ALMOST [w]anishing: THE ELUSIVE /v/-/w/ CONTRAST IN EDUCATED INDIAN ENGLISH

Robert Fuchs

University of Hamburg, Germany
Robert.fuchs@uni-hamburg.de

ABSTRACT

English is an institutionalised second language in India used by at least 23% of the population. In a process common to postcolonial Englishes, it is undergoing a process of standardisation that has not yet led to codification but has given rise to identifiable prestige varieties. Addressing the ongoing debate over whether the /v/-/w/ contrast is maintained in this prestige variety, this study analyses the speech of 20 speakers of Indian and 10 of British English with regard to their production of /v/ and /w/.

The acoustic analysis, based on spectral centroid, F2 and F3, reveals, for Indian English, (1) no difference in spectral centroid and (2) a small difference in F2 and F3, which is associated with slightly more pronounced lip rounding and velar constriction in /w/ than in /v/. This difference might be too small for reliable perception, especially in comparison with the British control group, where the phonemes differ markedly in spectral centroid, F2 and F3.

Keywords: Indian English, postcolonial Englishes, bilabial consonants, spectral centroid, standardisation

1. INTRODUCTION

While the sociophonetic study of vowels is widespread, there is comparatively little work on variation in consonants [33]. The present analysis addresses this gap by focussing on the sounds transcribed phonemically as /v/ and /w/, which are a labio-dental fricative and a labio-velar approximant, respectively, in British English (BrE). The study (1) determines which acoustic cues distinguish /v/-/w/ in BrE and (2) whether these cues also distinguish /v/-/w/ in Indian English (IndE).

The /v/-/w/ contrast is a particularly interesting case because it presumably relies on several acoustic cues. The identity of these cues and how they are influenced by phonological context is not well understood for either IndE or BrE, although previous research on other languages as well as American English suggests that acoustic correlates of lip-rounding and frication might play an important role [7, 27, 32]. In IndE, the status of /v/-/w/ is disputed and some argue that it has undergone a merger [30].

IndE is a postcolonial variety of English that has its roots in the colonial rule of the British Empire over India. Despite post-independence attempts to reduce the local importance of English, it continues to be widely used as a Lingua Franca in this multilingual country and enjoys an important role in government, administration, the courts, the economy and education. As is the case with many postcolonial Englishes, IndE is not spoken by all Indians. Around 23 % of the population have at least basic knowledge of English and 4 % are fluent [11], which suggests that there were around 50 million fluent speakers at the time of the 2010 census and very likely even more today.

A catch-all term such as IndE obscures the considerable variation that exists along variables such as education, age of acquisition of English and first language [15, 19]. However, it is possible to identify a subvariety of IndE that locally enjoys overt prestige and is used in education and presented as a target for students to strive towards [5, 15].

While this nascent standard IndE still lacks full official recognition, it is the de facto standard taught in schools and universities [23]. However, this standard has not yet been fully codified, and in such a context the kind of language used by educated speakers can be used to determine what is considered acceptable by the speech community [34].

The phonology of IndE differs in a number of respects from BrE [1, 12, 14-16, 19-21, 30], likely in part due to historical transfer from Indian languages [30]. In spite of phonological differences between Indian languages, the phonology of Educated IndE (as opposed to basi- and mesolectal varieties) is relatively homogeneous, regardless of the first languages (L1s) used by particular speakers [30].

Moreover, IndE is an appropriate choice for a case study into postcolonial Englishes. Compared to other varieties, its phonology is relatively well studied [acoustic studies include 12-16, 18-22, 31, 36], perhaps due to its large number of speakers and worldwide iconic and ambivalent language-ideological status [8]. Nevertheless, there is a lack of acoustic studies on IndE that take into account various first language (L1) and socio-economic backgrounds, rely on a greater number of informants and take into account at what age and where informants learned English [30]. Particularly the acoustic realisation of complex consonantal contrasts (defined here as involving

several acoustic dimensions) such as /v/-/w/ has not been studied in detail, neither for IndE, nor to any great extent in sociophonetics in general [33].

