DETAILING THE PHONETIC ENVIRONMENT: A SOCIOPHONETIC STUDY OF THE LONDON BENGALI COMMUNITY

Kathleen McCarthy[1], Bronwen G. Evans[2], Merle Mahon[1]

[1]Dept. of Developmental Science, [2]Dept. of Speech, Hearing and Phonetic Sciences, UCL, UK
kathleen.mccarthy@ucl.ac.uk

ABSTRACT

The present study investigated variability in the production of English by London Bengali adults who speak the Sylheti dialect of Bengali. Speakers had been resident in London for differing lengths of time. They were recorded producing /l/, /r/ and the eleven monophthongal vowels in English. Phonetically trained listeners rated speakers' productions of /l/ and /r/ in terms of manner and place. F1 and F2 formant frequency values, and duration were measured for all monophthongal vowels.

The results demonstrated that older speakers (first-generation immigrants) tended to use Sylheti-like variants when speaking English, whilst second-generation speakers tended to use native English-like variants. These findings will be used to inform studies of the role of phonetic input in child language acquisition from the London Bengali community.

Keywords: phonetic input, speech production, accent variation, phonetic development.

1. INTRODUCTION

The London Bengali community is the fourth largest ethnic group in London [8]. The majority of London Bengalis originate from the rural district of Sylhet in the north-east of Bangladesh, where the local vernacular is Sylheti. Whilst there is little published work on Sylheti, it has similarities with Standard Bengali (SB) [3]. SB has 7 monophthongal vowels, where vowel length is not contrastive. All vowels are on the periphery of the vowel space, with no central vowels. In contrast to English, plosives and affricates contrast in aspiration, and tap, trill and retroflex variants are common [7].

Continuous migration of Sylheti speakers to the UK since the 1950s has resulted in a complex multilingual community made up of first, second and subsequent generations. First-generation immigrants are often defined as learners of English as a second language. In contrast, their children, second-generation immigrants, have Sylheti at home and are schooled in English. English may then become their dominant language. However, due to the dense nature of the London Bengali community Sylheti remains the main language used both in the homes and within the local community. Consequently, even second and subsequent generation speakers born and raised within the UK acquire both English and Sylheti.

Sociolinguistic studies of other UK British Asian communities have found the use of first language (L1) phonetic features in the English of both first and second-generation speakers [5, 8]. However, there is very little published work that has investigated the phonetics of the London Bengali community.

The aim of the current study is to detail the English phonetic variants used by speakers of different generations from the London Bengali community. The study focuses on variation in the production of English monophthongal vowels and English /l/ and /r/, variants that have been found to vary in other British Asian communities [5, 8]. Adults from the London Bengali community recorded target sounds in carrier sentences. Phonetically trained listeners rated speakers' productions of /l/ and /r/ in terms of manner and place, and F1 and F2 formant frequency values and duration were measured for vowels.

2. MATERIALS AND METHOD

2.1. Participants

Forty subjects, aged 18-65yrs and from six language backgrounds were recorded: SB speakers, Sylheti controls, late first-generation Sylheti speakers (arrival in UK aged >25 yrs), early first-generation Sylheti speakers (arrival in UK aged <10 yrs), second-generation Sylheti speakers (born and raised in the UK), and monolingual Standard Southern British English (SSBE) controls. All subjects were recruited from the London Boroughs of Tower Hamlets and Camden.

2.2. Materials and apparatus

Target sounds were elicited using pictures in the carrier sentence *say* _____ *again.*

For /l/ the target sound was elicited in word-initial (WI) and word-final (WF) position, e.g., *laugh*, *pool*, and for /r/ in WI and word-medial (WM) position, e.g., *run*, *lorry*. All English monophthongs were elicited in a CVC context: *bead* /i/, *kid* /ɪ/, *bed* /e/, *cat* /æ/, *cup* /ʌ/, *bird* /ɜ/, *card* /ɑ/, *cot* /ɒ/, *court* /ɔ/, *boot* /u/, *book* /ʊ/. Due to varying literacy levels amongst participants, all words were selected to be imageable. Consequently, vowels could not be elicited within the same CVC frame, though initial alveolars were avoided in order to maximize comparison across vowel categories [6].

