ACQUIRING A MULTIETHNOLECT: THE PRODUCTION OF DIPHTHONGS BY CHILDREN AND ADOLESCENTS IN WEST LONDON

Rosie Oxbury & Kathleen McCarthy

Queen Mary University of London
r.f.oxbury@qmul.ac.uk; k.mccarthy@qmul.ac.uk

ABSTRACT

To date, much is still unknown about the acquisition of sociophonetic variation. This paper presents an acoustic-phonetic analysis of the English diphthongs FACE and PRICE in the spontaneous speech of adolescents and children who are acquiring Multicultural London English (MLE). The diphthongs are analysed in terms of their first and second formant frequencies at onset, and trajectory length. We find that the children show similar onset formant frequencies for the two diphthongs as the adolescents, suggesting that both age groups have acquired this feature of the MLE diphthongs. We also found differences between the age groups: specifically, the adolescents show a more monophthongal realisation of both diphthongs than the children. Taken together, these findings indicate that within this community MLE is acquired early in life, but that the adolescents are in possession of a more focused speech variety than the children.

Keywords: Diphthongs, acoustic-phonetics, London English, sociophonetic acquisition.

1. INTRODUCTION

Urban cities such as London, U.K., offer a unique opportunity to study the emergence of new language varieties [5, 32]. With increased immigration, these cities are often made up of complex multilingual and multidialectal communities. Children who grow up in such cities are typically exposed to speakers from many different language backgrounds. The emerging contact variety in such cases will likely be influenced by speakers from multiple heritage language backgrounds, with the new variety emerging from a "feature pool" of different input languages [24]. Such varieties are termed multiethnolects [6, 20, 25]. One such multiethnolect is Multicultural London English (MLE), developing in East London [5].

The majority of multiethnolect research has focused on adolescents, as most studies of language change in progress have typically found that adolescents use the highest rates of innovative variants [19]. Although children are recognized as playing an important role in new language and new dialect formation [15, 17], little is known about their role in the development of multiethnolects. Children

have been shown to acquire stable sociolinguistic variables, such as the deletion of apical stops in final clusters, from caregivers even by age 4 [22]. But where there is a difference between caregiver speech and community norms, studies have suggested that children initially acquire the accents or dialects of their caregivers, and only show signs of local community sociolinguistic variation at the age of 8-9 [17, 13]. In contrast, a study in London found that children aged 4-5 already showed the same vowel system as adolescents in their community [5]. The authors suggested that the multilingual nature of the community led the children to orient to their peers as a target for language acquisition at an earlier age than has been found in more monolingual communities.

This paper reports on a project whose aims were twofold: to see if MLE has diffused beyond East London, to West London; and to see if young children there appear to be acquiring MLE. MLE is supposed to have emerged in East London, where local heritage languages include Sylheti, Turkish, and African and Caribbean varieties of English. The West London fieldsite, by contrast, has significant Somali, Polish and Irish populations. In this paper, we compare the children's vowel production in English to that of the adolescents. Indeed, if MLE emerges through indirect language contact among children [5, 31], and if West London adolescents are participating in the diffusion of MLE across London, we might expect stark differences in the productions of the children and adolescents in this study.

The variables for this study are the diphthongs FACE, /eɪ/, and PRICE, /aɪ/ [30]. These diphthongs have been the focus of previous studies of MLE, and the changes in the diphthong system are thought to be the most salient feature of MLE [5, 11, 16, 18]. Here, FACE and PRICE are analyzed both in terms of the position of the onset, and vowel dynamics, i.e. how diphthongal or monophthongal the vowels are [14]. Previous research has shown that MLE adolescent and child speakers show "reversal of Diphthong Shift", i.e. the onsets of FACE and PRICE shift anti-clockwise in the vowel space; and in terms of dynamics, MLE adolescent speakers show monophthongization of these diphthongs [16].

2. METHODOLOGY

2.1. Participants

The child participants were aged 5-7 years (N=14, 7 female, 7 male). They were at school together in an area of West London, and all lived within the borough. The adolescent participants were aged 16–19 years (N=14, 7 female, 7 male). They were recruited and interviewed at a youthclub near the primary school. All adolescent participants were born in the UK except for one boy who arrived aged 3 years. All child participants were born in the UK except for one girl who arrived aged 3. All participants except three adolescent girls and one male child had exposure to at least one language other than English at home. These languages included Irish, Somali, Patwa, Farsi, Lingala, Arabic, Swahili, Portuguese, Urdu and Tamil.

