Bilingualism: Language and Cognition

cambridge.org/bil

Mixed language processing increases cross-language phonetic transfer in Bengali–English bilinguals

Auromita Mitra[1] and Indranil Dutta[2]

[1]New York University, New York, USA and [2]Jadavpur University, Kolkata, India

Research Article

Cite this article: Mitra A, Dutta I (2023). Mixed language processing increases cross-language phonetic transfer in Bengali–English bilinguals. *Bilingualism: Language and Cognition* **26**, 896–909. https://doi.org/10.1017/S1366728923000159

Received: 9 July 2021
Revised: 19 January 2023
Accepted: 1 February 2023
First published online: 27 February 2023

Keywords:
phonetic transfer; vowel quality; bilingual speech production; L1-L2 interaction; Indian English; Bengali

Address for correspondence:
Auromita Mitra,
Department of linguistics,
New York University,
10 Washington Pl,
New York, NY – 10003
E-mail: auromita.mitra@nyu.edu

Abstract

This study investigates the phonetic outcome of mixed-language speech in Bengali and Indian English, towards understanding cross-language interaction in highly proficient bilinguals. We compare spectral properties of L2 vowels [æ] (common to L1, L2) and [ʌ] (absent in L1) in code-switched (mixed) vs. nonswitched productions. Results reveal asymmetrical shifts in both vowels during mixed productions, towards related L1 categories. We interpret this as a temporary increase in cross-language phonetic interaction during mixed-language use, leading to a shift towards L1 norms, evidencing transfer effects on L2 vowels. We elicit mixed productions through two tasks: cued picture-naming and code-switching, to assess if experimental paradigm independently influences the behavior under study. Results reveal parallel patterns, but small magnitude differences, across paradigms. We discuss these findings in light of recent proposals about asymmetries in short-term phonetic interaction, postulated discursive factors in code-switching, and the issue of comparability between paradigms in transfer studies.

1. Introduction: Phonetic transfer

Bi/multilingual speakers can distinguish between the phonetic norms of their languages and maintain separate sound categories for each language (Bosch & Sebastián-Gallés, 2003; Caramazza, Yeni-Komshian, Zurif, & Carbone, 1973; MacLeod & Stoel-Gammon, 2010). However, these categories are not autonomous – they influence each other across languages in both perception and production (e.g., Flege, 1995; Flege, MacKay, & Piske, 2002; Fowler, Sramko, Ostry, Rowland, & Hallé, 2008), and the nature of such interaction provides crucial insights into how language 'systems' are cognitively represented and processed. Based on the productions of proficient bilingual speakers of Bengali and English, this study reports that cross-language phonetic influence temporarily increases during mixed-language use.

Cross-language influence at the level of sounds can be studied in broadly two kinds of conditions: (i) while a bilingual speaker is operating in any one of their languages, by comparing bilingual speech to monolingual norms (e.g., Caramazza et al., 1973; Flege, 1987; Guion, 2003); (ii) when both languages of a bilingual speaker are co-activated, by comparing speakers' mixed-language speech to their own norms while using a single language (e.g., Bullock & Toribio, 2009; Elias, McKinnon, & Milla-Muñoz, 2017; Grosjean & Miller, 1994; Simonet, 2014). The present study is concerned with the latter. In production, this kind of influence has been variously termed "transfer", "drift", "accommodation", and "interference". We use the term "transfer" here, to indicate any interaction between two sets of phonetic norms.

Existing research on the phonetic effects of mixed-language production has largely focused on a limited set of phonologically related languages. A majority of these studies use temporal properties of consonants (in particular, voice onset time or VOT) to measure transfer. However, reported results vary greatly across studies, as discussed in the following subsections, and appear to be contingent upon both language-specific features and language experience of the participants. This emphasizes the importance of considering data from a wider variety of populations, language pairs, and phonetic features in order to make meaningful generalizations. Widespread multilingualism in the Indian subcontinent suggests that phonetic behavior in these populations can be particularly valuable towards understanding the nature of such cross-language interactions, as it is likely to reflect real-world experience with mixed-language processing. However, there is no work yet on short-term phonetic transfer in the Indian subcontinent, or in any Indo-Aryan language.

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 baseline production of L2. Mixed-language data are elicited in two switching paradigms: cued picture-naming and code-switching. We compare these results to

© The Author(s), 2023. Published by Cambridge University Press. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

assess if differences between the paradigms independently influence the outcome of phonetic interaction. The results demonstrate a shift in L2 vowel quality during mixed-language production, with parallel patterns but different magnitude of transfer between the two paradigms. We discuss these findings in light of recent proposals about asymmetries in short-term phonetic interaction and the role of connected speech in introducing discursive factors to transfer studies. In subsequent sections, the acoustic feature examined in a study is often indicated in parentheses following its citation, for clarity.

1.1. Bengali and English in India

Demography

Bengali is an Indo-Aryan language primarily spoken in India and Bangladesh. In India, more than 97 million people speak Bengali as a first language (Office of the Registrar General & Census Commissioner, India, 2011), mostly in the state of West Bengal. A majority of this population also speaks additional languages. Indian English (IE) refers to the variety of English spoken in the Indian subcontinent. It is spoken as an L2 by 129 million people (Office of the Registrar General & Census Commissioner, India, 2011). English is one of two official languages, used in education, law, media, as a lingua franca mainly by an educated elite in most metropolitan regions, and carries high prestige value (Kachru, 1983; Kachru & Smith, 1981; Pandey, 2015; Tollefson & Tsui, 2014).

Indian English is not a monolithic entity; there are large regional variations in the population, and the literature on IE phonology is thus neither "standardized" nor uncontested. Recent research suggests that IE has a target phonology that is distinct from the varied L1s of its speakers, as well as from other native varieties of English (see Sirsa & Redford, 2013 for a review). For the present study, we focus on "General(ized) Indian English" (Masica, 1972). We focus specifically on short-term phonetic interaction during mixed-language use, because:

(i) In multilingual populations, long-term representations of sound categories are expected to be affected by multiple languages.
(ii) Given the rarity of IE phonology without L1 'influence' (it is almost exclusively spoken as an L2) it is more meaningful to think of cross-language transfer in IE as relative to a speaker's own production in a baseline condition.

The socio-linguistic facts around English in India also suggest that mixed-language processing of English and an L1 is an ecologically valid paradigm to understand short-term phonetic interaction.

Vowel systems

The vowel inventory of Western Bangla consists of [i, e, æ, a:, ɔ, o, u], and their nasalized counterparts (Rubino & Garry, 2001). The vowel chart in Figure 1 (from Ghosh, 2012) shows their distribution in the F1×F2 formant space. Note that there is no mid-central vowel category.The vowel system of IE contains the monophthongs [I, i, E, e, æ, ə/ʌ/ɜ, a, ɔ, o, ʊ, u] (Masica, 1972; Wells & Wells, 1982). The vowel chart in Figure 2 shows their distribution in the F1×F2 formant space. A single mid-central vowel corresponds to the categories [ʌ,ə,ɜ], which are treated as distinct in many varieties of English (Bansal, 1969; Hickey, 2005; Nihalani, Tongue, Hosali, & Crowther, 1979; Wells & Wells, 1982). 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 (e.g., Speech Learning Model – Flege, 1995, 2007; Perceptual Assimilation Model-L2 – Best & Tyler, 2007), from interaction during accessing, processing, or articulation of these categories (transfer vs. interference (Grosjean, 2012); competence vs.

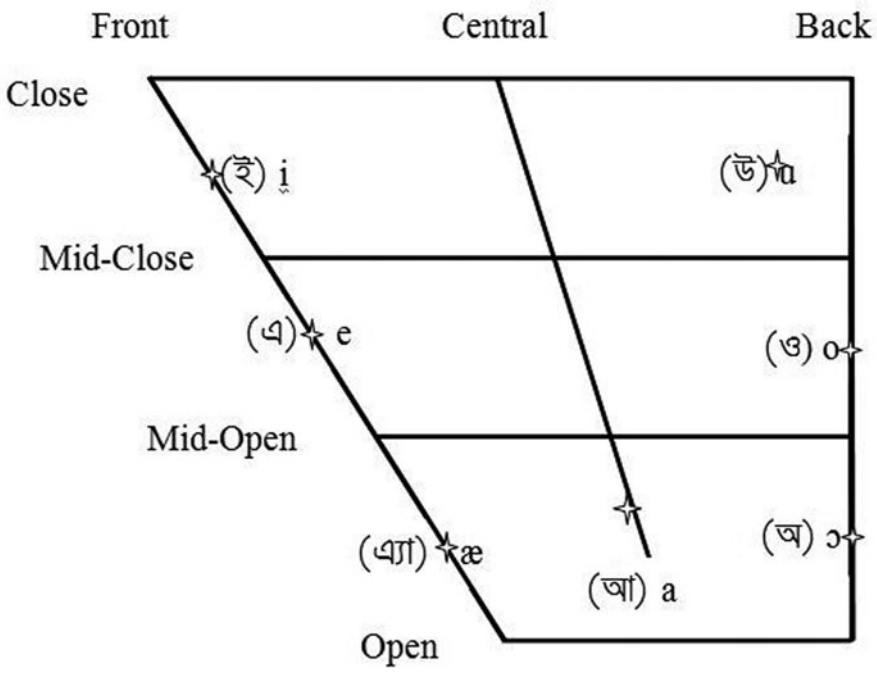


Fig. 1. Vowel chart for oral monophthongs in Bengali, from Ghosh (2016)

https://doi.org/10.1017/S1366728923000159 Published online by Cambridge University Press

