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.

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 (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 (cf 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 add to the understanding 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

1

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 next section provides 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 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 prestige value (Tollefson & 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 & 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, 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 in this study) consists of [i, e, æ, a, ɔ, o, u], and their nasalized counterparts (Garry et al. 2001). Note that there is no mid-central vowel category.

The vowel system of IE contains the monopthongs [ɪ, i, ɛ, e, æ, ə/ʌ, aː, ɔ, o, ʊ, u], represented by the lexical set KIT, FLEECE, DRESS, FACE, and TRAP, STRUT, PALM, LOT/CLOTH, GOAT, FOOT, GOOSE (Wells & Wells 1982; Masica 1972). There is a single mid-central vowel corresponding to the categories [ʌ,ə,ɜː], 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 between changes in category REPRESENTATIONS due to the acquisition of multiple sound systems (cf SLM (Flege 1995; 2007); PAM-L2 (Best & Tyler 2007)), and 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 & 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 & 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 & Toribio (2009); Antoniou et al. (2011); Elias et al. (2017); Šimáčková & 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 (cf section Methodology). 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 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 & 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 & 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 & Koops 2015; Antoniou et al. 2011; Šimáčková & Podlipskỳ 2015; Goldrick et al. 2014), L2 on L1 (Tsui et al. 2019; Elias et al. 2017; Olson 2013), bidirectional convergence (Bullock & Toribio 2009; Olson 2016), divergence (Bullock & Toribio 2009; Šimáčková & Podlipskỳ 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 & 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 & Toribio (2009); Antoniou et al. (2011); Chang (2012). In general, short-lag VOTs seem to resist accommodative shift. Bullock & 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 & 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 [aː], 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 [aː], at 55% into the vowel. Comparisons of Lobanov-normalized F1 and F2 suggests that the Bengali category [æ] is lower, and [aː] 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 & 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 & 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 extra-linguistic 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 (sec. **??**) 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 [ae] 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 & 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:
   - Unilingual sentence: That is a yellow [Target Word].
   - Mixed-language sentence: o-ta æk-ta kaːlo [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 & Weenink 2016), and the first and second formant frequencies (F1 and F2) of the target vowels [ʌ, ae] 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. 2015b), 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.

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
Context (corresponding to the experimental conditions)	e (English: unilingual), b (Bengali: mixed)

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 & 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 (Figure1) 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.

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

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

	F1_5	F1_15	F1_25	F1_35	F1_45
	Dependent variable:				
	(Model 1)	(Model 2)	(Model 3)	(Model 4)	(Model 5)
(Intercept)	0.447**	0.718***	0.750***	0.863***	0.921***
	(0.211)	(0.176)	(0.153)	(0.152)	(0.150)
Gendermale	−1.171***	−1.526***	−1.586***	−1.668***	−1.681***
	(0.237)	(0.184)	(0.168)	(0.177)	(0.175)
Contexte	−0.010	−0.015	0.033	0.025	0.037
	(0.058)	(0.057)	(0.042)	(0.040)	(0.035)
Vowel\vt	0.022	−0.070	−0.034	−0.131	−0.216
	(0.170)	(0.175)	(0.170)	(0.174)	(0.168)
Tasks	0.319***	0.235***	0.172**	0.133*	0.102
	(0.091)	(0.083)	(0.081)	(0.077)	(0.073)
Contexte:Vowel\vt	−0.126**	−0.138**	−0.185***	−0.191***	−0.192***
	(0.056)	(0.060)	(0.055)	(0.064)	(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: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

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.