2.2. Procedure

The adolescents were interviewed in pairs in a quiet space in their youthclub, using a Zoom H4 recorder with Audio-Technica lavalier microphone (sampling rate 44,100Hz, 16-bit resolution). The children were recorded using the same equipment. The children were recorded doing a modified Diapix task [1, 12]. In addition to FACE and PRICE, tokens of the point vowels FLEECE, TRAP, LOT and FOOT were elicited in order to map the boundaries of participants' vowel spaces. The target vowels and the keywords used to elicit them in the Diapix task are shown in Table 1. These were controlled for age of acquisition, word frequency and imageability, and selected from [21, 29]. The images representing the keywords were taken from standardised databases [7, 28].

Table 1: Keywords used in the Diapix task

Diphthongs		Point vowels	
Vowel	Keyword	Vowel	Keyword
FACE	Cake	FLEECE	Sheep
	Baby		Cheese
	Gate	FOOT	Football
	Table		Book
PRICE	Kite	LOT	Dog
	Five		Sock
	Spider	TRAP	Cat
	tiger		Hat

2.3. Analysis

The recordings were transcribed in ELAN and force-aligned using FAVE [27]. Vowel tokens were manually segmented in Praat [3]. Measurements of F1 and F2 frequencies were extracted using hand-corrected LPC analyses, and these were taken at the 20% and 80% duration points of the diphthong [10]. For the point vowels, F1 and F2 measurements were taken at the midpoint. 1614 tokens of FACE were analyzed (523 child data; 1091 adolescent data) and 2339 tokens of PRICE (704 child data, 1635 adolescent data). Five tokens per speaker for each of FLEECE, TRAP, LOT and FOOT were analyzed. The vowel tokens were normalized using the Watt-Fabricius method [8] in R [25].

FACE and PRICE were analyzed in terms of the first and second formant frequencies. An MLE realization of FACE is indicated by a lower F1 at onset, i.e. more close realization [16]. For PRICE, MLE shows a more front onset diphthong, i.e. higher F2 [16]. MLE shows monophthongization of both these diphthongs, i.e. the change in F1 and F2 from onset to offset is less [16]. Therefore, two measures were important for the analysis of the diphthongs:

1. **Onset F1/F2 frequency**, defined as the first or second formant frequency at the 20% time point
2. **Trajectory length**, calculated as Euclidean distance in the F1 and F2 dimensions between the 20% and 80% time points. This is the same as VL as defined by [10].
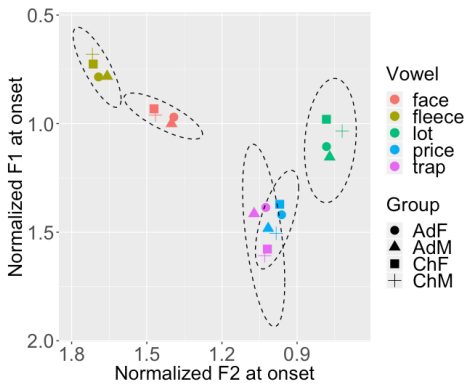
3. RESULTS

Separate mixed-effects linear regression models were run using lme4 [2] for the onset F1 of FACE, the trajectory length of FACE, the onset F2 of PRICE, and the trajectory length of PRICE. Fixed effects were age (binary: child or adolescent), gender and the duration of the vowel segment (log-transformed), while participant and word were included as random effects. Interaction terms were also included for age and gender, as it was expected that the adolescent boys would lead with respect to the MLE-related changes [5], and for duration and age, as impressionistically, it seemed that the nature of the Diapix task led children to produce words slowly and clearly. Subsequently type III analysis-of-variance tables were generated using the CAR package [9].

3.1. MLE diphthongs

An initial look at the vowel space (shown in Figure 1) indicates that the adolescents and children appear to have acquired MLE FACE and PRICE, in terms of the position of the diphthong onsets in the vowel system. The onset of PRICE overlaps with TRAP, as indicated by the overlapping ellipses. Similarly, the onset of FACE is close, as indicated by its proximity to FLEECE. The following sections will examine the onset formant frequencies of the adolescents and children,

and compare the two groups with respect to trajectory length.