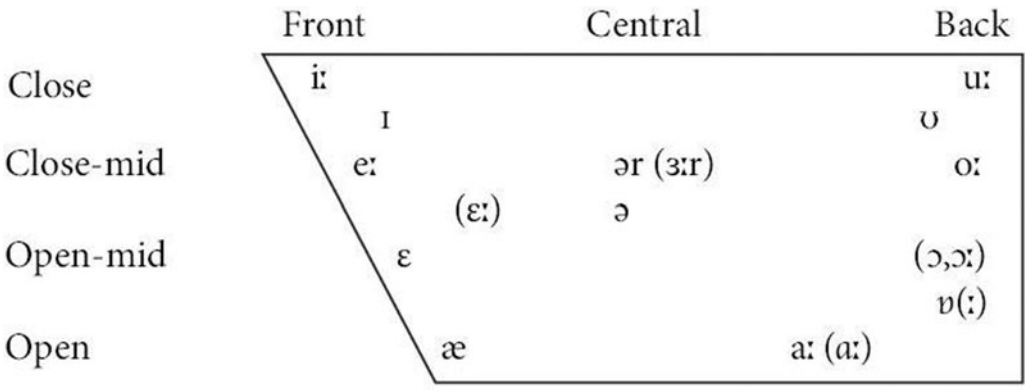


Fig. 2. Vowel chart for monophthongs in General(ized) Indian English, from Masica (1972)

performance interference (Paradis, 1993)). Given their transient nature, dynamic changes in production during mixed-language use are generally attributed to the latter, e.g., online processing costs (Olson, 2013; Šimáčková & Podlipskỳ, 2015; Tsui, Tong, & Chan, 2019, VOT), language mode (Simonet, 2014, vowel quality), context-awareness (Khattab, 2013, phonological variables). What triggers this interaction? Olson (2016) argues that cross-language phonetic effects could have two potential sources: (i) the local point of switch, and (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, 2014; Simonet & Amengual, 2020, vowel quality);
(ii) (however, language mode is not the sole source of influence during mixed language use – studies comparing switched and nonswitched tokens produced in the same test block (identical language mode) (Olson, 2016; Tsui et al., 2019, VOT), or spontaneous conversation (Piccinini & Arvaniti, 2015, VOT), 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, VOT) found no additive effects of language mode, Olson (2013, VOT) found a 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 nonswitched tokens in a monolingual test block, separated by a few hours to days (Elias et al., 2017, vowel quality), (Antoniou, Best, Tyler, & Kroos, 2011; Bullock & Toribio, 2009; Schwartz, Balas, & Rojczyk, 2015; Šimáčková & Podlipskỳ, 2015, 2018, VOT).

Since participants of this study are in an environment containing mixed-language input and multilingual interlocutors, we expect that phonetic transfer during everyday language use takes place in a bilingual mode (see Grosjean, 1998). We elicited all utterances in the same bilingual test block. Observed differences in this paradigm would arguably result from interaction during online processing of sounds while switching between languages. In bilingual mode, both language systems are expected to be nearly equally accessible throughout the test block. Thus, a consistent difference between switched and nonswitched 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 of transfer effects

Longitudinal studies and between-subject comparisons of speakers with different durations of L2 exposure suggest that interaction between bilingual sound systems is dynamic (e.g., Bohn & Flege, 1992). Changes due to transfer are not necessarily unidirectional or irreversible. Sancier and Fowler (1997, VOT) first demonstrated that spending 2-5 months in an L1 or L2 environment causes productions in both languages to "drift" towards the ambient language, showing that cross-language interaction can be triggered in an order of months, and reversed within a similar time range. A recent study by Tobin, Nam, and Fowler (2017, VOT) reported comparable effects in an even shorter duration (2-4 weeks). Based on a short-term longitudinal study, Chang (2012, VOT) demonstrates that over the course of the first five weeks of learning an L2, there is a gradual convergence of L1 towards L2. For some sounds and features (but not others, see sec.1.5), transfer was additive over time. How long these effects last in the absence of regular L2 input was not tested.

Studies discussed above concern situations where a speaker is operating in any one of their languages. For mixed-language use, the majority of existing studies only analyze the switched (target) token – there are few direct measurements of the duration of short-term transfer effects. One study which measured this in a code-switching paradigm (Bullock & Toribio, 2009, VOT) did not find any residual effects on the matrix language following a switch, suggesting that transfer during code-switching is localized. Indirect evidence for this comes from experiments that have elicited switched and nonswitched tokens in the same test block and still reported differences between the two (e.g., Olson, 2013; Tsui et al., 2019, VOT), suggesting that changes due to transfer are quickly 'reset' – in an order of seconds. Note that all the studies of mixed-language production discussed above measure VOT, which is a temporal feature. We have no a priori 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

https://doi.org/10.1017/S1366728923000159 Published online by Cambridge University Press

rapid shifts in vowel quality within comparable time-frames (e.g., Babel, 2010, 2012; Pardo, 2010). In this study, we present switched and nonswitched 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

A recurring pattern in the literature on short-term phonetic transfer is that the extent and patterns of phonetic shift are not equivalent across the two languages of the bilingual speaker. Studies vary greatly in the exact differences they report. The sources of such asymmetry are of interest because they point towards the factors that mediate cross-language phonetic interaction. In this section, we discuss two proposed sources of such asymmetry that are relevant to the present study:

L1 vs. L2 status

Flege's Speech Learning Model (SLM) (1995, 2007) proposes that the sound categories of a bilingual speaker exist in a common phonological space – in principle, both L1 and L2 categories can influence one another. Patterns of cross-language transfer depend on how L2 phonemes are mapped in relation to the existing L1 categories. In nonswitched production, this is evidenced through pervasive L1 influence on non-native contrasts that are perceptually linked to an existing L1 contrast (see equivalence classification: Flege, 1987; Flege & Hillenbrand, 1984), and the observation that both L1 and L2 sound systems of bilinguals differ from those of corresponding monolingual speakers (e.g., Guion, 2003, vowel quality).

In mixed-language production, the role of language status is less clear – studies have reported unidirectional influence of L1 on L2 (Antoniou et al., 2011; Balukas & Koops, 2015; Goldrick, Runnqvist, & Costa, 2014; Šimáčková & Podlipskỳ, 2015, VOT), L2 on L1 (Olson, 2013; Tsui et al., 2019, VOT), (Elias et al. (2017), vowel quality), bidirectional convergence (Bullock & Toribio, 2009; Olson, 2016, VOT), divergence (Bullock & Toribio, 2009; Šimáčková & Podlipskỳ, 2018, VOT), and no influence (Muldner, Hoiting, Sanger, Blumenfeld, & Toivonen, 2019, vowel quality), (Schwartz et al., 2015, phonological process). 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.

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 established that, beyond language-specific differences, the L1 vs L2 status of the language does mediate transfer. They interpret this asymmetry in terms of the Inhibitory Control Model (ICM) (Green, 1998): to select a phonetic realization from one language, the other must be inhibited. Greater inhibition is required on L1 while using an L2. This greater initial inhibition on L1 means that switching into L1 incurs a greater "switch cost" (more cross-language influence) than switching into L2. Lower switch costs translate to a lack of visible cross-language transfer effects on L2. Tsui et al. (2019) report 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 cost in both languages) due to greater experience with language switching.

Since differences between long-lag and short-lag languages independently affect the phonetic realization of VOT, focusing on other phonetic features as sites for transfer can avoid this conflation, and clarify the precise effect of language status. However, given the implications of the ICM, it is necessary to first establish that these L2 categories can indeed be affected by dynamic interference, particularly in bilinguals with extensive language-switching experience. There are relatively few studies on other phonetic features such as vowel quality (see Elias et al., 2017; Muldner et al., 2019; Simonet, 2014) or phonological processes (see Schwartz et al., 2015; Simonet & Amengual, 2020), and many of these focus on L1 categories as the target for transfer. This study builds on existing work by examining whether the L2 vowel quality of proficient bilinguals can be affected by mixed-language processing.