2. THE /v/-/w/ CONTRAST

2.1. The /v/-/w/ contrast in British English

The labio-dental fricative /v/ and the labio-velar approximant /w/ are distinct phonemes in BrE. Both /v/ and /w/ occur in syllable onsets, but only /v/ in codas. The acoustic cues to this contrast in BrE have not been investigated in detail.

However, research on other languages and American English suggests that there might be several acoustic cues. As a fricative, /v/ has more frication and more energy in higher frequencies than /w/, which can be measured by the spectral centroid [7, 27]. For /w/, lip-rounding and the velar constriction lower the second formant (F2). Lip-rounding also lowers the third formant (F3), which is however raised again by the velar constriction, so that /w/ has an average F3 compared to vowels [32].

2.2. The /v/-/w/ contrast in Indian English

In IndE, /v/ and /w/ are widely considered to have merged into a labio-dental approximant /ʋ/ [4, 8], sometimes also described as a frictionless continuant [5]. This generalisation might not account for sociolinguistic variation within IndE, as Sailaja [29] suggested that only non-standard varieties of IndE have a complete merger. Standard (or educated) IndE, on the other hand, was said to have a phonemic distinction between /v/ (realised as the labio-dental approximant [ʋ]) and /w/ (realised as [w]).

Empirical studies on the /v/-/w/ contrast in IndE are rare and restricted to impressionistic evidence [10, 28, 36]. These studies indicate that (1) some speakers might maintain the contrast, (2) that /w/ is commonly realised as [ʋ] and (3) /v/ as [v]. This supports Sailaja's [29] contention that Educated IndE might not have a /v/-/w/ merger. In order to resolve the debate, more - in particular acoustic - research is needed to determine the status of /v/ and /w/ in Educated IndE.

2.3. /v/ and /w/ in Indian languages

It is conceivable that IndE speakers' L1s have some influence on how /v/ and /w/ are realised, as previous studies found limited L1-based differences in the phonology of IndE, although L1-background turns out to be much less important than sometimes assumed [31]. The linguistic landscape of India is diverse but also shows many signs of convergence in a common linguistic area or sprachbund.

The most widely spoken languages in Northern India are the Indo-Aryan languages Hindi and Bengali. Dravidian languages are mostly spoken in the South, with Telugu and Malayalam being the most widely spoken languages of the largest subfamilies [17]. It is these L1 backgrounds that the present study focusses on. More than 50% of the population speak one of these languages [9], which lack a phonemic /v/-/w/ contrast, and are described as having only a single bilabial approximant /ʋ/ [3, 6]. This would suggest that L1-based differences are unlikely in the case of IndE /v/-/w/. However, given (1) the contradictory reports of the status of /v/-/w/ in IndE and (2) conjectures of regional/L1-basd variation in /v/-/w/ realisation [35], the present study will control for L1 background as a potential confounding factor.

3. AIMS AND METHODS

3.1. Aims

Given the controversial status of the /v/-/w/ contrast in Educated IndE, this study will provide an acoustic analysis of these speech sounds in the read and spontaneous speech of 20 speakers of IndE as well as, for comparison, 10 speakers of BrE. The following questions will be investigated:

1. What are the acoustic cues of /v/-/w/ in BrE?
2. Does Educated IndE have a phonemic distinction between /v/ and /w/?
3. Does phonological context (preceding/ following V or C) favour realisation of /w/ as [w] or [v], and of /v/ as [w] or [v]?
4. Do speakers with different L1s differ in the patterns described under (1) and (2)?

3.2. Data

Recordings of a text read by 10 speakers of Standard Southern BrE (all male) and 20 speakers of IndE (11 male, 9 female) were used. In addition, spontaneous data was collected in a semi-structured interview. The BrE data was taken from the DyViS database [25], and the IndE speakers were recorded reading the same text. All speakers were university students at the time of recording.

The IndE speakers (aged 20-28) were equally divided between four different L1 groups, i.e. Hindi, Bengali, Telugu or Malayalam. With the exception of one speaker each, they had exclusively attended English-medium schools and universities, and had not resided outside of India.

3.3. Methods

The present analysis is part of a larger project that studies the acoustic properties of /v/ and /w/ in all

Figure 1: Energy concentration in the spectrum (spectral centroid) of /v/ and /w/ in BrE and four IndE L1 backgrounds (Bengali, Hindi, Malayalam, Telugu) in post-consonantal and post-vocalic position.

