Multicultural London English in Ealing: sociophonetic and discourse-pragmatic variation in the speech of children and adolescents

Rosamund Oxbury

Submitted in partial fulfilment of the requirements

of the degree of Doctor of Philosophy

March 2021

School of Languages, Linguistics and Film
Queen Mary University of London

Statement of originality

I, Rosamund Frances Oxbury, confirm that the research included within this thesis is my own work or that where it has been carried out in collaboration with, or supported by others, that this is duly acknowledged below and my contribution indicated. Previously published material is also acknowledged below.

I attest that I have exercised reasonable care to ensure that the work is original, and does not to the best of my knowledge break any UK law, infringe any third party's copyright or other Intellectual Property Right, or contain any confidential material.

I accept that the College has the right to use plagiarism detection software to check the electronic version of the thesis.

I confirm that this thesis has not been previously submitted for the award of a degree by this or any other university.

Signature:

Date: 13th November 2020

Abstract

This thesis is about Multicultural London English (MLE), a multiethnolect that is spoken by young people in London (Cheshire et al., 2011) and potentially beyond (Drummond, 2016). The thesis investigates MLE in the speech of young people in a relatively understudied part of London: Ealing, a borough of West London. The speech of adolescents and children is compared to see if in Ealing, MLE features are used as part of an adolescent speech style, or are also acquired by children. Because a different range of heritage languages are spoken in Ealing compared to East London, the thesis also asks whether there are linguistic innovations in Ealing that have not been found in previous studies of MLE.

Using a variationist sociolinguistic framework, the project analyses MLE in the speech of 24 young people aged 16–24 and 14 children aged 5–7. The diphthongs FACE, PRICE and GOAT are analysed acoustically for both age groups. There is also a qualitative analysis of epistemic phrases (phrases related to *I swear*) in the adolescent data, motivated by the adolescents' use of *wallah* – an Arabic borrowing that has also been found in other European youth languages (Opsahl, 2009).

It is found that the children's and adolescents' diphthongs are similar in the quality of the onset, and similar to the emerging MLE system described by Kerswill et al. (2008). Among the adolescents, differences in the diphthongs pattern with language-internal effects as well as social factors including speaker sex and community of practice membership. The comparison between adolescents and children reveals that the children have acquired the same diphthong onset qualities as the adolescents – replicating Cheshire et al.'s finding in Hackney. However, the children have not acquired monophthongisation of the diphthongs. These findings have implications both for the study of multiethnolects and MLE, and for research on children's acquisition of sociolinguistic variation.

Dedication

This thesis is dedicated to the children and young people who generously gave their time and voices.

Acknowledgements

Thanks must go first of all to: the young people who were generous enough to help with the project, not just by taking part in recordings but allowing a weird university person to sit in awkwardly on their conversations and pester them with questions; their youth workers; the school teachers for accommodating me, allowing me to sit in on their lessons, and finding places and times when I could do recordings; the children for wanting to be involved and their parents for giving me the time of day and letting their children take part.

The fieldwork for this project began in January 2017. During the time that I was doing sociolinguistic interviews, the UK saw four terrorist attacks, a general election, and the Grenfell tower disaster. Negotiations for Britain's self-removal from the EU began. This was a time when it was hard to feel any kind of optimism for the future of the country. I would like to thank those who took part in this study for giving me hope. Sociolinguistic interviews have no agenda other than to get the participant to talk as much as possible. But that is only from the interviewer's point of view! Many of these teenagers used the microphone self-consciously to explain those aspects of their lives that they thought my university needed to know, and to put their best foot forward by telling me their plans and ambitions for the future. This thesis is dedicated to them, and to their primary school counterparts.

Thank you to my supervisors, Jenny and Kathleen. I feel extremely privileged to have been one of the last PhD students of Jenny Cheshire, and the first PhD student of Kathleen McCarthy! Thank you both for your patience, your encouragement and your attention to detail. It was a blessing to have two supervisors who are fantastic academics in their respective fields, but also just such nice people.