Differences between sound systems of the languages

In addition to the cognitive factors discussed above, many researchers have attributed asymmetries to language-internal factors, such as the shape, size, and composition of the sound inventory. For example, frequently reported asymmetries between long- and short-lag VOT languages (see, e.g., Antoniou et al., 2011; Bullock & Toribio, 2009; Chang, 2012; Olson, 2016; Tobin et al., 2017) and postulated frequency effects (Antoniou et al., 2011) highlight that beyond interaction during the cognitive processing of sounds, phonetic transfer is ultimately a linguistic phenomenon, subject to language-specific phonological constraints. Therefore, it is imperative to consider data from a variety of language pairs. Existing research on phonetic transfer largely focuses on a limited set of phonologically related languages. The present study extends the scope of this research to a new pair of languages – Indian English and Bengali.

1.5. Asymmetries between sounds

Asymmetries in the extent and patterns of transfer have not only been observed between languages, but also between different sounds/features of a language, in both long-term and transient interactions. Studies that have examined multiple sound categories have found that interactions between individual sound pairs do not necessarily parallel the overall pattern of global (system-wide) shift (e.g., Chang, 2012; Elias et al., 2017, vowel quality), suggesting that 'extent of transfer' cannot treated as an atomic measure. In light of the discussion about VOT in section 1.4, this is not surprising – if a general linguistic principle of 'room for movement' and contrast constrains transfer, then we should expect it to apply to individual sounds too. Once again, this emphasizes the importance of examining a wider range of sound contrasts.

In the present study, we focus on 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:

1. Position in the IE vowel space: compared to [æ], [ʌ] exists in a part of the vowel space that has lower vowel density. This affords a greater latitude for movement without risking the loss of a contrast, particularly in the vowel height (F1) dimension. Thus, considering purely phonological constraints on IE,

https://doi.org/10.1017/S1366728923000159 Published online by Cambridge University Press

we 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 backness (F2) dimension.

2. Target of transfer: existing research on cross-language phonetic interaction suggests that changes in production during mixed-language use are not random, but rather targeted with respect to categories in the other language. Thus, another source of asymmetry is the fact that the category [ʌ] is absent in Bengali, whereas [æ] is a common category across both languages (see sec. 1.1).

Flege's Speech Learning Model (SLM) (1995, 2007) posits 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, participants' Catalan [ɔ] shifted towards an acoustically close category [o], which is common to Catalan and Spanish, suggesting that movement might be towards a set of phonetic norms rather than a specific corresponding category.

We expect the vowel [ʌ] to shift towards an acoustically/perceptually close category in IE – a probable candidate being the low vowel [a:], as anecdotal observations of Bengali-accented English suggest that the sounds are perceived as close by Bengali speakers.

Our research questions in this study concern whether L2 vowels are produced differently in nonswitch vs switch conditions. Existing literature suggests that such shifts are directional and targeted. Although our primary concern is whether productions in the two conditions differ, we wanted to understand how the direction of shift, if any, relates to corresponding Bengali vowel categories. To estimate average baseline differences in vowel quality between the relevant English and Bengali categories, we processed and used existing speech data from a freely available Bengali corpus, SHRUTI (Das, Mandal, & Mitra, 2011).

We extracted formant values from syllable-initial productions of [æ] and [a:] at 55% into the vowel, and compared these to our participants' productions in the nonswitched English condition. This is visualized in Figure 3. The F1×F2 plot indicates that the Bengali category [æ] is more fronted, 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?

This section motivates a secondary aim of this study: to verify if the extent and patterns of phonetic interaction in bilingual participants are independently mediated by the choice of experimental paradigm. Studies of phonetic transfer during mixed-language use have largely employed spontaneous or scripted code-switching (CS) to elicit mixed productions. CS is defined as the use of multiple languages within a single utterance, i.e., in connected speech (Myers-Scotton, 1993). Other paradigms have included 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 reflect patterns of language practice, rather than cognitive behavior. This finds support from studies that report transfer effects before the switch point (Bullock & Toribio, 2009), or other evidence for planning such as hyperarticulation (Muldner et al., 2019) and interlocutor-awareness in spontaneous CS (Khattab, 2013). To avoid such confounds, Olson (2013) endorses 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. This observation has important implications for how the existing literature on short-term phonetic interaction is interpreted: if the discursive properties of a CS paradigm independently affect the phonetic

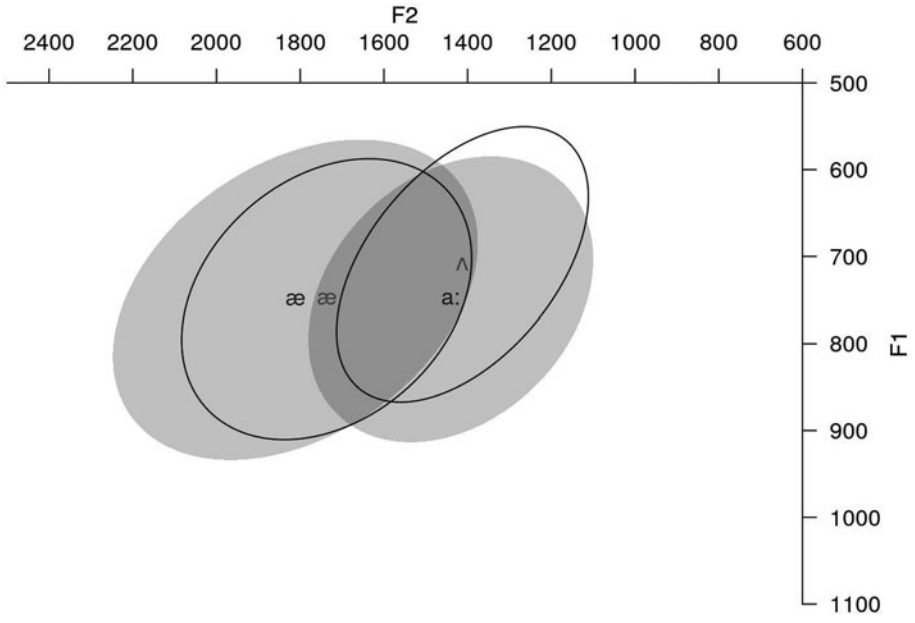


Figure 3: Participants' unilingual English vowels (shaded ellipses) compared to Bengali categories from the SHRUTI corpus (Das et al., 2011)

https://doi.org/10.1017/S1366728923000159 Published online by Cambridge University Press

outcome of mixed-language use, this calls to question the comparability of results across studies using these different paradigms. Because the reported outcomes of short-term phonetic interaction vary so greatly across studies (see section 1.4), generalizations about the mechanisms that underlie these patterns often rely crucially on the assumption of such comparability. Therefore, identifying paradigm-specific differences is particularly relevant for this line of research.

Note that while there is evidence for planning in CS, precisely how this affects phonetic interaction 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 participant group converged 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 (indicating some degree of planning), and did not evidence a significant shift during mixed-language use. Moreover, recall that the findings of Olson (2016) suggest that there is a limit on the extent of transfer – multiple factors do not necessarily result in additive effects. Thus, the presence of additional factors in CS does not necessitate that this should alter the extent of observable transfer compared to any other paradigm. It is possible that the outcomes in the studies discussed above are better attributed to some other aspect of the experimental setup, rather than the act of code-switching in particular.

The precise contribution of various (linguistic and extra-linguistic) factors and their interactions to the phonetics mixed-language use merits detailed investigation in its own right. In the present study, however, our focus is limited to verifying whether, in an experimental setting, the postulated discursive elements of a CS paradigm alter the extent of observable transfer compared to a paradigm that arguably lacks these: cued picture-naming. We are less concerned with the exact nature of differences than with their comparability. While researchers have successfully employed each of these paradigms to induce phonetic interaction, comparison across studies is difficult since they also differ in other factors. We provide a starting point for such comparison by eliciting productions of the same tokens from the same set of participants, in both paradigms. A difference in results will suggest that the planning involved in connected speech and code-switching affects transfer above and beyond the phonetic outcome of switching itself, which has implications that future work should consider.

Research questions and hypotheses

RQ1 – L2 vowels: Are L2 vowel categories of proficient bilingual speakers affected by dynamic phonetic transfer during mixed-language production?

H1 – L2 vowels: We expect to see differences in the spectral properties of L2 vowels in switched vs nonswitched tokens

RQ2 – Asymmetry: Do patterns of transfer differ between the two vowel categories?