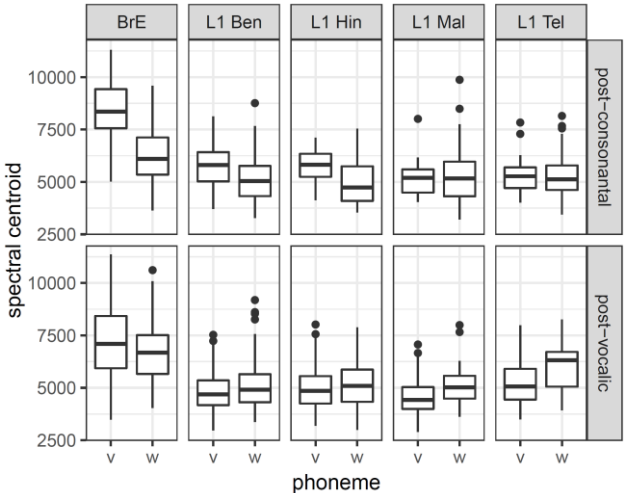
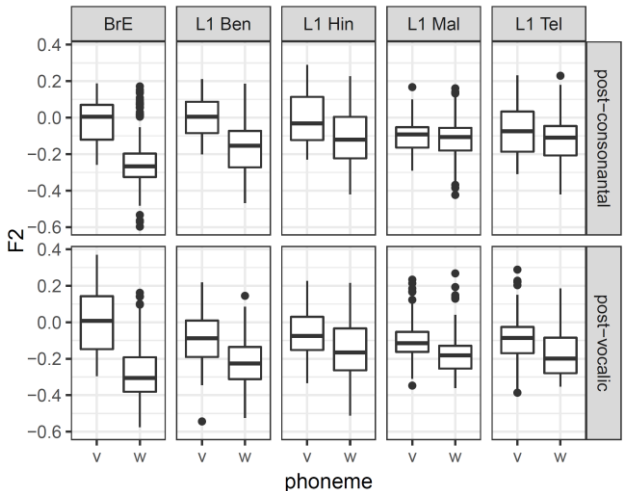


Figure 2: Lip-rounding and velar constriction (normalised F2) of /v/ and /w/ in BrE and four IndE L1 backgrounds (Bengali, Hindi, Malayalam, Telugu) in post-consonantal and post-vocalic position.



positions in BrE and IndE. Here, only results for syllable onsets are reported. The analysis distinguishes between post-consonantal and -vocalic contexts.

In total, 1,115 occurrences of /v/ and 1,642 of /w/ in the reading task, and 664 occurrences of /v/ and 1,358 of /w/ in the spontaneous condition were analysed. In order to normalise formant measurements for individual speakers, 11,613 /i/, 8,778 /u/, 5,364 /m/ occurrences were analysed.

The following acoustic criteria were used to determine where on a scale between a labio-velar approximant [w] and a labio-dental fricative [v] an individual /w/ or /v/ token is situated:

(1) Spectral centroid above 1,500 Hz (SC): The frequency with the highest intensity in the spectrum, with a high value for [v] and a low value for [w] (high-pass filtered at 1,500 Hz, with 100 Hz smoothing; cf. [7, 28]).

(2) Normalised second formant (F2n): The normalised frequency of the second formant as a measure of lip-rounding and velar constriction, where a lower F2n indicates more lip-rounding and a greater velar constriction. As formants are only meaningful for sonorants, this measure was only calculated for the more sonorant part of the data, defined as those instances that have a spectral centroid typical of BrE /w/. Normalisation for differences in vocal tract size was achieved by calculating, for each /w/ and /v/ token, its relative difference with the grand mean of the second formant of all /i/, /u/ and /m/ phonemes produced by that speaker:

$$F2_n(w) = -\frac{MEAN\_F2 - F2(w)}{MEAN\_F2}$$

(where $F2_n(w)$ is the normalised second formant of a phoneme, F2(w) is the acoustic second formant of a phoneme in Bark, and MEAN_F2 is the grand mean of the second formant of all /i,u,m/ phonemes uttered by the speaker in question)

(3) Normalised third formant (F3n): The normalised frequency of the third formant as a measure of lip-rounding and velar constriction, where a lower F3n indicates the presence of lip-rounding without a velar constriction. Computed in analogy to F2n.