Subjects recorded each word twice in a randomized order. All recordings were made in a quiet room and recorded onto a H2 Zoom recorder with a sampling rate of 44.1 kHz, 16-bit resolution.

3. ANALYSIS

3.1. /l/ and /r/: auditory analysis

Phonetically trained listeners rated each word for place and manner. For place, /l/ and /r/ were rated on a 5 point alveolar-retroflex scale. For manner, /l/ was rated on a 9 point clear-dark-vocalized scale, and /r/ was rated on a 5 point tap-approximant scale, with an additional option for trill. Where raters disagreed, the researcher categorized the token using a spectrogram. If raters felt the sound did not match the given scales, they were instructed to categorize it as *'other'*. All *'other'* tokens were removed from the current analysis.

A series of Pearson's correlations demonstrated that the raters were significantly correlated with each other (range; $r = 0.181$, $p < 0.05$ to $r = 0.921$, $p<0.001$), and so an average rating for each speaker for each variable in each word position was calculated. This value was used in all future analyses.

3.2. English monophthongs: acoustic analysis

F1, F2 and duration were measured using PRAAT [2]. Stimuli were located manually, and then F1 and F2 were extracted using hand corrected LPC analyses. Formant frequencies were measured from the midpoint of the steady state portion of the vowel. All duration measurements were taken from the beginning to the end of the F2 transitions. So the data from male and female speakers could be compared, Lobanov's z-score transformation [9] was used to normalize F1 & F2 values.

4. RESULTS

4.1. /l/ variable

Manner: A repeated measures ANOVA with word position coded as a within-subjects variable and group as a between-subjects variable investigated whether there were any differences in /l/ production. There was a main effect of word position, $F(1,34)=65.17$, $p<0.001$, where /l/ was clear WI and dark-vocalized in WF. There was a main effect of group, $F(5,34)=10.13$, $p<0.001$, and a significant interaction between word position and group, $F(5,34)=10.13$, $p<0.001$. Inspection of the data revealed that this was because the second-generation Sylheti speakers and SSBE controls used a variety of dark-vocalized variants in WF position, whereas all other groups used clear-dark variants.

Place: A repeated measures ANOVA, with word position coded as a within-subjects variable and group as a between-subjects variable investigated whether there were any differences in place of articulation. There was a main effect of word position, $F(1,34)=65.17$, $p<0.001$, and group, $F(5.34)=10.13$, $p<0.05$, and there was a significant interaction between word position and group, $F(5,34)=11.70$, $p<0.001$. Inspection of the data revealed that this was because second-generation Sylheti speakers and SSBE controls used retracted variants WF.

4.2. /r/ variable

Manner & Place: The potential differences in place and manner of articulation of /r/ were tested in separate repeated measured ANOVA analyses. Word position was coded as the within-subject variable, and group as the between-subjects variable. For place, there was a main effect of group $F(5,34)=10.24$, $p<0.001$. There were no other main effects or interactions ($p>0.05$). For manner, there was a main effect of group, $F(5,34)=10.55$, $p<0.05$, and a significant interaction between word position and group, $F(5,34)=3.31$, $p<0.05$. Inspection of the data revealed that this was because the Sylheti controls and SB speakers used tap and trill variants in WI position and approximant in WM position. The

other groups consistently used approximant WI and WM.

4.3. Monophthongs

F1 and F2 values are displayed in Fig 1. The early first-generation Sylheti, second-generation Sylheti and SSBE speakers have a native English-like vowel space. In contrast, the Sylheti controls, late first-generation Sylheti and SB speakers have a vowel space similar to their L1.

The differences in F1, F2 and duration were tested using separate repeated measures ANOVA with vowel coded as the within-subject variable and group coded as the between-subject variable. To avoid multiple statistical tests, vowels were divided into 3 groups; front /i, ɪ, e , æ/, central /ɜ, ʌ/ and back /u, ʊ, o, ɒ, ɑ/.