Figure 1: Vowel plot of the mean normalized onset F1 and F2 frequencies, aggregated by gender and age group. 95% Confidence intervals represented as ellipses.
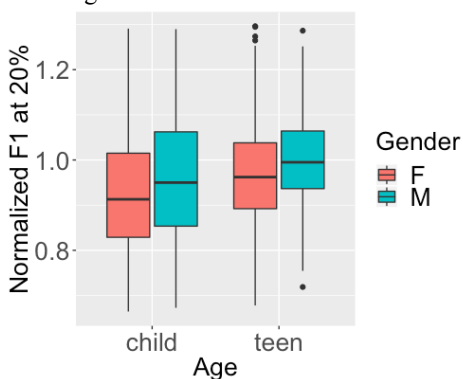

3.2. FACE

3.2.1. F1 onset

The model revealed a significant main effect of duration on onset F1 ($\chi^2(1)$=29.97, p<0.001), indicating that on average, tokens with a longer duration also had a higher F1 at onset, i.e. were more open at onset. Neither age, gender, nor their interactions with each other or duration were found to be significant (all p> .05). This suggests that there are not substantial differences between the two age groups nor between males and females in the onset of their FACE diphthongs.
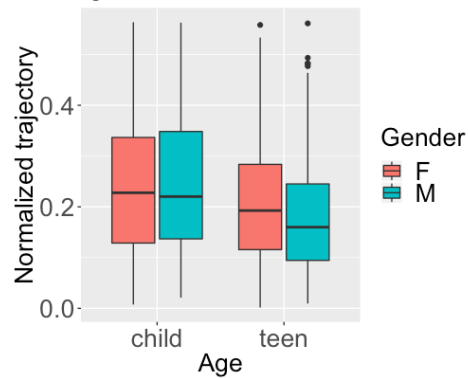
Figure 2: Boxplots of normalized F1 onset of FACE by age and gender.

3.2.2. Trajectory length

As regards trajectory, there was a significant main effect of duration (χ^2 (1)=44.73, p<0.001): a longer vowel duration predicts a longer trajectory, i.e. a more diphthongal realization of FACE. Age was not found to be significant, nor the interaction between age and duration, nor gender.

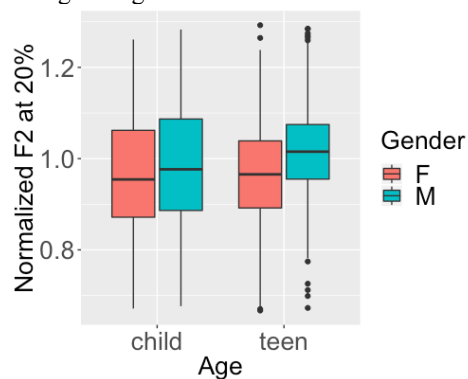**Figure 3:** Boxplots of normalized trajectory of FACE by age and gender.

3.3. PRICE

3.3.1. F2 onset

For PRICE F2 onset, the model revealed a significant main effect of duration (χ^2 (1)=51.93, p<0.001), indicating that the longer the duration of the vowel segment, the lower the F2 at onset tends to be. Age and duration showed a significant interaction (χ^2 (1)=14.62, p<0.001): an increase in vowel duration decreases the F2 at a steeper rate for the children than for the adolescents. Age on its own was not significant, meaning that children and adolescents do not differ significantly in their realization of the onset of PRICE. Gender and its interactions were not found to be significant.

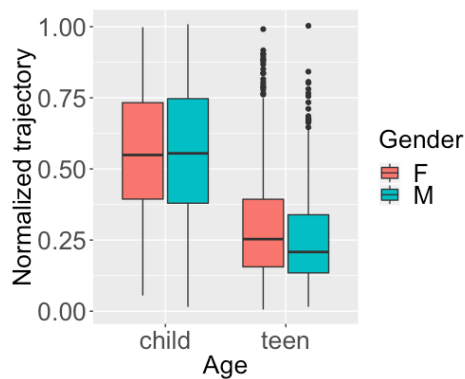**Figure 4:** Boxplots of normalized F2 onset of PRICE by age and gender.

3.3.2. Trajectory

For the PRICE trajectory, there was a significant main effect of duration (χ^2 (1)=155.24, p<0.001), with a longer duration predicting a longer trajectory, i.e. a more diphthongal realization of PRICE. There was also a significant main effect of age (χ^2 (1)=41.37, p<0.001). This means that adolescents have a shorter trajectory for PRICE than the children, when duration is kept constant. Therefore, adolescents show a more MLE-like realization of PRICE in terms of

monophthongization. The interaction between age and duration was not found to be significant. Neither gender nor its interactions were found to be significant.