	Dependent variable:				
	F2_5	F2_15	F2_25	F2_35	F2_45
	(Model 1)	(Model 2)	(Model 3)	(Model 4)	(Model 5)
(Intercept)	0.607***	0.838***	0.778***	0.827***	0.826***
	(0.222)	(0.213)	(0.211)	(0.203)	(0.204)
Gendermale	−0.839***	−0.891***	−0.807***	−0.750***	−0.748***
	(0.124)	(0.106)	(0.107)	(0.116)	(0.109)
Contexte	0.003	−0.088**	−0.065**	−0.063	−0.103***
	(0.036)	(0.036)	(0.032)	(0.044)	(0.033)
Vowel\vt	−0.217	−0.615***	−0.651***	−0.801***	−0.836***
	(0.175)	(0.183)	(0.175)	(0.168)	(0.166)
Tasks	−0.234***	−0.194***	−0.152***	−0.141***	−0.092**
	(0.029)	(0.035)	(0.050)	(0.047)	(0.045)
Contexte:Vowel\vt	0.031	0.110***	0.128***	0.091**	0.147***
	(0.050)	(0.041)	(0.040)	(0.046)	(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
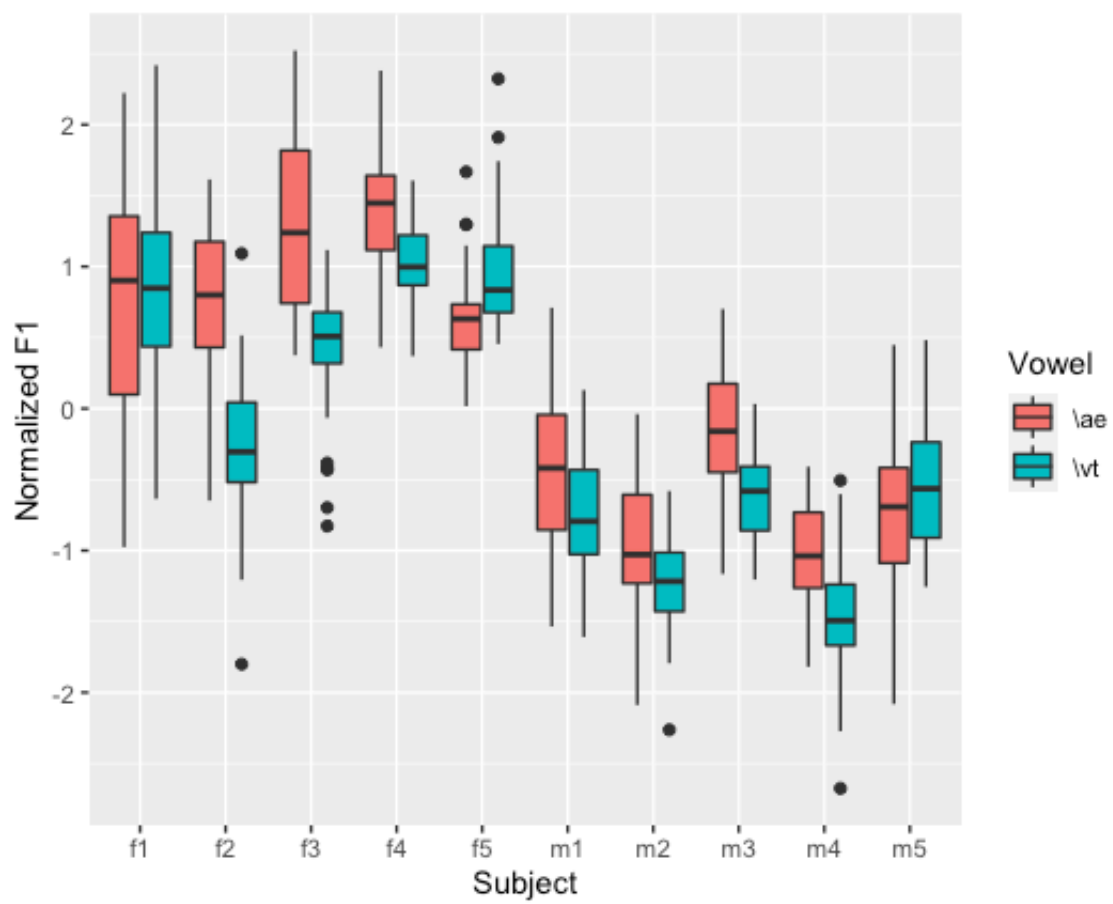
Figure 1: Subject-wise mean F1; unilingual English.
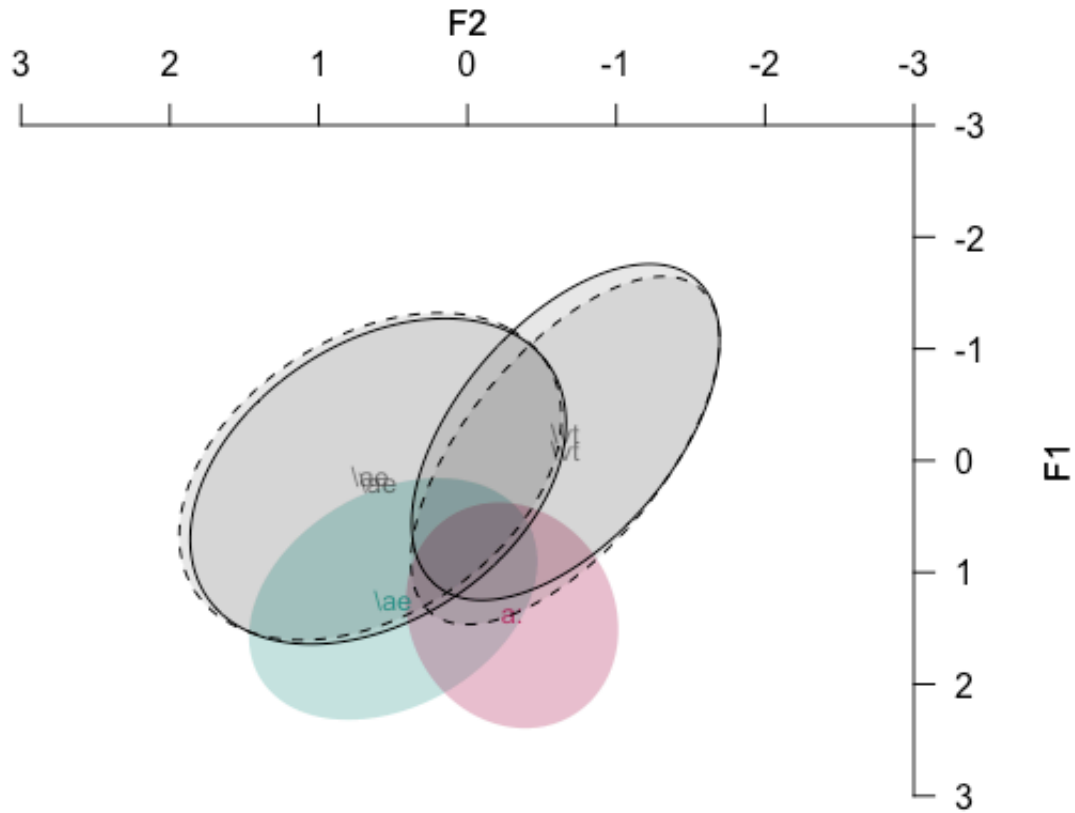
Figure 2: Vowel categories in unilinual English (solid lines), mixed condition (dotted lines), and baseline Bengali (colored ellipses).

