports asymmetrical transfer effects not only between languages, but also between 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 (Chang 2012; Elias et al. 2017), suggesting that ‘extent of transfer’ cannot be seen as an atomic measure. In light of the discussion in section 1.4.2, this is not surprising – if a general linguistic principle of ‘room for movement’ constrains transfer, then we should 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 categories in Indian English– the mid-central vowel [ʌ] and the low front vowel [æ]. There are two plausible sources of asymmetry between these:

- i. Position in the vowel space of IE: [ʌ] is in a less crowded part of the vowel space, and thus has more latitude for movement without risking contrast, particularly in the height (F1) dimension. Thus, considering purely phonological constraints on IE, we would expect that a greater degree of shift is possible in [ʌ] compared to [æ]. Any movement in [æ] is expected to be primarily in the front-back (F2) dimension.
- ii. Target of transfer: Since we do not elicit unilingual L1 utterances in this study, we cannot directly compare the target vowels to baseline L1 vowels. However, the body of existing research on phonetic transfer suggests that shifts are targeted with respect to categories in the other language, rather than random. Thus, another source of asymmetry is the fact that [ʌ] is not present as a phoneme in Bengali, whereas [æ] is a common category across both languages (cf sec.1.1.2).

Flege’s Speech Learning Model (SLM) (Flege 1995; 2007) suggests that sound categories which are common across languages influence each other because they share a common acoustic-phonetic space. Thus, if Bengali and English differ in their canonical realizations of [æ], then we should expect the English [æ] to shift towards the corresponding Bengali category in the mixed condition.

There is no obvious competing L1 category 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 (Simonet 2014). 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.

The vowel [ʌ] is thus 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.

As noted above, we did not elicit unilingual Bengali productions from the participants in this study. To estimate a baseline value for the relevant Bengali vowels, productions in the unilingual English condition were compared to data from five female Bengali speakers from Dutta et al. (2016). Formant values were extracted from syllable-initial productions of [æ] and [a:], at 55% into the vowel. Comparisons of Lobanov-normalized F1 and F2 suggests that the Bengali category [æ] is lower, and [a:] lower and more fronted, compared to the English categories [æ] and [ʌ] respectively.

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; Myers-Scotton 1993) to co-activate the two languages. Other paradigms include cued picture-naming, delayed repetition, interpreting across

languages, reading word-lists etc. [Olson \(2013\)](#) 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 ([Bullock and Toribio 2009](#)), or other evidence of planning such as hyperarticulation ([Muldner et al. 2019](#)) and interlocutor-awareness in spontaneous CS ([Khattab 2013](#)). 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 ([Bullock and Toribio 2009](#)) found tokens immediately preceding the switch point to be phonetically distinct 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 extralinguistic factors). The participants in Muldner et al.’s [2019](#) study hyperarticulated switched tokens (suggesting some degree of planning), and showed no significant shift. Moreover, recall that Olson’s [2016](#) 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 compared to any other paradigm. It is possible that the effects in the studies mentioned above could be attributed to some other aspect of the experimental setup, rather than a result of code-switching in particular.

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 thus 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.

1.7 Research questions and hypotheses

- Are L2 vowel categories of proficient bilingual speakers affected by dynamic phonetic transfer during mixed-language production?
- If so, we expect an asymmetry between the vowels in the extent and direction of shift. Based on comparisons with Bengali vowels from [Dutta et al. \(2016\)](#) we expect [ʌ] to be lowered and fronted in the mixed condition. [æ] is expected to show lowering, but to a lesser extent than [ʌ], as it is in a part of the IE vowel space that has greater vowel density.
- Is there an observable difference in the extent of transfer between the cued picture-naming and the code-switching paradigms?

2 Methodology

The experiment had the following factorial design: 10 participants*20 target words*2 conditions *2 tasks *4 iterations = 3200 target items

2.1 Participants

10 Bengali-English bilingual speakers (male = 5, age range 19 to 28), completed the tasks. All the participants speak other languages in addition to Bengali and English. At the time of recording, they were living in or around the EFL University campus in Hyderabad, which has a linguistically diverse population. Speakers were selected on the basis of responses to a Language Background Questionnaire to ensure comparable LSRW skills in L1 and L2 across participants, and were compensated for their time.