Figure 5: Boxplots of normalized trajectory of PRICE by age and gender.

4. DISCUSSION

The key question for the study was whether the children would show signs of acquiring the same system as the adolescents. The results for both FACE and PRICE indicate that the children show similar production patterns to the adolescents in terms of the onset formant frequencies: the analyses confirmed that adolescents and children do not differ significantly with respect to the F1 at onset of FACE, nor with respect to the F2 at onset of PRICE. This offers support for Cheshire et al.'s [5] suggestion that in highly multilingual communities where many children are acquiring English as a second language, children adopt the linguistic norms of their community at an early age. This is in contrast to studies in more monolingual communities, which found that children initially acquire sociolinguistic variation from their caregiver, and only adopt community language features at around age 8-9 [13, 17].

Moreover, we found the same pattern as Cheshire et al. [5] in a different area of London, where a different array of heritage languages are spoken locally. This suggests that the pattern is indeed related to community multilingualism, and may not be dependent on the particular heritage languages involved. To explore this further, more research is needed on other multilingual communities, as well as comparative studies on specific heritage language groups within the same community.

We also investigated children's acquisition of the ongoing change of monophthongization of the diphthongs FACE and PRICE. The findings from the analysis of diphthong trajectory length show that the adolescents had a noticeably more monophthongal

PRICE vowel than did the children, but the two age groups did not differ for FACE. This means that in terms of trajectory length, the adolescents had a much more MLE-like realization of PRICE. This indicates that a diphthongal production of PRICE is acquired first, and the incoming change of monophthongization appears later. Again, it would be useful to have further research into the input these children receive, so as to know whether they are exposed to a diphthongal PRICE-like vowel in their caregivers' speech.

Why is there an age difference in monophthongization of PRICE, but not FACE? Prior literature on MLE does not compare the vowel dynamics of young children and adolescents, so we cannot say whether there is a precedent for this finding. However, the difference between FACE and PRICE is reminiscent of Cheshire et al.'s (2011) findings for the GOOSE vowel. GOOSE-fronting is a globally diffusing change. They found that the positions of FACE and PRICE in the vowel space of 4-5 yr olds were similar to those in the speech of adolescents in the same community, yet the adolescents had a much fronter GOOSE vowel than the 4-5 year olds (Cheshire et al. 2011). As with Cheshire et al.'s (2011) findings, it could be the case in the current data that the anticlockwise movement of FACE and PRICE is an endogenous change – e.g. due to the role of universals in language contact – while monophthongization of PRICE is a change diffusing across London.

We intend to build on these findings by exploring other diphthongs in the children's production; this may shed light on whether PRICE is an anomaly here, or whether there are general differences in children's acquisition of "reversal of Diphthong Shift" vs. the acquisition of monophthongization [16].

5. CONCLUSION

The present study compared FACE and PRICE in the speech of adolescents and children acquiring MLE. The children did not differ significantly from the adolescents in terms of their onset formant frequencies for FACE and PRICE, nor in terms of monophthongization of FACE, indicating that even by age 5-7, they have acquired incoming changes in their community. The adolescents and children differed only in one feature, the monophthongization of PRICE. Taken together, these findings suggest that MLE is acquired early in life. Further research is needed to explore the role of caregiver and peer input in the acquisition process.

6. REFERENCES

[1] Baker, R., & Hazan, V. (2011). DiapixUK: task materials for the elicitation of multiple spontaneous speech dialogs. *Behavior Research Methods*, *43*(3), 761–770.

[2] Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48.

[3] Boersma, P., & Weenink, D. (2018). Praat: doing phonetics by computer (Version 6.0.43).

[4] Brentari, D., & Goldin-Meadow, S. (2017). Language Emergence. *Annual Review of Linguistics*, *3*(1), 363–388.

[5] Cheshire, J., Kerswill, P., Fox, S., & Torgersen, E. (2011). Contact, the feature pool and the speech community: The emergence of Multicultural London English. *Journal of Sociolinguistics*, *15*(2), 151–196.

[6] Cheshire, J., Nortier, J., & Adger, D. (2015). Emerging multiethnolects in Europe. *Queen Mary's OPAL*, (33).

[7] Duñabeitia, J., Crepaldi, D., Meyer, A., New, B., Pliatsikas, C., Smolka, E., & Brysbaert, M. (2018). MultiPic: A standardized set of 750 drawings with norms for six European languages. *Quarterly Journal of Experimental Psychology*, *71*(4), 808–816.