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 prop-

erties 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 between 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 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 coarticulatory 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

1989. *Oxford English dictionary*. Oxford: Oxford University Press.

Antoniou, Mark, Catherine T Best, Michael D Tyler & Christian Kroos. 2011. Interlanguage interference in vot production by l2-dominant bilinguals: Asymmetries in phonetic code-switching. *Journal of Phonetics* 39(4). 558–570.

Babel, Molly. 2010. Dialect divergence and convergence in new zealand english. *Language in Society* 437–456.

Babel, Molly. 2012. Evidence for phonetic and social selectivity in spontaneous phonetic imitation. *Journal of Phonetics* 40(1). 177–189.

Babel, Molly & Dasha Bulatov. 2012. The role of fundamental frequency in phonetic accommodation. *Language and speech* 55(2). 231–248.

Babel, Molly, Brianne Senior & Sophie Bishop. 2019. Do social preferences matter in lexical retuning? *Laboratory Phonology: Journal of the Association for Laboratory Phonology* 10(1).

Baker, Mark, Kyle Johnson & Ian Roberts. 1989. Passive arguments raised. *Linguistic Inquiry* 20. 219–251.

Balukas, Colleen & Christian Koops. 2015. Spanish-english bilingual voice onset time in spontaneous code-switching. *International Journal of Bilingualism* 19(4). 423–443.

Bates, Douglas, Reinhold Kliegl, Shravan Vasishth & Harald Baayen. 2015a. Parsimonious mixed models. *arXiv preprint arXiv:1506.04967* .

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015b. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. https://doi.org/10.18637/jss.v067.i01.

Bell, Allan. 2006. Speech accommodation theory and audience design .

Bertelson, Paul, Jean Vroomen & Béatrice De Gelder. 2003. Visual recalibration of auditory speech identification: a mcgurk aftereffect. *Psychological Science* 14(6). 592–597.

Best, Catherine T. 1995. A direct realist view of cross-language speech perception. *Speech perception and linguistic experience* 171–206.

Best, Catherine T & Michael Tyler. 2007. Nonnative and second-language speech perception. *Language experience in second language speech learning: In honour of James Emil Flege* 13–34.

Bhattacharya, Tanmoy. ???? Related languages .

Birdsong, D, LM Gertken & M Amengual. 2012. Bilingual language profile: An easy-to-use instrument to assess bilingualism. coerll, university of texas at austin. web. 20 jan. 2012.

Blevins, Juliette. 2004. *Evolutionary phonology*. Cambridge: Cambridge University Press.

Bobaljik, Jonathan. 2012. *Universals in comparative morphology*. Cambridge, MA: MIT Press.

Boersma, P & D Weenink. 2016. Praat: Doing phonetics by computer (version 5.4. 15)[computer program]. retrieved april 7, 2016.

Bohn, Ocke-Schwen & James Emil Flege. 1992. The production of new and similar vowels by adult german learners of english. *Studies in Second Language Acquisition* 131–158.

Brauer, Markus & John J Curtin. 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.

Bruggeman, Laurence & Anne Cutler. 2020. No l1 privilege in talker adaptation. *Bilingualism: Language and Cognition* 23(3). 681–693.

Bullock, Barbara E & Almeida Jacqueline Toribio. 2009. Trying to hit a moving target: On the sociophonetics of code-switching. *Multidisciplinary approaches to code switching* 41. 189–206.

Caha, Pavel. 2017. How (not) to derive a *ABA: the case of Blansitt's generalization. *Glossa* 2. 84.1–32.

Caramazza, Alfonso, Grace H Yeni-Komshian, Edgar B Zurif & Ettore Carbone. 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.

Casali, Roderic. 1998. Predicting ATR activity. *Chicago Linguistic Society (CLS)* 34(1). 55–68.

Chan, Leighanne, Khia Johnson & Molly Babel. 2020. Lexically-guided perceptual learning in early cantonese-english bilinguals. *The Journal of the Acoustical Society of Amer-*

ica 147(3). EL277–EL282.

Chang, Charles B. 2012. Rapid and multifaceted effects of second-language learning on first-language speech production. *Journal of phonetics* 40(2). 249–268.

Chatterji, Suniti Kumar. 1921. Bengali phonetics. *Bulletin of the School of Oriental Studies, University of London* 2(1). 1–25.

Chládková, Kateřina, Václav Jonáš Podlipskỳ & Anastasia Chionidou. 2017. Perceptual adaptation of vowels generalizes across the phonology and does not require local context. *Journal of Experimental Psychology: Human Perception and Performance* 43(2). 414.

Chomsky, Noam. 1981. *Lectures on government and binding.* Dordrecht: Foris.

Chomsky, Noam. 1986a. *Barriers.* Cambridge, MA: MIT Press.

Chomsky, Noam. 1986b. *Knowledge of language.* New York: Praeger.