H2 – Asymmetry: Since existing literature demonstrates targeted and phonologically-constrained shifts, we expect transfer effects to be category-specific. Based on the position of IE vowels compared to those of Bengali vowels from Das et al. (2011), we expect [ʌ] to be lowered and fronted in the mixed condition. [æ] is expected to be fronted. We hypothesize the extent of shift in [æ] to be lesser than [ʌ], as it is in a part of the IE vowel space that has greater vowel density.

RQ3 – Paradigm: Is there an observable difference between the extent of transfer in a cued picture-naming vs. a code-switching paradigm?

H3 – Paradigm: While researchers have discussed code-switching as a source of discursive factors likely to affect transfer behavior, this has not been systematically studied. Existing literature does not allow us to predict precisely what form this effect might take (see discussion in Section 1.6). Thus, this question is exploratory, and meant to serve as a starting point for more within-study comparisons.

2. Methodology

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

2.1. Participants

Participants were recruited through an open call at the EFL University campus in Hyderabad. Volunteers completed a brief language background questionnaire, which was used to screen participants.

The university campus has a linguistically diverse population with varying L1s and varieties of English. Our aim was to minimize between-participant random variability, while maintaining ecological validity considering the actual day-to-day settings in which transfer behaviors likely take place. We chose two metrics for this: (i) where the participant grew up; and (ii) Medium of Instruction (MoI) in school. The first metric ensures that the variety of Bengali that individuals were exposed to is similar across participants, resulting in similar target representations for Bengali vowel categories. This was especially important because our study does not elicit unilingual Bengali vowels, instead using existing corpus data to estimate the position of Bengali vowels in the formant space. Since we expect L2 vowels to shift towards related L1 categories during mixed language utterances, we wanted to ensure that any between-participant asymmetries reflected individual differences in transfer, rather than different acoustic targets. Another reason to control for location was that it is a fairly reliable predictor for language(s) of education (see supplementary materials for details). The second metric reflects sustained formal education in a language, which predicts LSRW (listening, speaking, reading, writing) skills. Since this study uses written stimuli, we wanted to ensure that all participants were sufficiently comfortable with reading and speaking both Bengali and English, to avoid variability due to task effort.

Given these considerations, we recruited participants who (i) grew up and lived in the state of West Bengal for the majority of their lives; (ii) had Bengali-speaking parents; (iii) received at least 5 years of formal education in Bengali (all respondents had 5+ years of education in English and were pursuing degrees taught in English); and (iv) self-reported LSRW proficiency in Bengali.

28 bilingual speakers of Bengali and English (18 female) responded to the recruitment call. Using the inclusion criteria reported above, 10 volunteers (5 female, 5 male, age range 19–28) participated in the study, and were compensated for their time. We followed up with a detailed language background survey at a later date to learn more about the language profiles of the participants. This survey was a modified and consolidated version of three popular bilingualism profiling tools: the Bilingual Language

https://doi.org/10.1017/S1366728923000159 Published online by Cambridge University Press

Table 1. Key biographical information summarized across participants. Scales: age 0(since birth)—5(since age 5) etc.; likert scale: 0(not proficient at all)—6(highly proficient)

	Bengali			English		
	Mean	SD	Range	Mean	SD	Range
Age of acquisition	0.25	0.71	0–2	2.6	2.13	0–6
Years of formal education	11.8	4.92	8–15	16+	4.15	7–20+
Self-rated proficiency speaking	6	0	6–6	6	0	6–6
listening	5.87	0.35	5–6	6	0	6–6
reading	4.87	1.24	3–6	6	0	6–6

Profile (BLP) (Birdsong, Gertken, & Amengual, 2012), the Language Experience and Proficiency Questionnaire (LEAP-Q) (Marian, Blumenfeld, & Kaushanskaya, 2007), and the Bilingual Switching Questionnaire (BSWQ) (Rodriguez-Fornells, Kramer, Lorenzo-Seva, Festman, & Münte, 2012). Together, these incorporate questions about language history, usage, self-assessed proficiency, attitudes, and switching experience. Table 1 summarizes some key biographical data from the survey. The complete questionnaire, detailed discussion of the linguistic situation in India, and language profiles of participants are included as Appendix S1 in the OSF repository for this study.

Although existing studies in the field have used comparable participant numbers, we acknowledge that this is a small sample size. As detailed above, even after a recruitment call that already controlled for language background and education level, we were able to include only 10 out of 28 volunteers (35.7%) in the final study. The criteria used here establish only a minimal level of between-participant similarity in exposure, dominance, habits, and proficiency, such as is extremely common in most comparable studies. This demonstrates the degree of linguistic variability in the present population. Norms for experimental research on bilingualism have largely been established in a Western context. In settings where linguistic variability at the population level is much greater than these canonically studied groups, establishing comparable levels of experimental control requires significantly greater resources. Given that small sample sizes are likely to have low statistical power, replication with larger sample sizes are important to confirm the observed effects, and provide much-needed data diversity. A discussion of power analyses in bilingualism research is included as Appendix S2 in the OSF repository.

2.2. Stimuli

Twenty monosyllabic English words, ten each containing the vowels [ʌ] and [æ], were selected as target words. To minimize coarticulatory effects on the vowel, pre-vocalic consonants were all voiced plosives ([b] or [d]). We avoided English words that are already lexicalized loanwords in Bengali. 10 monosyllabic English words that do not contain the target vowels served as filler items. The complete list of target words can be found in the "Materials" component of the OSF repository for this project.

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. Participants completed four iterations each of two tasks – cued picture-naming and code-switching. Utterances in the two conditions (nonswitch and switch/mixed) were randomly interspersed within each iteration. In the nonswitch condition, the target word was preceded by another English word in the utterance. In the mixed condition, the target was preceded by a Bengali word and thus produced as a switch from Bengali. Every target word appeared twice during a test block; once in a nonswitch and once in a switch condition. Participants alternated between the two tasks, and were allowed to take a short break after each test block. The tasks are described below:

Cued picture-naming: Participants were presented with slides in the following sequence (see fig. 4a,4b)

- Language cue: a word in either English or Bengali orthography, read silently
- A picture displayed for 50 ms; named aloud by the participant in either English or Bengali, depending on the language cue. The picture of a fish was used in all trials, prompting the participant to say either [fIʃ] (English) or [maː tʃʰ] (Bengali)
- Target: an English word; read aloud by the participant
- Distracter math problem

The language cue was used to manipulate the word preceding the target, giving target words in nonswitched and switched (mixed) utterances.

Code-switching: Each slide contained one English target word embedded in either an English (nonswitch) or a Bengali (mixed; target produced as a switch from Bengali) carrier2. 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-ʈa æk-ʈa kalo [target word].

Gloss and translation for Bengali sentence:

o-ʈa æk-ʈa kalo [target word]
that-CLF one-CLF black [target word]
"that is a black [target word]

Carrier sentences were such that the sound preceding the target word was uniform across all sentences (the final sound in "yellow" is [o] in Indian English), to minimize differences in coarticulatory effects on the target vowel. We chose a specifier-noun construction, which has identical word-order in Bengali and English. This gave a uniform prosodic context for the target word across

https://doi.org/10.1017/S1366728923000159 Published online by Cambridge University Press

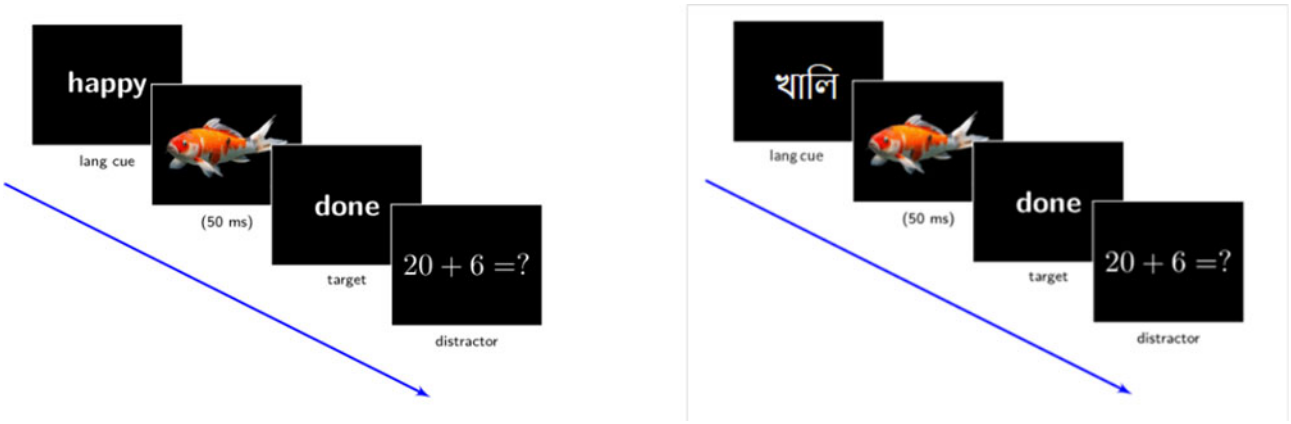


Figure 4: (a) English (nonswitch) trial (b) Bengali-English (switch/mixed) trial Picture-naming task sample sequence: language cue (not read out), picture (named as per language cue), target (read out), distractor

conditions. While presenting the mixed-language utterances, we used mixed orthography.