After extracting these acoustic measures from the recordings with a Praat script, mixed-effects models were computed in R [26]. Dependent variables in the analysis were acoustic measures (1)-(3). The independent variables that were considered are SPEAKING STYLE (read vs. spont.) and PHONOLOGICAL CONTEXT/PHONEME (/v/ and /w/, prec./foll. context). For the dependent variables SECONDFORMANT and THIRDFORMANT, rounding of preceding and following vowels was additionally considered as an independent variable. SPEAKER and WORD were used as random factors. In addition to these variables, regression models included either the independent variable VARIETY (BrE/IndE) or L1 (BrE/L1 Bengali/L1 Hindi/L1 Malayalam/L1 Telugu), in order to determine which of these accounted better for the data. Model selection was based on minimisation of the Bayesian Information Criterion [2].

4. RESULTS

4.1. Energy concentration in the spectrum (spectral centroid)

Overall, the results indicate that BrE has a significantly higher spectral centroid (SC) for /v/ than for /w/, whereas this tends not to be the case for IndE. After consonants (see Fig. 1, top row), /v/ has a significantly higher SC than /w/ in BrE (df=1593, t=12.9, p<0.0001), whereas the difference is not significant in any of the IndE L1 groups. In the L1-Telugu and Malayalam groups, the size of the difference

is negligible, while it is slightly larger in the L1-Hindi and Bengali groups. The differences between the L1 groups were not significant.

The post-vocalic context (see Fig. 1, bottom row) reveals a slightly different pattern. In BrE, /v/ has a somewhat higher SC than /w/, but the difference is not significant. By contrast, in all IndE L1 groups, /v/ has an (insignificantly) lower SC than /w/ in this context.

Apart from the interaction discussed above, the analysis also suggests an interaction between SPEAKING STYLE and L1. This interaction (which was included in the analysis as per the BIC criterion) did not reveal any significant differences, suggesting that there is no notable variation in speaking style.

In summary, the analysis of SC reveals that /v/ and /w/ do not differ significantly in degree of frication in IndE. By contrast, /v/ has more frication in BrE compared to /w/, but the difference is only significant in post-consonantal position.

4.2. Lip-rounding and velar constriction (F2/F3)

This part of the analysis was restricted to approximant-like occurrences, defined as a low SC (see section 4.1). This criterion was met by 630 (99.7%) IndE /w/ tokens, 321 (99.7%) BrE /w/ tokens, 997 (83.0%) IndE /v/ tokens, and 300 (54.5%) BrE /v/ tokens. For SECONDFORMANT, the independent variables PHONOLOGICAL CONTEXT/PHONEME (without rounding of preceding or following vowels) and VARIETY, as well as an interaction between them, were selected in the final model.

Results reveal that both BrE and IndE speakers produce /v/ with a higher F2n than /w/. In the post-consonantal context (see Fig. 2, top panel), /v/ has a significantly higher F2n than /w/ in both BrE ($df=1151$, $t=5.3$, $p<0.0001$) and IndE ($df=1227$, $t=5.2$, $p<0.0001$). Next, in the post-vocalic context (Fig. 2, bottom panel), BrE /v/ has a significantly higher F2n than BrE /w/ ($df=1626$, $t=7.3$, $p<0.0001$). In IndE, the difference is significant, too, but much smaller ($df=1428$, $t=7.3$, $p<0.0001$). Finally, speakers of Hindi and Bengali have a slightly larger difference in F2n between /v/ and /w/, compared to the two other L1-backgrounds, but these L1-based differences are not significant.

Broadly, IndE /v/ and /w/ are intermediate in F2 between BrE /v/ and /w/; median values for IndE /v/ are lower than for BrE /v/, and median values for IndE /w/ are higher than for BrE /w/, regardless of phonological context. This means that, notwithstanding significant differences between IndE /v/ and /w/ in F2n, the size of the difference between the two sounds is much smaller in IndE than in BrE.