Cooper, Angela & Ann Bradlow. 2018. Training-induced pattern-specific phonetic adjustments by first and second language listeners. *Journal of phonetics* 68. 32–49.

Cutler, Anne, Ann Burchfield & Mark Antoniou. 2019. A criterial interlocutor tally for successful talker adaptation? In *19th international congress of phonetic sciences (icphs 2019).* 1485–1489. Australasian Speech Science and Technology Association Inc.

Cutler, Anne, L Burchfield & Mark Antoniou. 2018. Factors affecting talker adaptation in a second language. In *The 17th australasian international conference on speech science and technology (sst 2018).* 33–36.

Cutler, Anne, Nuria Sebastián-Gallés, Olga Soler-Vilageliu & Brit Van Ooijen. 2000. Constraints of vowels and consonants on lexical selection: Cross-linguistic comparisons. *Memory & cognition* 28(5). 746–755.

De Belder, Marijke. 2011. A morphosyntactic decomposition of countability in Germanic. *Journal of Comparative Germanic Linguistics* 14. 173–202. https://doi.org/10.1007/s10828-011-9045-0.

de Swart, Henriëtte. 1998. Aspect shift and coercion. *Natural Language and Linguistic Theory* 16. 347–385. http://www.jstor.org/stable/4047954.

Dmitrieva, Olga & Indranil Dutta. 2020. Acoustic correlates of the four-way laryngeal contrast in marathi. *Phonetica* 77(3). 209–237.

Drozdova, Polina, Roeland van Hout & Odette Scharenborg. 2014. Phoneme category retuning in a non-native language. In *Fifteenth annual conference of the international speech communication association.*

Drozdova, Polina, Roeland Van Hout & Odette Scharenborg. 2016. Lexically-guided perceptual learning in non-native listening. *Bilingualism: Language and Cognition* 19(5). 914–920.

Dutta, Indranil, Irfan. Shailendra & K.R.S. Harsha. 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*.

Eisner, Frank & James M McQueen. 2005. The specificity of perceptual learning in speech processing. *Perception & psychophysics* 67(2). 224–238.

Elias, Vanessa, Sean McKinnon & Ángel Milla-Muñoz. 2017. The effects of code-switching and lexical stress on vowel quality and duration of heritage speakers of spanish. *Languages* 2(4). 29.

Fadlon, Julie, Chuchu Li, Anat Prior & Tamar H Gollan. 2019. Using what's there: Bilinguals adaptively rely on orthographic and color cues to achieve language control. *Cognition* 191. 103990.

Flege, James E. 1995. Second language speech learning: Theory, findings, and problems. *Speech perception and linguistic experience: Issues in cross-language research* 92. 233–277.

Flege, James E. 2007. Language contact in bilingualism: Phonetic system interactions. *Laboratory phonology* 9. 353–382.

Flege, James Emil. 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, James Emil & James Hillenbrand. 1984. Limits on phonetic accuracy in foreign language speech production. *The Journal of the Acoustical Society of America* 76(3). 708–721.

Franken, Matthias K, Frank Eisner, Jan-Mathijs Schoffelen, Daniel J Acheson, Peter Hagoort & James M McQueen. 2017. Audiovisual recalibration of vowel categories. In *Interspeech 2017*. 655–658.

Franks, Steven. 2005. Bulgarian clitics are positioned in the syntax. Ms. Indiana University.

Gargesh, Ravinder. 2008a. Indian english: Phonology. *Varieties of English 4: Africa, South and Southeast Asia* 231–243.

Gargesh, Ravinder. 2008b. Indian english: Phonology. *Varieties of English* 4. 231–243.

Garry, J., C.R.G. Rubino, H. W. Wilson, A.B. Bodomo, R. French & A. Faber. 2001. *Facts about the world's languages: An encyclopedia of the world's major languages, past and present* (A New England Publishing Associates book). H.W. Wilson. https://books.google.co.in/books?id=SGtiAAAAMAAJ.