2.4. Acoustic analysis

The first iteration of each task was treated as a trial, and excluded from further analyses. An additional 67 tokens (2.79%) were excluded due to errors in presentation/production, giving 2333 target word tokens ([ʌ]: 1137, [æ]: 1196) 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, 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. Individual vocal tract differences affect formant frequencies across speakers. To reduce between-speaker differences, we used z-score normalization (using the PhonR package (McCloy, 2016) in R (R Core Team, 2013), representing each speaker's formant frequencies as the scaled distance from their own mean).

Plotting these vowels shows a difference between male and female speakers, such that male speakers have overall lower formant frequencies (Figure S1a). Since gender-related differences in formant frequencies are not of interest in this study, we normalized individual values by gender (representing each value as a scaled distance from the group mean; Figure S1b).[1] Statistical analyses were performed on these normalized values.

3. Results

All statistical analyses were performed in R (R Core Team, 2013). We used linear mixed effects models to test the effect of experimental conditions on vowel height (F1) and backness (F2) respectively. Since formants were extracted from five points over the course of articulating the vowel, time was coded numerically (1-5), and treated as an independent variable. This was scaled and centered prior to analysis and treated as a continuous predictor. We fitted models using the lme4 package (Bates, Mächler, Bolker, & Walker, 2015), and used lmerTest (Kuznetsova, Brockhoff, & Christensen, 2017) to obtain estimates of p values for individual variables. For all mixed models, the alpha criterion was set at $|t| > 2$. The dependent variables were normalized F1 and F2, and we fitted separate models for these. The independent variables or fixed factors, along with their levels, are given in Table 2. Subject and Item were treated as random factors.

3.1. Models

Fixed effects

Our study involved two IE vowels. Vowel is a fixed effect, since it is expected to systematically affect formant frequencies. The experimental treatment is Context, i.e., whether the English target item was presented in an English (nonswitch) or Bengali (switch/mixed) context.

Our first research question (*RQ1 – L2 vowels*) asks whether L2 vowel quality of proficient bilinguals is affected by the Context of utterance. We hypothesize that the direction of such shift in the mixed condition, if any, is towards a corresponding L1 category. Therefore, we expect the effect of Context on formant frequencies, i.e., the direction of shift, to be moderated by the vowel category (see Section 1.5, leading to our second research question *RQ2 – Asymmetry*). Thus, the main effect of interest is the interaction term Context*Vowel. This corresponds to the transfer pattern. Our third research question (*RQ3 – Paradigm*) asks whether the language-mixing paradigm, or Task, affects the extent and

Table 2. Fixed effects

Variable name	Levels
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: nonswitch), b (Bengali: switch/mixed)
Time	continuous variable, range −1.6 to 1.6, scaled and centered

[1]"Since our experimental treatments were between-subject and statistical analyses use mixed-effects regression, this step would not affect the results, as pointed out by a reviewer. This was done to allow for more intuitive values and visualizations, rather than statistical necessity. The code and analyses using non-normalized values can be found in the OSF repository. "

https://doi.org/10.1017/S1366728923000159 Published online by Cambridge University Press

direction of transfer. If this is the case, we expect Task to interact with the transfer term, giving a three way Task*Context*Vowel interaction. Since we measured formant frequencies at five points in the vowel, Time is expected to affect the formant frequency. However, the precise effect of formant dynamics is expected to depend on the vowel category. Thus, we expect the interaction of Vowel*Time to predict formant frequency. If transfer patterns change across time, then we should expect Time to interact with the transfer term, giving a Context*Vowel*Time interaction.

3.2. Vowel Height – F1

The sequence of models fitted can be found as Appendix S3 in the OSF repository. The optimal model was: f1 ~ time*vowel + vowel*context*task + time*task + (1 + context|word) + (1 + context + vowel|subject). Table 3 summarizes the model coefficient estimates (β), standard errors (S.E.), t-values, and p-values for this model. p-vales smaller than 0.05 are starred.

Model interpretation

F1 values are significantly predicted by Time in the model, reflecting the dynamic nature of articulation which causes acoustic variation over the course of time. Vowel is not a significant predictor of F1. This is initially surprising, given that the phonological category is expected to systematically affect vowel quality. However, a boxplot visualization (Figure S2) shows that the participants in this study vary greatly both in the relative heights of the two IE categories [ʌ, æ] and the extent of height difference between them. These measures are only from the nonswitch condition, suggesting that this variability is a feature of the participants' IE vowel categories, independent of the effects of switching.

The negative slope of the estimate for Context shows that F1 is lower, i.e., vowels are higher, in the nonswitch (e) condition compared to the mixed (b) condition. More intuitively, vowels are lowered in the mixed condition. However, the lack of significance indicates that using English words in mixed utterances does not lead to a uniform lowering/raising across vowels. Instead, there is an asymmetrical shift in the categories [ʌ] and [æ], as shown by the significant interaction between Vowel and Context. Specifically, the Vowel*Context interaction has a negative slope, showing that the lowering is greater in [ʌ]. The Vowel*Context*Time interaction is not significant: transfer patterns do not change across time.

Independent of experimental condition, there is a significant effect of Task: the sentence task (code-switching) leads to higher F1 values in both vowels relative to the cued picture-naming task. This is surprising, as we have no a priori reason to expect a uniform difference in vowel quality across tasks. A significant interaction with Time shows that this effect is strongest at the beginning of the vowel and decreases with time, indicating a local effect. The Vowel*Context*Task interaction is of interest, as it relates to our research question of how the test paradigm affects transfer. The significant effect of this interaction indicates a difference in transfer patterns such that transfer in the F1 (height) dimension is greater in the code-switching (s) task compared to the picture-naming (p) task.

3.3. Vowel backness – F2

The optimal model was: f2 ~ vowel*time + vowel*context*task + task*time + (1 + context|word) + (1 + context+vowel|subject). Table 4 summarizes the model coefficient estimates (β), standard errors (S.E.), t-values, and p-values for this model. p-values smaller than 0.05 are in starred.

Model interpretation

As with F1, we find that Time significantly influences F2 values as expected. Vowel is a significant predictor of F2, with [ʌ] having lower F2, i.e., being more backed, than [æ]. The negative slope of the estimate for Context indicates that on average, vowels are fronted in the mixed condition. The lack of significance for this effect shows that mixed productions do not lead to uniform fronting or backing across vowels. This aligns with our expectation for movement to be towards the corresponding L1 category, and thus differ across vowels. Instead, the significant Vowel*Context interaction indicates that backing is greater in [æ]. The lack of a significant Vowel*Context*Time interaction shows that transfer patterns do not change across time, suggesting that the underlying acoustic targets are different in the two conditions.

We see a significant effect of Task on F2: vowels in the code-switching (s) task have lower F2, i.e., are more backed than those in the picture-naming (p) task. As with F1, this is an unexpected result. The significant interaction with Time

Table 3. Optimal model: F1

	Estimate	Std. Error	t value	p value
(Intercept)	−0.053	0.179	−0.300	0.769
Time	0.139	0.012	12.040	<0.001***
vowel2	−0.147	0.223	−0.661	0.519
contexte	−0.013	0.068	−0.196	0.846
tasks	0.330	0.027	12.414	<0.001***
time:vowel2	−0.171	0.014	−12.688	<0.001***
vowel2:contexte	−0.137	0.060	−2.251	0.031*
vowel2:tasks	0.009	0.038	0.234	0.815
contexte:tasks	0.073	0.038	1.951	0.051.
time:tasks	−0.122	0.013	−9.028	<0.001***
vowel2:contexte:tasks	−0.196	0.054	−3.600	<0.001***

Table 4. Optimal model: F2

	Estimate	Std. Error	t value	p value
(Intercept)	0.232	0.221	1.05	0.304
vowel2	−0.386	0.148	−2.60	0.015*
time	0.043	0.008	5.63	<0.001***
contexte	−0.017	0.028	−0.61	0.548
tasks	−0.087	0.017	−5.00	<0.001***
vowel2:time	−0.163	0.009	−18.35	<0.001***
vowel2:contexte	0.110	0.028	4.01	<0.001***
vowel2:tasks	−0.090	0.025	−3.55	<0.001***
contexte:tasks	−0.110	0.025	−4.44	<0.001***
time:tasks	0.062	0.009	7.03	<0.001***
vowel2:contexte:tasks	0.033	0.036	0.94	0.34652

https://doi.org/10.1017/S1366728923000159 Published online by Cambridge University Press