2.2 Stimuli

Twenty monosyllabic English words containing the vowels [ʌ] or [æ] were selected as target words. To minimize coarticulatory effects on the vowel, the pre-vocalic consonants were all voiced plosives ([b] or [d]). 10 monosyllabic English words that do not contain the target vowel served as filler items. English words which are lexicalized loanwords in Bengali were avoided.

2.3 Procedure

Participants were recorded individually in a sound attenuated room. Productions were recorded in Praat ([Boersma and Weenink 2016](#)) at a 44.1K sampling rate. Every participant performed four iterations each of two tasks– cued picture-naming and code-switching. Utterances in the two conditions (unilingual and mixed) were randomly interspersed within each iteration. In the unilingual condition, the target word was preceded by another English word in the utterance, whereas in the mixed condition, the target was produced as a switch from Bengali. Every target word appeared twice during a test block; once in a unilingual and once in a mixed condition. Participants alternated between the two tasks, and were allowed to take a short break after each test block.

The tasks are described below:

- i. 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 – an English word; read out by the participant
 - Distracter math problem

Since all the target words were in English, the language cue was used to manipulate the preceding word, giving target words in non-switched (unilingual) and switched (mixed) utterances. The orthographies of English and Bengali are visually distinct, and were thus used to cue language.

- ii. Code-switching: Each slide contained one target word embedded in either an English (unilingual) or a Bengali (mixed; target produced as a switch from Bengali) carrier sentence which was read out by the participant. The carrier sentences are given below:

Table 1: Fixed factors.

Factor	Levels
Gender	male, female
Vowel	[ʌ], [æ]
Task (corresponding to the two paradigms)	p (Picture: cued picture-naming, s (Sentence: code-switching)
Context (corresponding to the experimental conditions)	e (English: unilingual), b (Bengali: mixed)

- Unilingual sentence: That is a yellow [Target Word].
- Mixed-language sentence: o-ʈa æk-ʈa kalo [Target Word].

Gloss and translation for Bengali sentence:

o-ta Ek-ta kalo [target word]
 that-CL one-CL black [target word]
 that is a black [target word]

The sound preceding the target word was uniform across all sentences (the final sound in ‘yellow’ is produced as [o] in Indian English), to minimize any possible differences due to coarticulatory effects 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. Mixed orthography was used for the mixed-language utterances.

2.4 Acoustic analysis

The first iteration of each task was treated as a trial, and excluded from analysis. An additional 67 tokens (2.79%) were excluded due to errors in presentation/production, giving 2333 target word items for analysis. Recordings were manually segmented and annotated in Praat (Boersma and Weenink 2016), and the first and second formant frequencies (F1 and F2) of the target vowels [ʌ, æ] were measured at five points– in 10% increments starting at 5% into the vowel. This gave measures of F1 and F2 at 5%, 15%, 25%, 35% and 45% into the vowel, allowing us to observe any dynamic changes in the extent of transfer over the course of producing the vowel.

3 Results

All statistical analyses were performed in R (R Core Team 2016). The dependent variables were vowel height (F1) and backness (F2). We fitted a series of linear mixed effects models using the *lme4* package (Bates et al. 2015), to examine the effect of experimental conditions on lobanov normalized F1 and F2 at each of the five points in the vowel (sec.2.4). For all mixed models, the alpha criterion was set at $|t| > 2$. The independent variables or fixed factors, along with their levels, are given in table 1. Subject and Item were treated as random factors.

Since we wanted to model height and backness at different points in the vowel using a single model, allowing us to track if/how transfer changes over time, we arbitrarily chose a point in the vowel (F1 at 25%; henceforth *F1_25*) to identify an optimal model, and then fitted this to all other points.