Goldrick, Matthew, Elin Runnqvist & Albert Costa. 2014. Language switching makes pronunciation less nativelike. *Psychological science* 25(4). 1031–1036.

Green, David W. 1986. Control, activation, and resource: A framework and a model for the control of speech in bilinguals. *Brain and language* 27(2). 210–223.

Green, David W. 1998. Mental control of the bilingual lexico-semantic system. *Bilingualism: Language and cognition* 1(2). 67–81.

Grosjean, François. 1998. Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and cognition* 1(2). 131–149.

Grosjean, François. 2012. An attempt to isolate, and then differentiate, transfer and interference. *International Journal of Bilingualism* 16(1). 11–21.

Grosjean, François & Joanne L Miller. 1994. Going in and out of languages: An example of bilingual flexibility. *Psychological science* 5(4). 201–206.

Guion, Susan G. 2003. The vowel systems of quichua-spanish bilinguals. *Phonetica* 60(2). 98–128.

Gullifer, Jason W & Debra Titone. 2019. Characterizing the social diversity of bilingualism using language entropy. *Bilingualism: Language and cognition* 94(1). 1–18.

Hanulíková, Adriana & Jenny Ekström. 2017. Lexical adaptation to a novel accent in german: A comparison between german, swedish, and finnish listeners. In *Interspeech*. 1784–1788.

Iverson, Gregory K. 1989. On the category supralaryngeal. *Phonology* 6. 285–303.

Khattab, Ghada. 2009. *Phonetic accommodation in children's code-switching*. na.

Khattab, Ghada. 2013. Phonetic convergence and divergence strategies in english-arabic bilingual children. *Linguistics* 51(2). 439–472.

Kim, Donghyun, Meghan Clayards & Eun Jong Kong. 2020. Individual differences in perceptual adaptation to unfamiliar phonetic categories. *Journal of Phonetics* 81. 100984.

Kim, Midam, William S Horton & Ann R Bradlow. 2011. Phonetic convergence in spontaneous conversations as a function of interlocutor language distance. *Laboratory phonology* 2(1). 125–156.

Kostić, Djordje & Rhea Stagner Das. 1972. *A short outline of bengali phonetics*. Statistical Pub. Society.

Kraljic, Tanya, Susan E Brennan & Arthur G Samuel. 2008a. Accommodating variation: Dialects, idiolects, and speech processing. *Cognition* 107(1). 54–81.

Kraljic, Tanya & Arthur G Samuel. 2005. Perceptual learning for speech: Is there a return to normal? *Cognitive psychology* 51(2). 141–178.

Kraljic, Tanya & Arthur G Samuel. 2006. Generalization in perceptual learning for speech. *Psychonomic bulletin & review* 13(2). 262–268.

Kraljic, Tanya, Arthur G Samuel & Susan E Brennan. 2008b. First impressions and last resorts: How listeners adjust to speaker variability. *Psychological science* 19(4). 332–338.

Lahiri, Aditi (ed.). 2000. *Analogy, leveling, markedness: Principles of change in phonology and morphology* (Trends in Linguistics 127). Berlin: Mouton de Gruyter.

Lecumberri, Maria Luisa Garcia, Martin Cooke & Anne Cutler. 2010. Non-native speech perception in adverse conditions: A review. *Speech communication* 52(11-12). 864–886.

Li, Rulong. 1999. Minnan fangyan de daici [Demonstrative and personal pronouns in southern Min]. In Rulong Li & Song-Hing Chang (eds.), *Daici* [Demonstrative and personal pronouns], 263–287. Guangzhou: Ji'nan University Press.

Llompart, Miquel & Eva Reinisch. 2019. Imitation in a second language relies on phonological categories but does not reflect the productive usage of difficult sound contrasts. *Language and Speech* 62(3). 594–622.

Mack, Molly. 1989. Consonant and vowel perception and production: Early english-french bilinguals and english monolinguals. *Perception & Psychophysics* 46(2). 187–200.

Masica, Colin. 1972. The sound system of indian english. monograph no. 7. .