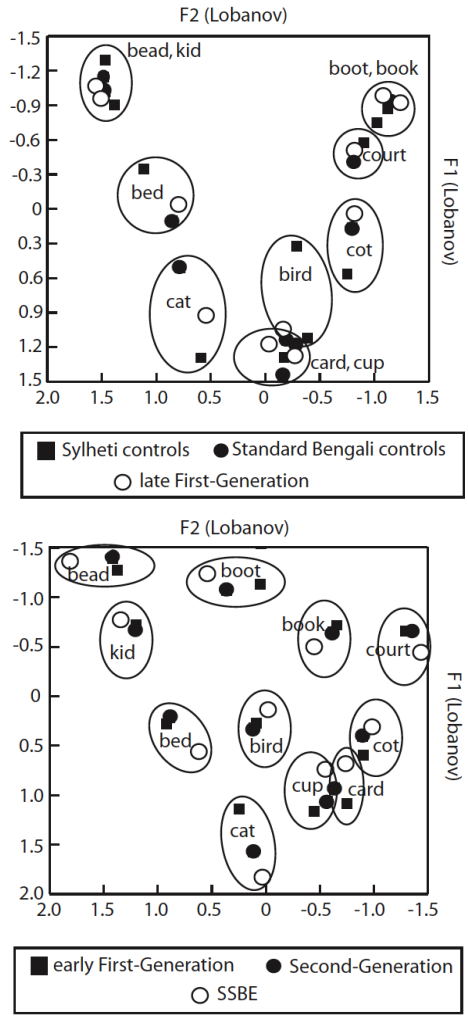
Front vowels: There was a main effect of vowel for F1, $F(1,34)=456.7$, $p<0.001$, and F2, $F(1,34)=223.4$, $p<0.001$, confirming that speakers were producing different vowels. There was a main effect of group for F1, $F(5,34)=3.89$, $p<0.001$, and F2, $F(5,34)=8.94$, $p<0.001$, and a significant interaction between vowel and group for F1, $F(5,34)=3.89$, $p<0.05$, and F2, $F(5,34)=2.29$, $p<0.05$. Inspection of the data revealved that this was because the Sylheti controls, late first-generation Sylheti and SB speakers merged *bead* and *kid*.

For duration there was a main effect of vowel, $F(3,34)=45.3$, $p<0.001$. There was no main effect of group ($p>0.05$), but there was a significant interaction between vowel and group, $F(15,34)=4.67$, $p<0.001$. Inspection of the data revealed that this was because the Sylheti controls tended to use a longer duration for *bed*, *cat* and *kid*, compared to the other groups.

Central vowels: There was a main effect of vowel for F1, $F(1,34)=42.12$, $p<0.001$, and F2, $F(1,34)=33.43$, $p<0.001$ confirming that speakers produced *cup* and *bird* with a different F1 and F2. There was a main effect of group for F1, $F(1,34)=10.81$, $p<0.001$, but not for F2 ($p>0.05$), confirming that groups produced *cup* and *bird* with a different F1. There was also a significant interaction between vowel and group for F2, $F(5,34)=6.14$, $p<0.001$, but not for F1 ($p>0.05$). Inspection of the data revealed that this was because the Sylheti controls, late first-generation Sylheti, and SB speakers produced *bird* with a lower F2 and *cup* with a higher F2 than did the other groups.

For duration there was a main effect of vowel, $F(1,34)=178.73$, $p<0.001$. There was no main effect of group ($p>0.05$) or significant interaction between vowel and group ($p>0.05$).

Figure 1: F1 and F2 formant frequency plots for subjects' productions of target English monophthongs.



Back vowels: There was a main effect of vowel for F1, $F(1,34)=712.3$, $p<0.001$, and F2, $F(1,34)=22.7$, $p<0.001$, confirming that speakers produced these vowels with a different F1 and F2. There was a main effect of group for F1, $F(5,34)=2.82$, $p<0.001$, and F2, $F(5,34)=5.75$, $p<0.001$, and a significant interaction between vowel and group, for F1, $F(5,34)=2.82$, $p<0.001$, and F2, $F(5,34)=9.68$, $p<0.001$. Inspection of the data revealed that this was because the Sylheti controls, late first-generation, and SB speakers produced *book* and *boot* with a lower F1 and F2 than did the other groups, resulting in a merged category for these vowels. In contrast, early first-generation

Sylheti speakers, second-generation Sylheti speakers and SSBE controls produced *boot* with a higher F2, i.e., more front vowel, typical of GOOSE fronting. The Sylheti controls, late first-generation speakers, and SB controls also produced *court* with a higher F2 than did the other groups.