Vowel quality and gender are expected to systematically affect F1. Thus, we began with a model that had Vowel and Gender as fixed effects, random intercepts for Subject and Item, and random slopes for each within-unit variable that was relevant to our hypotheses—by-subject slopes for Task, Context, and Vowel; by-item slopes for Task and Context. Next, adding Context as a fixed effect significantly improved the model fit ($\chi^2(1) = 5.54, p = 0.01854$), suggesting that the experimental condition affects vowel height. However, since we expected the extent of change in vowel height between the conditions to depend on the identity of the vowel (cf sec.1.4), we added a two-way interaction factor between Vowel and Context. As expected, this was found to significantly improve model fit ($\chi^2(1) = 12.61, p = 0.0003837$). To check if the experimental paradigm affected F1, we added Task as a fixed effect, and this marginally improved the model ($\chi^2(1) = 3.74, p = 0.05302$). Following from the hypothesis in sec.1.6, we expected the extent of shift to differ across the two paradigms. However, no significant improvement was achieved by adding an interaction effect between Task and Context. Subsequent additions of other interactions, and removing each of the fixed effects in turn, did not lead to any further improvement.

Context and Vowel are both within-subject factors. Since our hypothesis concerns their interaction, we modified the random effects structure by adding a by-subject random slope for the interaction term (Brauer and Curtin 2018). This improved the model fit ($\chi^2(5) = 11.43, p = 0.04342$). Thus, we report on the model including Gender, Task, Context, Vowel, and a Context-Vowel interaction as fixed effects, random intercepts for Subject and Item, by-subject random slopes for Task, Context, Vowel and the Context-Vowel interaction, and by-item random slopes for Task and Context.

The following sections discuss the results for vowel height and backness respectively:

3.1 Vowel Height – F1

Table 2 summarizes the results of the fixed effects from the model for F1 at each point. As expected, we observe a significant effect of gender, such that male speakers have a lower mean F1 than female speakers. Vowel is not a significant predictor of F1. This seems surprising, given that the phonemic category is expected to systematically affect vowel quality. However, a boxplot visualization (Figure 1) shows that the participants in this study vary greatly both in the relative heights of the two categories and the extent of height difference between them. Note that these measures are only from the English (unilingual) condition, suggesting that the variability is a feature of the participants' IE vowel categories, independent of the effects of switching. The lack of any independent effect of Context on F1 suggests that using English words in mixed utterances does not lead to a uniform lowering/raising across vowels. Instead, the significant interaction between Context and Vowel indicates a lowering in [ʌ], but not in [æ], in the mixed-language condition. The results also reveal a significant effect of Task, such that the sentence task (code-switching) leads to higher F1 values in both vowels, relative to the cued picture-naming task. This effect is strongest at the beginning of the vowel, and disappears by 45%. This is an unexpected result, since we have no principled reason to expect code-switching to induce an across-board lowering of vowels. As noted earlier, there is no significant interaction of Task with any of the other fixed factors. We used a series of ANOVAs to compare the above model with a null model that differs in the absence of the main factor which concerns our hypothesis—Vowel*Context—at each point in the vowel. We find that the effect is not significant at 5%, and increases over the course of articulating the vowel.