Massam, Diane. 2000. VSO versus VOS: Aspects of Niuean word order. In Andrew Carnie & Eithne Guilfoyle (eds.), *The syntax of verb initial languages* (Oxford Studies in Comparative Syntax), 97–116. Oxford: Oxford University Press.

Maye, Jessica, Richard N Aslin & Michael K Tanenhaus. 2008. The weckud wetch of the wast: Lexical adaptation to a novel accent. *Cognitive Science* 32(3). 543–562.

McCarthy, John J. & Alan S. Prince. 1999. Prosodic morphology. In John A. Goldsmith (ed.), *Phonological theory: The essential readings*, 238–288. Malden, MA and Oxford: Blackwell.

McGurk, Harry & John MacDonald. 1976. Hearing lips and seeing voices. *Nature* 264(5588). 746–748.

McQueen, James M, Anne Cutler & Dennis Norris. 2006. Phonological abstraction in the mental lexicon. *Cognitive science* 30(6). 1113–1126.

Mitterer, Holger & Mirjam Ernestus. 2008. The link between speech perception and production is phonological and abstract: Evidence from the shadowing task. *Cognition* 109(1). 168–173.

Mitterer, Holger & James M McQueen. 2009. Foreign subtitles help but native-language subtitles harm foreign speech perception. *PloS one* 4(11). e7785.

Muldner, Kasia, Leah Hoiting, Leyna Sanger, Lev Blumenfeld & Ida Toivonen. 2019. The phonetics of code-switched vowels. *International Journal of Bilingualism* 23(1). 37–52.

Murray, Robert W. & Theo Vennemann. 1983. Sound change and syllable structure in Germanic phonology. *Language* 59(3). 514–528.

Myers-Scotton, Carol. 1993. Dueling languages: Grammatical structure in code-switching. claredon.

Norris, Dennis, James M McQueen & Anne Cutler. 2003. Perceptual learning in speech. *Cognitive psychology* 47(2). 204–238.

Olson, Daniel J. 2013. Bilingual language switching and selection at the phonetic level: Asymmetrical transfer in vot production. *Journal of Phonetics* 41(6). 407–420.

Olson, Daniel 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, Pramod. 2015. 17 indian english pronunciation. *The handbook of English pronunciation* 301.

Paradis, Michel. 1993. Linguistic, psycholinguistic, and neurolinguistic aspects of" interference" in bilingual speakers: The activation threshold hypothesis. *International Journal of Psycholinguistics* .

Pardo, Jennifer S. 2006. On phonetic convergence during conversational interaction. *The Journal of the Acoustical Society of America* 119(4). 2382–2393.

Pardo, Jennifer S. 2010. Expressing oneself in conversational interaction. *Expressing oneself/expressing one's self: Communication, cognition, language, and identity* 183–196.

Piccinini, Page & Amalia Arvaniti. 2015. Voice onset time in spanish–english spontaneous code-switching. *Journal of Phonetics* 52. 121–137.

R Core Team. 2013. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing Vienna, Austria. http://www.R-project.org/.

R Core Team. 2016. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing Vienna, Austria. https://www.R-project.org/.

Reinisch, Eva, Andrea Weber & Holger Mitterer. 2013. Listeners retune phoneme categories across languages. *Journal of Experimental Psychology: Human Perception and Performance* 39(1). 75.

Rissanen, Matti. 1999. Syntax. In Roger Lass (ed.), *Cambridge history of the English language*, vol. 3, 187–331. Cambridge: Cambridge University Press.

Rodriguez-Fornells, Antoni, Ulrike Kramer, Urbano Lorenzo-Seva, Julia Festman & Thomas F Münte. 2012. Self-assessment of individual differences in language switching. *Frontiers in Psychology* 2. 388.

Samuel, Arthur G. 1986. Red herring detectors and speech perception: In defense of selective adaptation. *Cognitive psychology* 18(4). 452–499.

Sancier, Michele L, Carol A Fowler et al. 1997. Gestural drift in a bilingual speaker of brazilian portuguese and english. *Journal of phonetics* 25(4). 421–436.