Thank you to the DDS at QMUL and Equality Focus for matching me with my mentor, Amina Ispahani. Amina has been a constant source of kindness, advice and encouragement over the last three years. It is no exaggeration to say that I would not have been able to produce this thesis without her. She has helped me break the PhD into month-by-month and week-by-week to-do lists, pushed me to plan my time, nudged me to tackle the tasks I really didn't want to do, talked me through break-ups and numerous house moves, and even got me to sign up for my first half marathon. Amina, thank you!

Thank you Erez, for first inviting me to QM to meet Jenny and discuss a $3+1$ application all the way back in 2014 and then for your dedicated attention to the various drafts of the PhD proposal. Thank you for suggesting I do an OIV in Copenhagen and for helping to make that happen.

I am grateful to have been part of the QMUL Linguistics community. In particular, thanks are owed to: Devyani, for teaching me Sociolinguistic Theory and encouraging me to write up my first journal article; Esther, for teaching me Phonetics and pushing me to write up my MA thesis; Colleen, for advice with fieldwork; Adam and Adib, for various bits of advice with the phonetic analysis. Thank you to the various iterations of the Linglab community who have brightened my day at different points: Zoe, Daniella, Shivonne, Annette, Melissa, Mohammed, Panpan, Chen; Nate, Christian, Maame; Ynda, Elisa, Scott, Tom Rausch; Elvis, Stamatina, Pietro, Chiara, Margaret, Lujain, Tom Meadows, Pedro, Adrian, Songyan, Antoaneta, Louis, Matthew, Liam, Celia, Jessica, Rosy Hall.

The Copenhagen department, especially Pia Quist and Nicolai Pharao for being so kind and welcoming, and the masters students for taking me under your wings. By pure luck, my time in Copenhagen coincided with a visit by Heini Lehtonen. If I hadn't met her, Chapter 8 would not exist in its current form.

Thank you Tanya, for telling me to keep going.

Thank you to those who got me through the (first) lockdown: friends and family who called me from the UK; Ikram, Andrea, Yao, Sylvette, Gita and Eva for looking out for me in Germany. Thank you to my friends, especially Sophie for being the Monica to my Rachel, and Sijana and Anna for all of the long, long phonecalls.

Matt: I don't know if this makes me a bad feminist, but I'm glad I did a PhD if for no other reason than that's how I met you! Thank you for putting up with me.

Thank you to my brother George and my own non-local caregivers. I have been hugely privileged to have a family as interested in language as I am! Thank you especially to my mother, who told me I would like Linguistics when I was trying to pick an undergrad degree ten years ago – I should have listened to you! My mother is actually one of the unsung heroes of Linguistics, having spent her career project-managing and copy-editing textbooks and handbooks for Wiley-Blackwell and CUP. I decided to spare her by not asking her to check my references...

Lastly, I'm grateful to the ESRC for the $3+1$ funding that made this project possible.

Contents

List of Tables

List of Figures

Transcription conventions

↑	Rise in intonation on that syllable
↓	Step down in intonation on that syllable
=	No gap between turns
[word]	Overlapped speech
<<word>>	Slow speech
>>word<<	Fast speech
((sound))	Extralinguistic information
(word)	Transcription uncertain
(. . .)	Unclear speech
(0.5)	Silence duration
(.)	Micropause
.	Final intonational contour
?	High rising terminal
¿	Less extreme rise than “?”
,	Less extreme rise than “¿”. Continuing intonation
:	Drawn out sound
word	Emphasis
WOrd	Emphasis and loudness
°word°	Quiet speech
wo(h)rd	Laughter breaking up a word
# word #	This utterance does not have laughter particles per se but it is audible that the speaker is smiling or holding in laughter