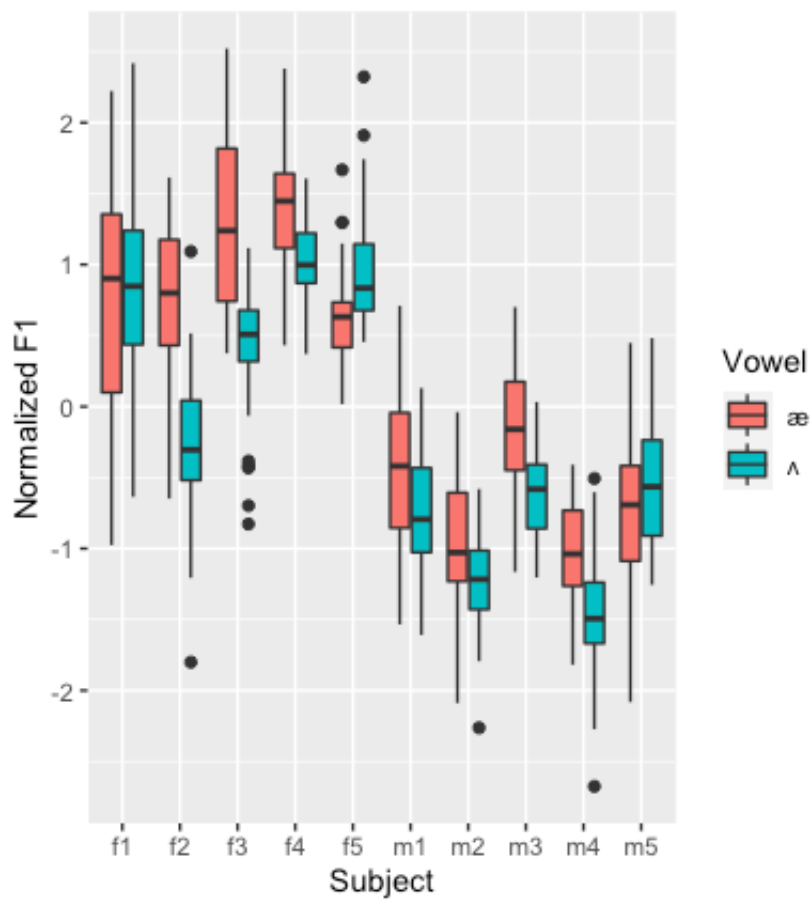


Figure 1: Subject-wise mean F1; unilingual English.

Table 2: Model outputs for F1 at each of the five points in the vowel.

	<i>Dependent variable:</i>				
	F1_5 (Model 1)	F1_15 (Model 2)	F1_25 (Model 3)	F1_35 (Model 4)	F1_45 (Model 5)
(Intercept)	0.447** (0.211)	0.718*** (0.176)	0.750*** (0.153)	0.863*** (0.152)	0.921*** (0.150)
Gendermale	-1.171*** (0.237)	-1.526*** (0.184)	-1.586*** (0.168)	-1.668*** (0.177)	-1.681*** (0.175)
Contexte	-0.010 (0.058)	-0.015 (0.057)	0.033 (0.042)	0.025 (0.040)	0.037 (0.035)
Vowel[Λ]	0.022 (0.170)	-0.070 (0.175)	-0.034 (0.170)	-0.131 (0.174)	-0.216 (0.168)
Tasks	0.319*** (0.091)	0.235*** (0.083)	0.172** (0.081)	0.133* (0.077)	0.102 (0.073)
Contexte:Vowel[Λ]	-0.126** (0.056)	-0.138** (0.060)	-0.185*** (0.055)	-0.191*** (0.064)	-0.192*** (0.066)
Observations	2,333	2,333	2,333	2,333	2,333
Log Likelihood	-1,771.558	-1,589.316	-1,278.310	-1,313.795	-1,380.083
Akaike Inf. Crit.	3,599.116	3,234.631	2,612.621	2,683.589	2,816.165
Bayesian Inf. Crit.	3,760.253	3,395.769	2,773.758	2,844.727	2,977.303

Note:

*p<0.1; **p<0.05; ***p<0.01

3.2 Vowel backness – F2

Table 3 summarizes the results of fixed effects from the model for F2 at each point. Once again, we see that male speakers have a lower mean F2 than female speakers, as expected. The estimates for Vowel show that excepting the beginning of the vowel, the category [ʌ] is consistently produced as more backed than [æ]. Unlike vowel height, Context has an independent effect on vowel backness, such that both the vowels in this study have lower F2 values in the English (unilingual) condition, i.e. they are fronted in the mixed condition. A significant interaction between Vowel and Context shows that there is a difference in the extent of fronting between the two vowels – the shift is greater in the case of [æ]. We see a significant effect of Task on F2, such that vowels in the sentence task (code-switching) were more backed than those in the picture-naming task. As with F1, this is an unexpected result. A lack of significant interaction with Context suggests that there is no difference in the extent of shift between the two paradigms. The results of ANOVAs comparing the full model to a null model which lacks the Vowel*Context interaction factor show that as with vowel height, the change in vowel backness between the two conditions is not significant at 5%, and increases over the course of the vowel.