Schertz, Jessamyn, Taehong Cho, Andrew Lotto & Natasha Warner. 2016. Individual differences in perceptual adaptability of foreign sound categories. *Attention, Perception, & Psychophysics* 78(1). 355–367.

Schuhmann, Katharina S. 2016. Cross-linguistic perceptual learning in advanced second language listeners. *Proceedings of the Linguistic Society of America* 1. 31–1.

Schwartz, Geoffrey, Anna Balas & Arkadiusz Rojczyk. 2015. Language mode vs. l2 interference: Evidence from l1 polish. In *Icphs*.

Šimáčková, Šárka & Václav J Podlipský. 2015. Immediate phonetic interference in code-switching and interpreting. In *Proceedings of the 18th international congress of phonetic sciences. glasgow, uk*.

Šimáčková, Šárka & Václav Jonáš Podlipský. 2018. Patterns of short-term phonetic interference in bilingual speech. *Languages* 3(3). 34.

Simonet, Miquel. 2014. Phonetic consequences of dynamic cross-linguistic interference in proficient bilinguals. *Journal of Phonetics* 43. 26–37.

Simonet, Miquel & Mark Amengual. 2020. Increased language co-activation leads to enhanced cross-linguistic phonetic convergence. *International Journal of Bilingualism*

24(2). 208–221.

Sirsa, Hema & Melissa A Redford. 2013. The effects of native language on indian english sounds and timing patterns. *Journal of phonetics* 41(6). 393–406.

Sjerps, Matthias J, Holger Mitterer & James M McQueen. 2011. Listening to different speakers: On the time-course of perceptual compensation for vocal-tract characteristics. *Neuropsychologia* 49(14). 3831–3846.

Stewart, Thomas W., Jr. 2000. *Mutation as morphology: Bases, stems, and shapes in Scottish Gaelic*. Columbus, OH: The Ohio State University dissertation.

Thundy, Zacharias. 1976. The origins of indian english. *CIEFL Bulletin (Hyderabad)* 12. 29–40.

Tobin, Stephen J, Hosung Nam & Carol A Fowler. 2017. Phonetic drift in spanish-english bilinguals: Experiment and a self-organizing model. *Journal of phonetics* 65. 45–59.

Tollefson, James W & Amy BM Tsui. 2014. Language diversity and language policy in educational access and equity. *Review of Research in Education* 38(1). 189–214.

Tsui, Rachel Ka-Ying, Xiuli Tong & Chuck Siu Ki Chan. 2019. Impact of language dominance on phonetic transfer in cantonese–english bilingual language switching. *Applied Psycholinguistics* 40(1). 29–58.

van Coetsem, Frans. 2000. *A general and unified theory of the transmission process in language contact*. Heidelberg: Winter.

Van Linden, Sabine & Jean Vroomen. 2007. Recalibration of phonetic categories by lipread speech versus lexical information. *Journal of Experimental Psychology: Human Perception and Performance* 33(6). 1483.

Wells, John C & John Corson Wells. 1982. *Accents of english: Volume 1*, vol. 1. Cambridge University Press.

Wiener, Seth & Rory Turnbull. 2016. Constraints of tones, vowels and consonants on lexical selection in mandarin chinese. *Language and speech* 59(1). 59–82.

Winter, Bodo. 2013. Linear models and linear mixed effects models in r: Tutorial 11. *arXiv preprint arXiv:1308.5499* .

Xie, Xin, Rachel M Theodore & Emily B Myers. 2017. More than a boundary shift: Perceptual adaptation to foreign-accented speech reshapes the internal structure of phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance* 43(1). 206.

Yu, Alan. 2003. *The morphology and phonology of infixation*. Berkeley, CA: University of California dissertation.

Zheng, Yi & Arthur G Samuel. 2019. The relationship between phonemic category boundary changes and perceptual adjustments to natural accents. *Journal of Experimental Psychology: Learning, Memory, and Cognition* .