Abbreviations

AAVE, AAE	African American Vernacular English, African American English
AoA	Age of acquisition
BNC	British National Corpus
CPH	Critical Period Hypothesis
CofP	Community of Practice
DS	Diphthong Shift
ESS	Effective Sample Size
F1, F2	First formant, second formant
GVS	Great Vowel Shift
HDI	Highest Density Interval
LI	Linguistic Innovators
MLE	Multicultural London English
MUBE	Multicultural Urban British English
Onset F1	First formant frequency at the 20% time point
Onset F2	Second formant frequency at the 20% time point
PD	Probability of direction
ROPE	Region of Practical Equivalence
RP	Received Pronunciation
SLI	Sociolinguistic interview
SSBE	Southern Standard British English
TL	Trajectory Length
TRP	Transition Relevance Place

Chapter 1

Introduction and background

This chapter introduces the research context for the current project and motivates the research questions for the thesis.

We begin with an introduction to research on multiethnolects (Section 1.1) – this is the area of sociolinguistics to which the project aims to contribute. Because the theoretical issues at stake in the study of multiethnolects are largely inherited from the study of ethnolects, this section also reviews work on ethnolects. Broadly speaking, ethnolects are varieties of a national language that are associated with one particular ethnic group, while multiethnolects are similar, but develop in cities where there is great linguistic and ethnic diversity. They are thought to index belonging to a multiethnic peer group (Wiese, 2020). Section 1.2 then reviews recent work on Multicultural London English (MLE), the variety of English that is the focus of the thesis. MLE is thought to be a multiethnolect that developed in inner-city London in the late twentieth century.

However, the "London" part of MLE is up for debate. Section 1.3 discusses work that examines the spread of MLE features beyond East London, particularly Drummond's (2016, 2018b) work in Manchester. The possible reasons for these changes are then discussed: whether similar language contact outcomes have occurred in different parts of London and even in different cities, or whether adolescents are diffusing MLE between these locations. One of the contributions of this project is to analyse the MLE spoken in a different part of London, where a different array of heritage languages are spoken locally, and compare it to other researchers' findings on MLE in East London (e.g. S. Fox, 2015; Cheshire et al., 2011; Gates, 2019).

In order to distinguish between these two possibilities, it is necessary to compare the speech of children with the speech of adolescents. But children are also important to the study of multiethnolects in their own right. Multiethnolects are thought to arise via indirect language contact between children, and the work of Cheshire et al. (2011) was seminal in showing that in apparent-time, children as young as 4–5 had already acquired the same vowel system as adolescents in their community, leading them to suggest that in highly multilingual communities such as East London, children might orient to peers as their target in language acquisition at an earlier age than would be expected in monolingual communities. As such, Section 1.4 reviews: the expected role of children in the Labovian framework of the transmission and incrementation of linguistic change; the findings from a recent study of children's acquisition of a new dialect in Milton Keynes (Kerswill & Williams, 2000a) – this study forms an important point of comparison with that of Cheshire et al. (2011); and other relevant studies on (a) children's acquisition of the community language when their caregivers do not come from the local area, and (b) the influences of starting school and being part of a peer group on children's sociolinguistic development.

Section 1.5 then sets out the research questions and the approach to be taken in the current study.

1.1 Multiethnolects and ethnolects

The backdrop to this thesis is recent research on multiethnolects in Europe (Clyne, 2000). Variationist sociolinguistics in the 20th century often focused on well-defined speech communities, comprising individuals whose families had lived in the area for generations (e.g. Labov, 1966). Increasingly, researchers have begun to pay attention to the contributions made by multilinguals and immigrants – mobility – in language change (Horvath & Sankoff, 1987; Kotsinas, 1988; Rampton, 1995). Arguably, there has been a switch from privileging "sedentarism" to fetishising "nomadism" in variationist sociolinguistics (Britain, 2016). In various European countries, much recent research has focused on the innovative speech practices associated with young people in multiethnic, multilingual friendship groups (for an overview, see Cheshire, Nortier, & Adger, 2015). Multiethnolects are argued to be a particular type of contact language (Dorleijn & Nortier, 2013).