Table 3: Model outputs for F2 at each of the five points in the vowel.

	<i>Dependent variable:</i>				
	F2_5 (Model 1)	F2_15 (Model 2)	F2_25 (Model 3)	F2_35 (Model 4)	F2_45 (Model 5)
(Intercept)	0.607*** (0.222)	0.838*** (0.213)	0.778*** (0.211)	0.827*** (0.203)	0.826*** (0.204)
Gendermale	−0.839*** (0.124)	−0.891*** (0.106)	−0.807*** (0.107)	−0.750*** (0.116)	−0.748*** (0.109)
Contexte	0.003 (0.036)	−0.088** (0.036)	−0.065** (0.032)	−0.063 (0.044)	−0.103*** (0.033)
Vowel[ʌ]	−0.217 (0.175)	−0.615*** (0.183)	−0.651*** (0.175)	−0.801*** (0.168)	−0.836*** (0.166)
Tasks	−0.234*** (0.029)	−0.194*** (0.035)	−0.152*** (0.050)	−0.141*** (0.047)	−0.092** (0.045)
Contexte:Vowel[ʌ]	0.031 (0.050)	0.110*** (0.041)	0.128*** (0.040)	0.091** (0.046)	0.147*** (0.043)
Observations	2,333	2,333	2,333	2,333	2,333
Log Likelihood	−1,283.531	−1,478.494	−1,174.087	−1,222.682	−1,019.033
Akaike Inf. Crit.	2,623.062	3,012.988	2,404.175	2,501.363	2,094.066
Bayesian Inf. Crit.	2,784.199	3,174.125	2,565.312	2,662.501	2,255.204

Note:

*p<0.1; **p<0.05; ***p<0.01

4 Discussion and Conclusion

With regard to our first research question, the results confirm that the L2 vowels of proficient bilinguals can be affected by mixed-language processing. Both vowels in this study showed systematic shifts in the mixed-language condition compared to baseline unilingual productions. This is significant in the context of an ICM-based analysis of between-language asymmetries in phonetic transfer as proposed by (Olson 2013), because it demonstrates that at least for vowels, the postulated lower switch cost on an L2 does not entirely mask transfer effects. Thus, following from the discussion in sec. 1.4, it might be fruitful for future studies to focus on vowels as a site for transfer. This would both expand the scope of the existing research (cf sec. 1.5) and avoid confounding factors arising from independent differences between long- and short-lag VOT languages.

We hypothesized an asymmetrical shift between the two vowels, owing to differences in (i) their positions in the Indian English vowel space, and (ii) the existence of a corresponding category in the L1 Bengali inventory. The significant interaction between Vowel and Context in both F1 and F2 confirm that the experimental condition affected both height and backness differently for each vowel. As expected, we find a greater degree of shift in [ʌ], possibly owing to its position in a less crowded part of the IE vowel space. Figure 2 shows F1xF2 plots for the vowels in the unilingual and mixed (dotted ellipses) conditions, in comparison to the related Bengali categories (colored ellipses) from Dutta et al. (2016). The category [ʌ] is lowered and fronted in the mixed language condition as expected. The direction of shift in [æ] is surprising– it is fronted in the mixed-language condition. While phonological constraints on IE suggested that a shift in the F1 dimension could potentially risk the loss of contrast and is thus expected to be minimal, Figure 2 shows that [æ] in Bengali is not fronted with respect to the English category. Thus, fronting in the mixed condition cannot be explained in relation the former, at least from the present data. Data from a larger number of speakers, both for IE and Bengali, is needed to explore this further.