For reasons of space, the results for F3n can only be sketched out here. BrE has a significantly lower F3n for /w/ compared to /v/ after consonants and unrounded vowels, whereas the difference is not significant after rounded vowels. In IndE, the difference is significant after consonants, but not after vowels, regardless of roundedness. Even where the difference is significant, it is much smaller in IndE than in BrE.

5. DISCUSSION AND CONCLUSION

This aim of this paper was to present a sociophonetic analysis of the /v/-/w/ contrast in IndE, subsequent to the determination of what acoustic cues are used in BrE to distinguish these sounds. Results indicate that, in BrE, /w/ is distinguished from /v/ by its nature as a fricative (lower spectral centroid) and the presence of lip-rounding and a velar constriction.

In IndE, similar acoustic cues are used and differences between /v/ and /w/ are statistically significant in some phonological contexts. However, even where they are, the difference is much smaller than in BrE, and this result holds across the four L1 backgrounds studied here. This adds to the existing evidence suggesting that L1-based variation in Educated IndE is, where it exists at all, limited [19, 31]. L1-based variation within Educated IndE appears to be smaller than differences between BrE and IndE as a whole.

The sound around which /v/ and /w/ converge in Educated IndE appears to be best described as the labio-dental approximant [ʋ]. The amount of frication involved in the production of IndE tokens intended as /v/ is smaller than even in BrE productions of [w], let alone [v], so that it must be described as an approximant on a phonetic level. Further, both instances intended as /v/ and /w/ in IndE appear to involve some lip-rounding, especially in post-vocalic position, further confirming that [ʋ] might be an accurate phonetic description of IndE /v/ and /w/. Contextual variation in the form of different realisations of /v/ and /w/ following vowels and consonants seems to be caused at least in part by coarticulatory effects in the form of greater amounts of frication following consonants.

Overall, the results suggest that the /v/-/w/ contrast might constitute a near-merger in IndE, a result that would account for the conflicting views present in the literature. In order to confirm this conclusion, in future work the author will (1) widen the acoustic basis of the analysis with a further measure of 'fricativeness' (Zero Crossing Rate) and (2) employ Support Vector Machines to model, and perception experiments to study, whether the acoustic cues used by speakers of IndE in the production of /v/ and /w/ permit reliable perception of the contrast.

6. REFERENCES

[1] Agnihotri, R. K., Sahgal, A. 1985. Is Indian English retroflexed and r-full? *Indian Journal of Applied Linguistics* 11.1, 97–108.

[2] Akaike, H. 1980. Likelihood & the Bayes procedure. *Trabajos de Estadistica y de Investigacion Operativa* 31.1, 143-166.

[3] Asher, R., Kumari, T. 1997. *Malayalam*. Abingdon: Routledge.

[4] Bansal, R. K. 1976. *Intelligibility of Indian English*. Hyderabad: Central Institute of English & Foreign Languages.

[5] Bansal, R.K., Harrison, J. B. 1994[1972]. *Spoken English. A manual of speech and phonetics*. London: Sangam.

[6] Bhattacharya, T. 2001. Bangla. In: Garry, J., Rubino, Carl (eds.), *Facts about the World's Languages. An Encyclopedia of the World's Major Languages, Past & Present*. New York/Dublin: Wilson, 65-71.

[7] Bjorndahl, C. 2012. The phonology & phonetics of ambiguity: A case study of /v/. Poster presented at *LabPhon 2012*, Stuttgart.

[8] Chand, V. 2009. [v] at is going on? Local & global ideologies about Indian English. *Language in Society* 38.4, 393-419.

[9] Chandramouli, C. 2011. Census of India 2011. Provisional Population Totals. New Delhi: Government of India

[10] Cowie, C., Murty, L. 2010. Researching & understanding accent shifts in Indian call centre agents. In: Forey, G., Lockwood, J. (eds), *Globalization, Communication & the Workplace*. London: Continuum, 125-146.

[11] Desai, S.B., Dubey, A., Joshi, B. L., Sen, M., Shariff, A., Vanneman, R. 2010. *Human Development in India*. Oxford: Oxford University Press.