Many studies of multiethnolects (e.g. Quist, 2008; Cheshire et al., 2011) take as their starting point Clyne's definition of a multiethnolect:

> The other type of ethnolect may be termed a 'multi-ethnolect' because several minority groups use it collectively to express their minority status and/or as a reaction to that status to upgrade it. [...] It is the expression of a new kind of group identity. (Clyne, 2000, p.86)

The idea of a multiethnolect is, therefore, parasitic on the concept of ethnolects, and so this section will review research on ethnolects as well as multiethnolects.

1.1.1 Ethnolects (US-based research)

Ethnolects are varieties of the majority language that are associated with a particular ethnic group, as the name suggests. Examples include: Chicano English in California (Eckert, 2008; Mendoza-Denton, 2008); Chinese and Italian English ethnolects in Toronto (Hoffman & Walker, 2010); the English of Polish New Yorkers (Newlin-Łukowicz, 2013, 2015, 2016); Jewish Russian (Verschik, 2007); Jewish Lithuanian (Verschik, 2010); Turkish German (Kern, 2015); Chicano English (Fought, 2003; Mendoza-Denton, 2008); Maori English (Szakay, 2012). For some of these, only one or two variables have as yet been studied, while others are known to show distinctive features at all levels of the grammar. For example, African American English (AAE) is known to comprise pronunciation features such as *r*-lessness, DH- and TH-stopping, copula deletion, and verbal markers such as habitual *be*, among other features (L. J. Green, 2002; Rickford & McNair-Knox, 1994).

The study of ethnolects represents a change in thinking about cases of language shift. Earlier studies in bilingualism typically looked at the outcomes of language shift in terms of bilinguals' degree of success in approximating the target language, while the retention of features linked to the L1 was seen as fossilisation (Selinker, 1972). For this reason Matras (2009, p.76) states that the "social aspect of interlanguage is often overlooked in traditional approaches, which tend to focus on the individual's role in the second-language learning process".

Hoffman and Walker (2010) distinguish strong and weak interpretations of "ethnolect". According to Hoffman and Walker (2010, p.42), the strong interpretation "predicts that members of the same ethn(olinguist)ic group should resemble each other in their linguistic behavior (regardless of generation or native-speaker status) while differing from the larger population". The weak interpretation holds that ethnolects "serve to differentiate speakers who wish to convey membership in a particular ethnic group ... with linguistic features that may or may not derive from substrate transfer". Thus, the key difference between the two interpretations is the role of identity. The strong interpretation sees the distinctive feature of the ethnolect as involuntary and a consequence of fossilisation in the wider community, while in the weak interpretation, successive generations have agency in their uptake of ethnolectal features depending on their degree of identification with their ethnic heritage. Hoffman and Walker (2010) claim that their own results support the weak interpretation, as the degree of use of ethnolect features among their participants appears most strongly related to ethnic orientation.

Sociolinguistic research in the 21st century has seen a move from etic to emic understandings of identity, and consequently practice-based approaches to linguistic variation (Mendoza-Denton, 2002). To better understand how participants construed their membership in ethnic groups, Hoffman and Walker (2010) devised an ethnic orientation questionnaire, based on three key tenets of ethnic identity: the perception of difference, both by in-group and out-group members; sharing qualities or values, e.g. "language, religion, race, homeland or origin, culture, interests, and goals"; and sharing practices related to these values (Hoffman & Walker, 2010, pp.40–41). From this, they conclude "Rather than speaking of ethnicity, we should speak of ethnicities, or degrees of ethnicity, which vary from individual to individual and from situation to situation" (Hoffman & Walker, 2010, p.41).

A further important contribution of Hoffman and Walker (2010) is that their analysis centres not on supposed substrate variables, but on two features of Canadian English, (t/d) deletion and the Canadian Vowel Shift, that could potentially be influenced by Italian or Cantonese as an L1. In recent studies of ethnolects, there has been a shift from looking at speakers' use of substrate features, to looking at the interaction between heritage language and participation in majority-language variation and change (e.g. Gnevsheva, 2020). This has led Newlin-Łukowicz (2015) to note that ethnic groups can signal in-group affiliation either by using substrate features in the majority language, or by resisting uptake of linguistic variation that is current in the majority language.