For duration there was a main effect of vowel, $F(4,34)=31.73$, $p<0.001$, and group, $F(5,34)=3.33$, $p<0.05$. There was also a significant interaction between vowel and group, $F(5,34)=2.24$, $p<0.001$. Inspection of the data revealed that this was because, in contrast to the other groups, the late-first generation, Sylheti controls and SB speakers tended to produce *cot* and *card* with a longer duration.

5. DISCUSSION

As expected, native-like SSBE phonetic inventories do not apply to the English spoken in this community. Late first generation speakers used categories similar to those in their L1 (i.e., Sylheti) when producing English /l/, /r/ and monophthongal vowels. Both early first- and second-generation speakers used more native English-like categories for /r/ and monophthongal vowels. However, for /l/ only second-generation speakers used variants similar to those used by SSBE speakers.

In line with previous studies of British Asian communities [5, 8] and L2 learners [1, 4], our results suggest that the late first-generation group use their Sylheti categories when speaking English. One possible explanation for this pattern is the effect of native language on the production of L2 phonetic categories [see e.g. 1, 4]. Late first-generation speakers likely used more accented variants because of the decreased likelihood that phonetic categories will be established for English sounds that do not have an exact phonetic counterpart in the L1. For example, Sylheti does not contain /ɪ/, only having the monophthong /i/. Consequently, native Sylheti speakers who have acquired English as a second language find it difficult to acquire the English /i/-/ɪ/ contrast, usually assimilating them into their existing /i/ category.

This study provides preliminary detail of some key aspects of the phonetic environment of the London Bengali community. We aim to use these findings to develop test materials for subsequent research that will investigate the phonetic development of children from this community.

The phonetic development of children that grow up in immigrant communities is likely to be different from that of their monolingual peers. In the London Bengali community, not only are the children exposed to more than one language (i.e., Sylheti and English), but they will also be exposed to foreign-accented variants of English. Moreover, the English phonetic input the children will be exposed to will vary depending on the child's family structure, e.g., whether the children have older second-generation siblings or first-generation grandparents living in the home. The main question for our next study concerns the consequences of this highly variable input for children's phonetic development.

6. REFERENCES

[1] Best, C., McRoberts, G. & Goodall E. 2001. American listeners perception of nonnative consonant contrasts varying in perceptual assimilation to English phonology. *J. Acous. Soc. Am.* 1097, 775-794

[2] Boersma, P. Weenik, D 2010. *Praat: Doing phonetics by computer*. Retrieved from http:// www.praat.org

[3] Chalmers, R. 1996. *Learning Sylheti*. London: The Centre for Bangladeshi studies

[4] Flege, J. E 1995. Second language and speech learning: Theory, findings, and problems. In W. Strange (ed) *Speech Perception and Language Experience: Issues in Cross-Language Research*. 233-277. York, Baltimore

[5] Heselwood, B. & McChrystal, L. 2000. Gender, accent features and voicing in Panjabi-English bilingual children. *Leeds Working Papers in Linguistics and Phonetics* 8, 45-70

[6] Hillenbrand, J. M., Clark, M. J., Nearey, T. M 2001. Effects of consonant environment on vowel formant patterns. *J. Acous. Soc. Am.* 109 (2) 748-763

[7] Khan, S. D. 2010. Bengali (Bangladeshi Standard). *JIPA* 40 (2) 221-225

[8] Lambert, K., Alam, K., & Stuart-Smith, J. 2007. Investigating British Asian accents: Studies from Glasgow. *Proc. XVIth, ICPhS*, Saarbrucken, pp. 1509-11

[9] Lobanov, B. M. 1971. Classification of Russian vowels spoken by different speakers. *J. Acous. Soc. Am.* 49, 606-608

[10] ONS 2009. *Office of National Statistics: population and migration. Population estimates by ethnic group*: http://www.statistics.gov.uk/statbase/Product.asp?vlnk=14238