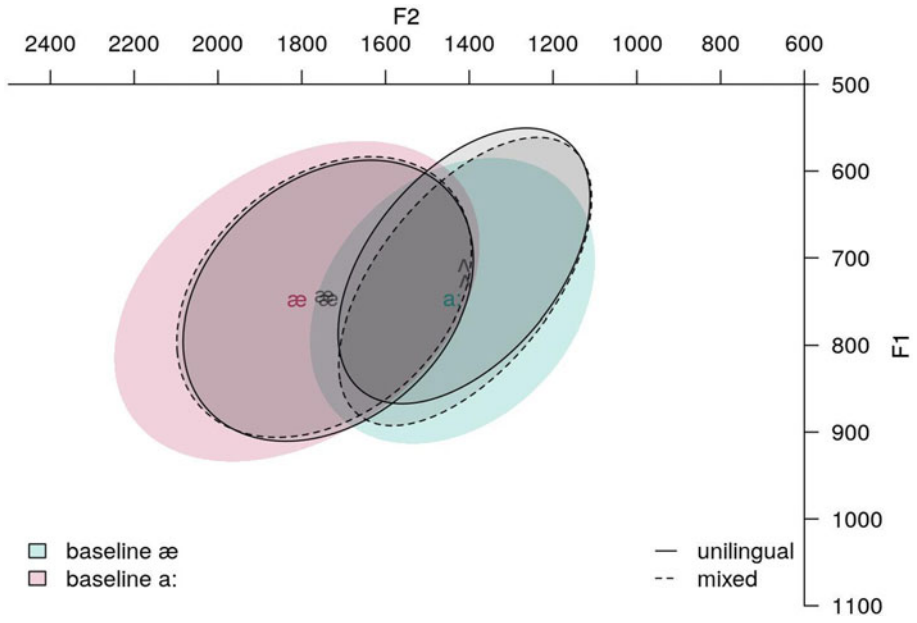


Figure 5: Vowel quality in unilingual English, mixed, and baseline Bengali productions

shows that this effect is local and decreases over time. Unlike vowel height, we do not find patterns of shift in vowel backness (F2) to differ across paradigms, as indicated by the lack of a significant effect for the Vowel*Context*Task interaction.

Statistical analyses reveal that [ʌ] is lowered and fronted, while [æ] is fronted, in the mixed condition. As discussed in section 1.5, we wanted to understand how the direction of this shift compares to corresponding Bengali categories. Figure 3 shows the nonswitched IE productions of the participants in relation to average positions of related vowel categories in Bengali. Figure 5 overlays productions in the mixed condition over Figure 3, showing F1×F2 plots for IE vowel productions in the nonswitch and mixed conditions, in comparison to the related Bengali categories from the SHRUTI corpus (Das et al., 2011). Given that the Bengali corpus data represent average vowel positions, and are not from the same set of speakers as the participants of this study, we refrained from including these in statistical analyses. However, Figure 5 provides a heuristic confirmation for our informal hypotheses in section 1.5 that the IE [æ] moves towards the corresponding [æ] category in Bengali, whereas the shift in [ʌ] is towards the related Bengali category [a:]. The following sections discuss the implications of these results.

4. Discussion

4.1. Phonetic transfer in L2 vowels

This study examined short-term phonetic transfer in L2, towards understanding language processing in proficient bilinguals. Two interrelated questions (section 1), bear on this:

(i) what causes transfer in mixed-language speech
(ii) what causes observed asymmetries between transfer in L1 vs L2

4.1.1. What causes transfer – language mode vs switching

We presented nonswitched English and code-switched Bengali–English stimuli in the same test block and found consistent differences in vowel quality. This confirms that the local act of switching triggers transfer, even in a bilingual language mode, paralleling results from Olson (2016); Piccinini and Arvaniti (2015); Tsui et al. (2019). These studies focused on a temporal feature, VOT. Our results demonstrate a similar effect in vowels. Since VOT is a one-dimensional measure, a difference across conditions is insufficient to infer whether the category has shifted towards a specific L1/L2 counterpart, or reflects a more general system-wide shift in articulatory norms. Vowel quality measures, having both directionality and magnitude, allow this. We find that productions of English vowels do not merely shift towards a Bengali counterpart, but rather towards a set of Bengali-like norms: a vowel category that doesn't exist in Bengali nonetheless moves towards a related (perceptually close) Bengali category. This parallels findings from Simonet (2014) for Catalan and Spanish (they manipulated language mode, without code-switching). The Simonet (2014) study showed that co-activating multiple languages within a discourse causes a transient increase in cross-language phonetic interaction, leading to changes in both L1 and L2. This is interpreted as follows: a bilingual mode temporarily co-activates both language systems, making both L1 and L2 representations available during processing, whose interactions lead to cross-language transfer. But what does this interaction look like, and how are these abstract representations processed during speech? Examining transfer during switching can provide insights into this.

The time course of such shifts is telling: we see a difference between conditions right from the start of articulation, suggesting different acoustic targets. This indicates that such interaction involves the speech planning stage. The pattern of transfer observed in this study, particularly for [ʌ], also demonstrates that beyond a simple L2-L1 mapping, transfer during online processing makes reference to entities above the level of individual sounds, demonstrating system-level interaction: a sound that does not exist in the L1 repertoire nonetheless acquires a different target, and the target is related to the L1 system in a meaningful way. This cannot be reasonably explained as articulatory inertia carried over from uttering L1 sounds. It is also difficult to account for in a purely exemplar framework, as the discourse does not contain any instances of "L1 [ʌ]" that could affect the L2 category. If representations are combined in creative ways during online

https://doi.org/10.1017/S1366728923000159 Published online by Cambridge University Press

speech processing, then can we think of intermediate representations not just for individual sounds, but rather for sound systems? While more detailed empirical data is needed to answer such questions, this approach aligns with sociophonetic understandings of bilingual phonetic variation, e.g., as discussed by Khattab (2009, 2013) based on the use of phonological variables by Arabic–English bilingual children. These visualize the bilingual phonetic system as "a rich store of phonological representations", comprising both language-internal and cross-language variation, from which the bilingual speaker is constantly choosing, guided by a variety of pressures. We propose that vowels are a good feature to examine the plausibility of such systems as cognitive models for language processing: in addition to continuous measures (unlike phonological variables), they also provide two-dimensional data (unlike VOT) that is suited to further test and refine predictions from such a model.

4.1.2. Differences between L1 vs L2 transfer

Existing studies on language-switching have often reported phonetic transfer from L2 to L1, no parallel transfer from L1 to L2. One postulated model consistent with this finding is the Inhibitory Control Model (ICM) (Green, 1986; Olson, 2013; Tsui et al., 2019). As discussed in section 1.2, vowels provide a fertile testing ground for this theory at the level of sounds. Both vowels in this study showed systematic shifts in the mixed-language condition compared to baseline nonswitched productions. In the context of an ICM-based analysis of between-language asymmetries in phonetic transfer, as proposed by Olson (2013), this demonstrates that the postulated lower switch cost on L2 does not entirely mask transfer effects on L2 vowels. There are a few possible explanations for this. The present study focuses solely on L1 (Bengali) to L2 (English) transfer, and not the converse. It is theoretically possible that these participants are L2-dominant. That would mean that switching from L1 into L2 induces a greater switch cost, leading to transfer effects on switched L2 productions. We should then expect a lack of transfer while switching from English to Bengali. While more empirical data could clarify this, the language background of these participants makes this scenario unlikely. All individuals report regularly using both English and Bengali in a variety of settings, and feeling equally comfortable using both languages. Listening and speaking skills are expected to be most relevant to phonetic transfer in daily language use, and participants' self-ratings for these skills are at ceiling for both languages. One point to note is that participants on average rated their reading proficiency in Bengali to be lower than in English. Given that the stimuli in this task were presented visually, differences in reading proficiency are a potential source of explanation. However, anecdotal observation suggests that phonetic transfer in L2 productions is common in daily language use too, even in the absence of reading. This raises questions about an ICM-based analysis of cross-language asymmetries in phonetic transfer that future research should investigate.