[8] Fabricius, A., & Watt, D. (2008). A new speaker-intrinsic vowel formant frequency normalization algorithm for sociophonetics. *The Journal of the Acoustical Society of America*, *123*(5), 3069–3069.

[9] Fox, J., & Weisberg, S. (2011). *An R Companion to Applied Regression* (2nd ed.). Thousand Oaks CA: Sage.

[10] Fox, R. A., & Jacewicz, E. (2009). Cross-dialectal variation in formant dynamics of American English vowels. *Journal of the Acoustical Society of America*, *126*(5), 2603–2618.

[11] Fox, S. (2015). *The new Cockney: new ethnicities and adolescent speech in the traditional East End of London*. Houndmills, Basingstoke, Hampshire: New York, NY: Palgrave Macmillan.

[12] Granlund, S. (2015). *Speech communication strategies in older children: acoustic-phonetic and linguistic adaptations to a hearing-impaired peer* (PhD Thesis). University College London, London.

[13] Habib, R. (2014). Vowel Variation and Reverse Acquisition in Rural Syrian Child and Adolescent Language. *Language Variation and Change*, *26*(1), 45.

[14] Jacewicz, E., Fox, R. A., & Salmons, J. (2011). Cross-generational vowel change in American English. *Language Variation and Change*, *23*(1), 45–86.

[15] Kegl, J., Senghas, A., & Coppola, M. (1999). Creation through Contact: Sign Language Emergence and Sign Language Change in Nicaragua. In M. DeGraff (Ed.), *Language Creation and Language Change: Creolization, Diachrony, and Development*. Cambridge, MA: MIT Press.

[16] Kerswill, P., Torgersen, E. N., & Fox, S. (2008). Reversing "drift": Innovation and diffusion in the London diphthong system. *Language Variation and Change*, *20*(3), 451–491.

[17] Kerswill, P., & Williams, A. (2000). Creating a New Town koine: Children and language change in Milton Keynes. *Language in Society*, *29*(1), 65–115.

[18] Khan, A. (2006). *A sociolinguistic study of Birmingham English: Language variation and change in a multi-ethnic British community* (PhD Thesis). University of Lancaster.

[19] Kirkham, S., & Moore, E. (2013). Adolescence. In J. K. Chambers & N. Schilling (Eds.), *The handbook of language variation and change* (2nd ed., pp. 277–296). Malden, MA: Wiley-Blackwell.

[20] Kotsinas, U.-B. (1988). Immigrant Children's Swedish: A New Variety? *Journal of Multilingual and Multicultural Development*, *9*(1–2), 129–140.

[21] Kuperman, V., Stadthagen-Gonzalez, H., & Brysbaert, M. (2012). Age-of-acquisition ratings for 30,000 English words. *Behavior Research Methods*, *44*(4), 978–990.

[22] Labov, W. (1989). The child as linguistic historian. *Language Variation and Change*, *1*(1), 85–97.

[23] Labov, W. (2007). Transmission and Diffusion. *Language*, *83*(2), 344–387.

[24] Mufwene, S. S. (2001). *The ecology of language evolution*. Cambridge, UK; New York; Cambridge University Press.

[25] Quist, P. (2008). Sociolinguistic approaches to multiethnolect: Language variety and stylistic practice. *International Journal of Bilingualism*, *12*(1–2), 43–61.

[26] R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org

[27] Rosenfelder, I., Frühwald, J., Evanini, K., Seyfarth, S., Gorman, K., Prichard, H., & Yuan, J. (2014). FAVE (forced alignment and vowel extraction) program suite (Version 1.2.2).

[28] Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology. Human Learning and Memory*, *6*(2), 174.

[29] Stadthagen-Gonzalez, H., & Davis, C. J. (2006). The Bristol norms for age of acquisition, imageability, and familiarity. *Behavior Research Methods*, *38*(4), 598–605.

[30] Wells, J. C. (1982). *Accents of English: 1, An Introduction*. Cambridge: Cambridge University Press.

[31] Winford, D. (2003). *An introduction to contact linguistics*. Malden, MA: Blackwell.

[32] Wolfram, W., Carter, P., & Moriello, B. (2004). Emerging Hispanic English: New dialect formation in the American South: Emerging Hispanic English. *Journal of Sociolinguistics*, *8*(3), 339–358.