Our third research question concerned the experimental paradigms. Olson (2013) proposes that in both scripted and spontaneous code-switching paradigms, “discursive properties may partially drive the production of code-switches, masking underlying effects of interaction”, and thus proposes cued picture-naming as a preferable alternative. This calls to question the comparability of results across studies using these different paradigms. We tested this by comparing productions of the same items from the same group of participants in a scripted CS paradigm and a cued picture-naming paradigm. We found no significant interaction of Task either with Context, or with the Vowel*Context interaction, suggesting that the extent of phonetic transfer was not affected by the paradigm. There are at least two plausible reasons for this– (i) it is possible that the planning effects of CS do not translate to an increase/decrease in observable transfer at all. This would mean that the hyperarticulation seen, for example, in Muldner et al. (2019) is better interpreted as resulting from some other aspect of the experimental setup, rather than the act of code-switching itself. While planning could additionally affect the words preceding the switch, the effects on the switched token itself do not change. (ii) Our study used a frame sentence to induce code switching, and not all the items led to pragmatically meaningful sentences. Thus, another possible explanation is that the discursive effects discussed by Olson (2013) are triggered by more naturalistic paradigms where attention to meaning and interlocutor are important. More within-study comparisons with less structured paradigms, and spontaneous code-switching data, can shed light on this. The results of the present study suggest that for the purposes of observing dynamic phonetic transfer, elicitation through cued-picture-naming vs (controlled) code-switching does not

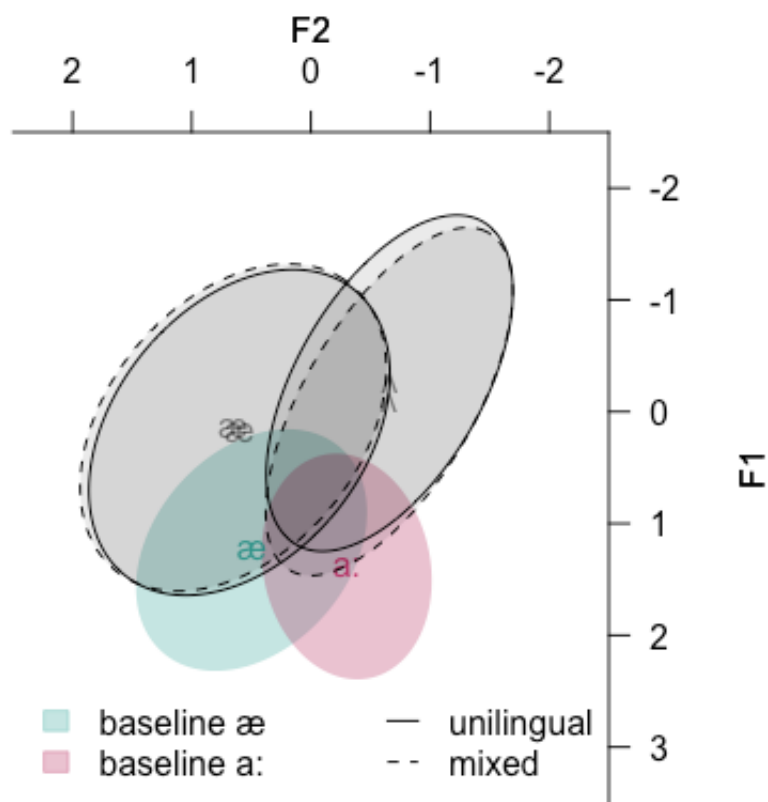


Figure 2: Vowel categories in unilingual English (solid lines), mixed condition (dotted lines), and baseline Bengali (colored ellipses).

seem to significantly alter the behavior under study, and thus the choice of paradigm can be based on other considerations. Moreover, this suggests that (all else being equal) the results from such studies can be compared directly.

Paradigm did affect vowel quality, but this applied uniformly across experimental conditions—all vowels were more backed and lowered in the code-switching task. Given that at least for F1, the effect is limited to the early stages of articulating the vowel, it is plausible that this lowering is an artifact of the particular stimuli used here—resulting from some corarticulatory effect that was absent in the picture-naming task. While this is intriguing, it is beyond the scope of the present study.

Mixed language processing temporarily increases cross-language phonetic transfer, causing an L2 vowel to converge towards a related L1 category. The extent and direction of such shift is mediated by language-specific distribution of sound categories. In a controlled experimental setup to study transfer effects, the postulated discursive factors in a code-switching paradigm does not change the patterns of observed transfer relative to a cued picture-naming paradigm.