Moreover, our findings differ from Tsui et al. (2019), who report on the basis of VOT measurements that balanced bilinguals do not evidence an effect of language-switching. Tsui et al. attribute this to greater language-switching experience in these participants, leading to lower switch-costs in both languages. As noted in sec. 2.1, participants in this study are highly proficient bilinguals with extensive experience of language switching. Moreover, their environment constantly exposes them to mixed-language stimuli: we expect that they are habituated to drawing on multiple language systems during online language processing. They nonetheless demonstrate cross-language transfer effects in L2 vowel production. There are a few different ways to approach this:

One possible explanation is that vowels are generally more susceptible to phonetic transfer than temporal features such as VOT. Following from the discussion in sec. 1.4, it might therefore be fruitful for future studies to focus on vowels as the site for transfer. More generally, these findings point to a need for expanding the range of features studied, in order to make robust generalizations about the mechanisms underlying observed patterns of transfer. This is an example where examining different features of speech (e.g., VOT vs. vowel quality) allows us to compare predictions from different underlying process models. Measuring multiple features within a single study promises to be a powerful tool for interpretation and theorizing, at little added cost to researchers. This would sidestep the unavoidable confounds that come with comparing results across studies.

Beyond methodological considerations, however, we argue that these differences ultimately raise broader questions about a purely inhibition-based understanding of transfer. Given the compelling body of literature on the socio-communicative significance of phonetic variation along the L1-L2 continuum (e.g., Agnihotri, 1979; Khattab, 2013; Stuart-Smith et al., 2011), is it wise to think of transfer as resulting from an inability to suppress and separate language systems? Seeing as even young children are capable of using minute phonetic variation in sophisticated ways, it might be more intuitive to think of transfer as a communication-enabling feature of language, even at a purely cognitive level. Research on the cognitive mechanisms underlying cross-language interaction and variation in bilingual speech is still an emerging and exciting line of inquiry. We believe that a communication-centered approach to bilingual variation is essential to understand how bilingual language processing functions in the real world. To this end, highly proficient native multilingual populations serve as an important source of data.

4.2. Asymmetries in Phonetic Transfer

Our results indicate differences across the two vowels in both magnitude and direction of shift. Overall, shift is greater in [ʌ]. Recall (Figure 2) that the category [æ] is in a denser part of the IE vowel space than [ʌ]. Specifically: [æ] is already at the lowermost part of the vowel space. Large movements upwards in the F1 (vowel height) dimension would risk a conflation of the category with the neighboring [ɛ]. In contrast, [ʌ] (indicated as [ə] in the chart) is in the central part of the vowel space and not flanked by any immediate neighboring categories. Thus, the difference in magnitude of transfer aligns with our hypothesis *H2 – Asymmetry*, that the extent of shift in individual categories is constrained by the linguistic principle of contrast preservation. An alternative way to interpret this result is that the initial distance between IE [ʌ] and its target L1 category [aː] was greater than that between the IE and Bengali [æ] categories, leading to a greater visible shift in the former. However, our results find parallels in previous studies with VOT that have consistently observed greater shifts in the long-lag VOT language (greater acceptable range of VOT) compared to the short-lag language (e.g., Antoniou et al., 2011; Bullock & Toribio, 2009). This suggests that a general principle of "room for movement" does play a role in moderating transfer. These results suggest that phonetic transfer (i) makes reference to system-level representations that go beyond individual sounds; and (ii) respects linguistic principles.

https://doi.org/10.1017/S1366728923000159 Published online by Cambridge University Press

This has implications both for our understanding of phonetic transfer as a phenomenon (e.g., which mental representations and at what level of processing are involved?) and language as a system (e.g., how have transfer phenomena shaped phonological systems in multilingual settings, where languages have co-existed for long periods of time; and what are the linguistic implications of increasingly cosmopolitan urban spaces and identities?).

Our results also indicate asymmetry in the direction of shift: the category [ʌ] is lowered and fronted in the mixed language condition, while [æ] is fronted. On the basis of previous literature indicating that both long-term and short-term shifts in categories of multilingual speakers are targeted with respect to the other language, we hypothesized that participants' IE [æ] would shift towards the Bengali [æ], and [ʌ] towards the Bengali [a:]. As Figure 5 indicates, comparing transfer patterns with Bengali categories from Das et al. (2011) suggests that this is borne out.

4.3. Switching paradigm

Olson (2013) posits that in both scripted and spontaneous code-switching, "discursive properties may partially drive the production of code-switches, masking underlying effects of interaction", proposing cued picture-naming as an alternative. We tested whether this might independently affect the outcome of transfer, by comparing productions from the same participants in these two tasks. The direction of transfer did not change across paradigms. However, there was a difference in the magnitude of transfer: in the F1 dimension, the extent of phonetic transfer is greater in the code-switching task compared to the cued picture-naming task. This asymmetry is not seen in the F2 dimension. It is possible that code-switching induces a greater degree of language co-activation because switches are being made within an utterance, which requires planning: participants see the entire sentence and then read it aloud. Compared to this, the cued picture-naming paradigm presents the language cue immediately prior to the target word, inducing extremely "local" switches and thus a smaller degree of language co-activation. While both of these are local switches and take place in a bilingual language mode, it is possible that this minute difference in time-frame induces slightly different degrees of cross-language phonetic interaction. This aligns with the findings of Bullock and Toribio (2009), who found evidence of transfer in pre-switch tokens, suggesting planning effects. In our study specifically, orthography was another factor that might have affected this. In the picture-naming task, participants saw one word at a time. The CS task in contrast presented mixed utterances in mixed orthography. It is possible that seeing mixed orthography, particularly for individuals accustomed to processing mixed-language visual input in their surroundings, triggered increased language co-activation, leading to greater cross-language interaction.

We do not find any differences in the pattern of transfer across paradigms. Our study used a frame sentence to induce code switching, and not all the items led to pragmatically meaningful sentences. Thus, one possible explanation for the overall lack of asymmetry 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 clarify this. The results of this study indicate that for the purposes of observing dynamic phonetic transfer in an experimental setup, elicitation through cued-picture-naming vs. (controlled) code-switching does not significantly alter the behavior under study. However, elicitation through code-switching allows for more planning, potentially increasing the magnitude of observable effects.

An unexpected result in this study is that the paradigm directly affected vowel quality: all vowels were more backed and lowered in the code-switching task. There is no principled reason to expect this. A closer look at the dynamics of this effect indicate that it is greatest at the beginning of articulation, and decreases with time. Given that it is localized at the beginning of the target vowel, it is likely an artifact of the particular stimuli used here. While intriguing, a more detailed investigation of this effect is beyond the scope of the present study.

Although the findings of this study are interesting, we acknowledge that it has a number of limitations, including

(i) a small sample size;
(ii) a frame sentence for eliciting code-switching such that not all the sentences were pragmatically well-formed – this could have blocked some of the discursive effects that might result from speech planning in ordinary connected speech, thus obscuring differences between the two paradigms under study;
(iii) focus on a single acoustic feature;
(iv) the absence of unilingual Bengali data from the same set of speakers as the participants of the study, which would have allowed us to compare each individual's transfer patterns to their own Bengali productions.

Future work should address these. Our results also raise a number of open questions for future research, particularly concerning paradigm differences and between-sound asymmetries in phonetic transfer.

5. Conclusion

This study investigated how the sound systems of highly proficient bilingual speakers interact during online speech processing, by investigating the phonetic outcome of mixed-language production in speakers of Bengali and Indian English. We examined the spectral quality of L2 vowel categories [ʌ] and [æ] produced in two experimental conditions: nonswitch (English) and switch/mixed (switching from Bengali to English). Comparisons of F1 and F2 across conditions reveal that mixed language processing temporarily increases phonetic interaction between the languages, causing the English [æ] of participants to converge towards the Bengali category [æ], and [ʌ] to converge towards a related L1 category [a:]. The extent and direction of shift is asymmetrical across the two categories, which parallels findings in prior studies. This asymmetry appears to be mediated by the language-specific distribution of vowel categories. These findings evidence an effect of phonetic transfer on L2 vowels, and raise questions for proposed inhibition-based accounts of cross-language phonetic inter-action. Mixed productions were elicited in two switching paradigms – code-switching and cued picture-naming – to verify whether the postulated discursive elements in code-switching independently influence the outcome of phonetic transfer. Results suggest that in an experimental setup, paradigm does not significantly alter the behavior under study. However, the magnitude of observed transfer is greater in the code-switching paradigm. These findings add to empirical data, highlight the need for diversifying the set of language pairs and populations in transfer studies, and encourage more within-study comparisons of paradigms and acoustic features, towards establishing better

https://doi.org/10.1017/S1366728923000159 Published online by Cambridge University Press

methodological norms. Ultimately, these results contribute to ongoing discussion about the implications of short-term phonetic interaction as a window to understanding language as a mental and social system.

Acknowledgments. We are grateful to Maumita Bhowmik and Anannya Mondal for their help with data collection and processing, and to all the participants of the study. We also thank the anonymous reviewers at BLC for their detailed feedback and engagement with the contents of this paper. This study was conducted at the English and Foreign Languages University (EFLU), Hyderabad.

Abbreviations. CLF = classifier.

Data availability. The materials, data files, and analysis code for this study are openly available in OSF at: https://doi.org/10.17605/OSF.IO/DSB2X.