[12] Fuchs, R. 2014. Integrating variability in loudness and duration in a multidimensional model of speech rhythm: Evidence from indian english and british English. In: Campbell, N., Gibbon, D., Hirst, D. (eds), *Proceedings of Speech Prosody 7* Dublin, 290–294.

[13] Fuchs, R. 2014b. Towards a perceptual model of speech rhythm: Integrating the influence of f0 on perceived duration. In: Li, H., Meng, H., Ma, B., Chng,E., Xie, L. (eds), *Proceedings of Interspeech 2014*, 1949-1953.

[14] Fuchs, R. 2015. You're not from around here, are you? - a dialect discrimination experiment with speakers of british and Indian English. In: Delais-Roussarie, E., Avanzi, M., Herment, S. (eds), *Prosody and Language in Contact: L2 Acquisition, Attrition and Languages in Multilingual Situations*. Berlin: Springer, 121–146.

[15] Fuchs, R. 2016. *Speech rhythm in Varieties of English. Evidence from Educated Indian English & British English*. Berlin: Springer.

[16] Fuchs, R., Maxwell, O. 2015. The Placement & acoustic realisation of primary & secondary stress in Indian English. *Proceedings of ICPhS* Glasgow, UK.

[17] Krishnamurti, J. 2003. *The Dravidian languages*. Cambridge: Cambridge University Press.

[18] Krivokapić, J. 2013. Rhythm & convergence between speakers of American & Indian English. *Laboratory Phonology* 4.1, 39–65.

[19] Maxwell, O. 2014. *The Intonational Phonology of Indian English. An Autosegmental-Metrical Analysis Based on Bengali & Kannada English.* PhD thesis. University of Melbourne.

[20] Maxwell, O. and Fletcher, J. 2009. Acoustic and durational properties of indian english vowels. *World Englishes* 28.1, 52–69.

[21] Maxwell, O., Fletcher, J. 2010. The acoustic characteristics of diphthongs in Indian English. *World Englishes* 29.1, 27–44.

[22] Moon, R. 2002. *A Comparison of the Acoustic Correlates of Focus in Indian English & American English*. MA thesis. University of Florida.

[23] Mukherjee, J. 2007. Steady states in the evolution of new englishes: Present-day indian english as an equilibrium. *Journal of English Linguistics* 35.2, 157–187.

[24] Nearey, T. 1977. *Phonetic Feature Systems for Vowels.* Dissertation. University of Alberta.

[25] Nolan, F., McDougall, K., de Jong, G., Hudson, T. 2006. A forensic phonetic study of dynamic sources of variability in speech: The dyvis project. In: Warren, P., Watson, C. (eds), *Proceedings of the 11th Australasian International Conference on Speech Science and Technology*,13–18.

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

[27] Ruinskiy, D., Dadush, N., Lavner, Y. 2010. Spectral & textural feature-based system for automatic detection of fricatives & affricates. *Proceedings Electrical & Electronics Engineers in Israel (IEEEI).*

[28] Sahgal, A., Agnihotri, R. 1988. Indian English phonology: A sociolinguistic perspective. *English World-Wide* 9.1, 51–64.

[29] Sailaja, P. 2009. *Indian English*. Edinburgh: Edinburgh University Press.

[30] Sailaja, P. 2012. Indian English: Features & sociolinguistic aspects. *Language & Linguistics Compass* 6.6, 359–370.

[31] Sirsa, H., Redford, M. A. 2013. The effects of native language on indian english sounds and timing patterns. *Journal of Phonetics* 41.6, 393–406.

[32] Stevens, K. 2000. *Acoustic Phonetics*. Cambridge/London: MIT Press.

[33] Thomas, E. 2010. *Sociophonetics: an introduction*. London: Palgrave Macmillan.

[34] Trudgill, P. 1999. Standard English: What it isn't. In: Bex, T., Watts, R. J. (eds), *Standard English: The Widening Debate*. London: Routledge, 177–128.

[35] Trudgill, P., Hannah, J. 2002. *International English. A guide to varieties of Standard English*. London: Arnold.

[36] Wiltshire, C. 2005. The "Indian English" of Tibeto-Burman language speakers. *English World-Wide* 26.3, 275-300.