References

- Antoniou, M., Best, C. T., Tyler, M. D., and Kroos, C. (2011). Inter-language interference in VOT production by L2-dominant bilinguals: Asymmetries in phonetic code-switching. *Journal of Phonetics*, 39(4):558–570.
- Babel, M. (2010). Dialect divergence and convergence in New Zealand English. *Language in Society*, pages 437–456.
- Babel, M. (2012). Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics*, 40(1):177–189.
- Balukas, C. and Koops, C. (2015). Spanish-english bilingual voice onset time in spontaneous code-switching. *International Journal of Bilingualism*, 19(4):423–443.
- Bates, D., Mächler, M., Bolker, B., and Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, 67(1):1–48.
- Best, C. T. and Tyler, M. (2007). Nonnative and second-language speech perception. In Bohn, O.-S. and Munro, M. J., editors, *Language experience in second language speech learning: In honour of James Emil Flege*, pages 13–34.
- Boersma, P. and Weenink, D. (2016). Praat: Doing phonetics by computer (version 5.4.15)[computer program]. retrieved april 7, 2016.
- Bohn, O.-S. and Flege, J. E. (1992). The production of new and similar vowels by adult german learners of english. *Studies in Second Language Acquisition*, pages 131–158.
- Bosch, L. and Sebastián-Gallés, N. (2003). Simultaneous bilingualism and the perception of a language-specific vowel contrast in the first year of life. *Language and speech*, 46(2-3):217–243.
- Brauer, M. and Curtin, J. J. (2018). Linear mixed-effects models and the analysis of nonindependent data: A unified framework to analyze categorical and continuous independent variables that vary within-subjects and/or within-items. *Psychological Methods*, 23(3):389.
- Bullock, B. E. and Toribio, A. J. (2009). Trying to hit a moving target: On the sociophonetics of code-switching. *Multidisciplinary approaches to code switching*, 41:189–206.
- Caramazza, A., Yeni-Komshian, G. H., Zurif, E. B., and Carbone, E. (1973). The acquisition of a new phonological contrast: The case of stop consonants in french-english bilinguals. *The Journal of the Acoustical Society of America*, 54(2):421–428.

- Chang, C. B. (2012). Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of phonetics*, 40(2):249–268.
- Dmitrieva, O. and Dutta, I. (2020). Acoustic correlates of the four-way laryngeal contrast in marathi. *Phonetica*, 77(3):209–237.
- Dutta, I., Shailendra, I., and Harsha, K. (2016). Using ANNs for vowel identification from V-to-V coarticulation in non-harmonic VCV sequences. In *15th LabPhon. July 13-16th, 2016, Cornell University, Ithaca, NY USA*.
- Elias, V., McKinnon, S., and Milla-Muñoz, Á. (2017). The effects of code-switching and lexical stress on vowel quality and duration of heritage speakers of spanish. *Languages*, 2(4):29.
- Flege, J. E. (1987). The production of “new” and “similar” phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of phonetics*, 15(1):47–65.
- Flege, J. E. (1995). Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research*, 92:233–277.
- Flege, J. E. (2007). Language contact in bilingualism: Phonetic system interactions. *Laboratory phonology*, 9:353–382.
- Flege, J. E. and Hillenbrand, J. (1984). Limits on phonetic accuracy in foreign language speech production. *The Journal of the Acoustical Society of America*, 76(3):708–721.
- Flege, J. E., MacKay, I. R., and Piske, T. (2002). Assessing bilingual dominance. *Applied Psycholinguistics*, 23(4):567–598.
- Fowler, C. A., Sramko, V., Ostry, D. J., Rowland, S. A., and Hallé, P. (2008). Cross language phonetic influences on the speech of french–english bilinguals. *Journal of Phonetics*, 36(4):649–663.
- Goldrick, M., Runnqvist, E., and Costa, A. (2014). Language switching makes pronunciation less nativelike. *Psychological Science*, 25(4):1031–1036.
- Green, D. W. (1998). Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and cognition*, 1(2):67–81.
- Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and cognition*, 1(2):131–149.
- Grosjean, F. (2012). An attempt to isolate, and then differentiate, transfer and interference. *International Journal of Bilingualism*, 16(1):11–21.
- Guion, S. G. (2003). The vowel systems of Quichua-Spanish bilinguals. *Phonetica*, 60(2):98–128.
- Khattab, G. (2013). Phonetic convergence and divergence strategies in English-Arabic bilingual children. *Linguistics*, 51(2):439–472.
- MacLeod, A. A. and Stoel-Gammon, C. (2010). What is the impact of age of second language acquisition on the production of consonants and vowels among childhood bilinguals? *International Journal of Bilingualism*, 14(4):400–421.
- Masica, C. (1972). The Sound System of Indian English. Monograph No. 7.
- Muldner, K., Hoiting, L., Sanger, L., Blumenfeld, L., and Toivonen, I. (2019). The phonetics of code-switched vowels. *International Journal of Bilingualism*, 23(1):37–52.
- Myers-Scotton, C. (1993). *Dueling languages: Grammatical structure in code-switching*. Clarendon Press.
- Olson, D. J. (2013). Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in VOT production. *Journal of Phonetics*, 41(6):407–420.
- Olson, D. J. (2016). The role of code-switching and language context in bilingual phonetic transfer. *Journal of the International Phonetic Association*, 46(3):263–285.