Supplementary material. All supplementary material accompanying this paper can be found at https://doi.org/10.1017/S1366728923000159.

List of supplementary materials

Figure S1: Acoustic data before and after gender normalization
Figure S2: Speaker-wise mean F1 in non-switched English productions
Appendix S1: Linguistic situation in India and detailed language background of participants
Appendix S2: Discussion on power analysis and sample size in the study
Appendix S3: Sequence of statistical models, outputs, and determining the optimal model

Competing interests. The authors declare none.

References

Agnihotri RK (1979) *Processes of assimilation: A sociolinguistic study of Sikh children in Leeds*. (Unpublished doctoral dissertation). University of York.

Antoniou M, Best CT, Tyler MD and Kroos C (2011) Inter-language interference in VOT production by L2-dominant bilinguals: Asymmetries in phonetic code-switching. *Journal of Phonetics* **39**, 558–570.

Babel M (2010) Dialect divergence and convergence in New Zealand English. *Language in Society*, 437–456.

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

Balukas C and Koops C (2015) Spanish-English bilingual voice onset time in spontaneous code-switching. *International Journal of Bilingualism* **19**, 423–443.

Bansal RK (1969) *The intelligibility of Indian English: Measurements of the intelligibility of connected speech, and sentence and word material, presented to listeners of different nationalities*. Central Institute of English;[available from Orient Longmans, Madras].

Bates D, Mächler M, Bolker B and Walker S (2015) Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* **67**, 1–48. doi: 10.18637/jss.v067.i01

Best CT and Tyler M (2007) Nonnative and second-language speech perception. In OS Bohn and MJ Munro (Eds.), *Language experience in second language speech learning: In honour of James Emil Flege* (pp. 13–34).

Birdsong D, Gertken L and Amengual M (2012) *Bilingual language profile: An easy-to-use instrument to assess bilingualism*. COERLL, University of Texas at Austin. web. 20 jan. 2012.

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

Bohn OS and Flege JE (1992) The production of new and similar vowels by adult German learners of English. *Studies in Second Language Acquisition*, 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**, 217–243.

Bullock BE and Toribio AJ (2009) Trying to hit a moving target: On the sociophonetics of code-switching. In K. D. B. Ludmila Isurin Donald Winford (Ed.), *Multidisciplinary approaches to code switching* (Vol. 41). John Benjamins Publishing Amsterdam, The Netherlands, pp. 189–206.

Caramazza A, Yeni-Komshian GH, Zurif EB 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**, 421–428.

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

Das B, Mandal S and Mitra P (2011) Bengali speech corpus for continuous auutomatic speech recognition system. In 2011 *international conference on speech database and assessments* (oriental cocosda) (pp. 51–55).

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**, 29.

Flege JE (1987) The production of "new" and "similar" phones in a foreign language: Evidence for the effect of equivalence classification. *Journal of Phonetics* **15**, 47–65.

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

Flege JE (2007) Language contact in bilingualism: Phonetic system interactions. *Laboratory phonology* **9**, 353–382.

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

Flege JE, MacKay IR and Piske T (2002) Assessing bilingual dominance. *Applied Psycholinguistics* **23**, 567–598.

Fowler CA, Sramko V, Ostry DJ, Rowland SA and Hallé P (2008) Cross language phonetic influences on the speech of French–English bilinguals. *Journal of Phonetics* **36**, 649–663.

Ghosh S (2016) Vowel sound change in progress: A cross-generational study of navaron dialect. *International Journal of English: Literature, Language Skills IJELLS* **5**, 2278–742.

Goldrick M, Runnqvist E and Costa A (2014) Language switching makes pronunciation less nativelike. *Psychological Science* **25**, 1031–1036.

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

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

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

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

Grosjean F and Miller JL (1994) Going in and out of languages: An example of bilingual flexibility. *Psychological Science* **5**, 201–206.

Guion SG (2003) The vowel systems of Quichua-Spanish bilinguals. *Phonetica* **60**, 98–128.

Hickey R (2005) *Legacies of colonial English: Studies in transported dialects*. Cambridge University Press.

Kachru BB (1983) *The indianization of English: the English language in India*. Oxford University Press.

Kachru BB and Smith L (1981) *English for cross-cultural communication*. London: Macmillan Press.

Khattab G (2009) Phonetic accommodation in children's code-switching. In BE Bullock and AJ Toribio (Eds.), *The Cambridge Handbook of Linguistic Code-Switching*. Cambridge University Press, p. 142– 160. doi: 10.1017/CBO9780511576331.010

Khattab G (2013) Phonetic convergence and divergence strategies in English-Arabic bilingual children. *Linguistics* **51**, 439–472.

Kuznetsova A, Brockhoff PB and Christensen RHB (2017) lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* **82**, 1–26. doi: 10.18637/jss.v082.i13

MacLeod AA 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**, 400–421.

Marian V, Blumenfeld HK and Kaushanskaya M (2007) *The language experience and proficiency questionnaire (leap-q): Assessing language profiles in bilinguals and multilinguals*.

Masica C (1972) *The Sound System of Indian English*. Monograph No. 7.

https://doi.org/10.1017/S1366728923000159 Published online by Cambridge University Press

McCloy DR (2016) phonr: tools for phoneticians and phonologists [Computer software manual]. (R package version 1.0-7)

Muldner K, Hoiting L, Sanger L, Blumenfeld L and Toivonen I (2019) The phonetics of code-switched vowels. *International Journal of Bilingualism* **23**, 37–52.

Myers-Scotton C (1993) *Dueling languages: Grammatical structure in code-switching*. Clarendon Press.

Nihalani P, Tongue RK, Hosali P and Crowther J (1979) *Indian and British English:A handbook of usage and pronunciation*. Oxford University Press Delhi.

Office of the Registrar General & Census Commissioner, India (2011) *Census of India 2011*. [Web Archive] Retrieved from https://censusindia.gov.in/census.website/

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

Olson DJ (2016) The role of code-switching and language context in bilingual phonetic transfer. *Journal of the International Phonetic Association* **46**, 263–285.

Pandey P (2015) Indian English Pronunciation. In *The handbook of English pronunciation*. Wiley Online Library, pp. 301–319.

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 JS (2010) Expressing oneself in conversational interaction. In E Morsella (Ed.), *Expressing oneself/expressing one's self: Communication, cognition, language, and identity*. London: Taylor & Francis, pp. 183–196.

Piccinini P and Arvaniti A (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 [Computer software manual]*. Vienna, Austria. Retrieved from http://www.R-project.org/

Rodriguez-Fornells A, Kramer U, Lorenzo-Seva U, Festman J and Münte TF (2012) Self-assessment of individual differences in language switching. *Frontiers in Psychology* **2**, 388.

Rubino CRG and Garry J (2001) *Facts about the world's languages: an encyclopedia of the world's major languages, past and present*. H.W. Wilson, New York.

Sancier ML and Fowler CA (1997) Gestural drift in a bilingual speaker of Brazilian Portuguese and English. *Journal of Phonetics* **25**, 421–436.

Schwartz G, Balas A and Rojczyk A (2015) Language mode vs. L2 interference: Evidence from L1 Polish. In *Proceedings of the International Congress of Phonetic Sciences (ICPhS)*. Glasgow, UK.

Šimáčková Š. and Podlipskỳ VJ (2015) Immediate phonetic interference in code-switching and interpreting. In *Proceedings of the 18th International Congress of Phonetic Sciences (ICPhS)*. Glasgow, UK.

Šimáčková Š. and Podlipskỳ VJ (2018) Patterns of short-term phonetic interference in bilingual speech. *Languages* **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**, 208–221.

Sirsa H and Redford MA (2013) The effects of native language on Indian English sounds and timing patterns. *Journal of Phonetics* **41**, 393–406.

Stuart-Smith J, Timmins C, Alam F, Gregersen F, Parrot J and Quist P (2011) *Hybridity and ethnic accents*. Selected Papers ICLaVE Amsterdam: Benjamins, **5**, 43–57.

Tobin SJ, Nam H and Fowler CA (2017) Phonetic drift in Spanish-English bilinguals: Experiment and a self-organizing model. *Journal of Phonetics* **65**, 45–59.

Tollefson JW and Tsui AB (2014) Language diversity and language policy in educational access and equity. *Review of Research in Education* **38**, 189–214.

Tsui R. K.-Y., Tong X and Chan CSK (2019) Impact of language dominance on phonetic transfer in Cantonese–English bilingual language switching. *Applied Psycholinguistics* **40**, 29–58.

Wells JC and Wells JC (1982) *Accents of English*: Volume 1 (Vol. 1). Cambridge University Press.

https://doi.org/10.1017/S1366728923000159 Published online by Cambridge University Press