- Pandey, P. (2015). Indian English Pronunciation. In *The handbook of English pronunciation*, pages 301–319. Wiley Online Library.
- Paradis, M. (1993). Linguistic, psycholinguistic, and neurolinguistic aspects of “interference” in bilingual speakers: The activation threshold hypothesis. *International Journal of Psycholinguistics*, 9:133–145.
- Pardo, J. S. (2010). Expressing oneself in conversational interaction. In Morsella, E., editor, *Expressing oneself/expressing one’s self: Communication, cognition, language, and identity*, pages 183–196. London: Taylor & Francis.
- Piccinini, P. and Arvaniti, A. (2015). Voice onset time in Spanish–English spontaneous code-switching. *Journal of Phonetics*, 52:121–137.
- R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Rubino, C. R. G. and Garry, J. (2001). *Facts about the world’s languages : an encyclopedia of the world’s major languages, past and present / edited by Jane Garry and Carl Rubino; contributing linguistics editors, Adams B. Dodomo, Alice Faber, Robert French*. H.W. Wilson New York.
- Sancier, M. L., Fowler, C. A., et al. (1997). Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of phonetics*, 25(4):421–436.
- Schwartz, G., Balas, A., and Rojczyk, A. (2015). Language mode vs. L2 interference: Evidence from L1 Polish. In *ICPhS*.
- Šimáčková, Š. and Podlipský, V. J. (2015). Immediate phonetic interference in code-switching and interpreting. In *Proceedings of the 18th International Congress of Phonetic Sciences*. Glasgow, UK.
- Šimáčková, Š. and Podlipský, V. J. (2018). Patterns of short-term phonetic interference in bilingual speech. *Languages*, 3(3):34.
- Simonet, M. (2014). Phonetic consequences of dynamic cross-linguistic interference in proficient bilinguals. *Journal of Phonetics*, 43:26–37.
- Simonet, M. and Amengual, M. (2020). Increased language co-activation leads to enhanced cross-linguistic phonetic convergence. *International Journal of Bilingualism*, 24(2):208–221.
- Sirsa, H. and Redford, M. A. (2013). The effects of native language on Indian English sounds and timing patterns. *Journal of phonetics*, 41(6):393–406.
- Tobin, S. J., Nam, H., and Fowler, C. A. (2017). Phonetic drift in Spanish-English bilinguals: Experiment and a self-organizing model. *Journal of phonetics*, 65:45–59.
- Tollefson, J. W. and Tsui, A. B. (2014). Language diversity and language policy in educational access and equity. *Review of Research in Education*, 38(1):189–214.
- Tsui, R. K.-Y., Tong, X., and Chan, C. S. K. (2019). Impact of language dominance on phonetic transfer in Cantonese–English bilingual language switching. *Applied Psycholinguistics*, 40(1):29–58.
- Wells, J. C. and Wells, J. C. (1982). *Accents of English: Volume 1*, volume 1. Cambridge University Press.