1.1.1.1 Repertoire vs. lect

Some of the ethnolects discussed above are distinguished by just one or two key features, while others consist of a range of features from different levels of the grammar. This raises the question of how different variables interact with one another and this raises more complex questions still, such as how to identify who counts as a user of an ethnolect (Benor, 2010, p.164) – this question also becomes relevant in relation to multiethnolects, where some studies prefer a quantitative approach (e.g. Cheshire et al., 2011), while others identify users and non-users of the multiethnolect (e.g. Quist, 2008).

According to Benor (2010), there is a need to be able to identify and name the distinctive ways of speaking that are associated with different ethnic groups, but at the same time, to acknowledge their inherent inter- and intraspeaker variability, whilst also avoiding reifying these ways of speaking and essentialising their speakers (Mendoza-Denton, 2002; Jaspers, 2008). Benor (2010) therefore proposes treating ethnolects as repertoires. This conception of repertoire is based partly on the concept of "verbal repertoire" (Gumperz, 1964) and partly on the idea of a "pool of resources" (Fought, 2006, p.21). In the repertoire approach, the distinctive linguistic features that are included in the repertoire are treated as relatively fixed, meaning that it is easier to take account of dynamic ethnic identification.

Benor's (2010) approach comes with a number of problems. The repertoire approach does not work so straightforwardly when ethnic groups are not differentiated by the use of different variants, but by different language-internal constraints on variation (Hoffman

& Walker, 2010; Schleef, Meyerhoff, & Clark, 2011). Another problem is that, while the repertoire approach emphasises individual agency and identity as dynamic rather than static, this comes at the expense of linguistic structure. Newman (2010) addresses this latter problem by using implicational scales to assess the degree of systematicity in New York Latino speakers' use of Spanish-substrate features in their English.

Sharma (2011) argues in favour of "ethnolinguistic repertoire" rather than "ethnolect", but also warns against interpretations that put too much emphasis on fluidity. Sharma (2011) identifies different types of repertoire amongst older and younger second generation Pakistani/Indian speakers in Southall, London, UK. An older man shows a differentiated repertoire, shifting between Indian and British varieties according to different communicative contexts; similarly, younger women employ Indian English variants in their speech in their home, but use British variants almost entirely in other contexts. Meanwhile, older women and younger men do not show such a "differentiated repertoire", but rather, show a "fused lect".

1.1.2 European research & multiethnolects

The term "multiethnolect" stems largely from European and, in the case of Clyne, Australian research (e.g. Clyne, 2000; Kotsinas, 1988). This term and its alternatives are contentious (see Section 1.1.2.2 below), but broadly, "multiethnolect" refers to a form of youth language that has emerged in multilingual and multiethnic neighbourhoods of major cities, and that is used by both by young people from marginalised ethnic groups and from the societally dominant ethnic group. Youth languages of this kind have been attested across Europe: London, the topic of this thesis (Cheshire et al., 2011); Berlin (Wiese, 2009, 2013; Freywald, Mayr, Özçelik, & Wiese, 2011); Amsterdam and Utrecht in the Netherlands (Appel & Schoonen, 2005; Nortier & Dorleijn, 2008); Flanders in Belgium (Marzo & Ceuleers, 2011); Oslo (Aasheim, 1995; Opsahl, 2009); Copenhagen and Køge in Denmark (Quist, 2005; Møller, 2009); Stockholm, Gothenburg and Malmö in Sweden (Bodén, 2010; Kotsinas, 1988). Similar youth languages exist in Kenya, DR Congo, Sudan, Ethiopia, Rwanda, Central African Republic, Zimbabwe and Uganda (see contributions in Nassenstein & Hollington, 2015).

The actual phenomena that get described as "multiethnolects" are extremely varied. Dorleijn, Mous, and Nortier (2015) compare youth languages in the Netherlands with the case of Sheng in Kenya. Whereas Dutch youth languages are identified by phonetic features and also grammatical deviations from Standard Dutch, Sheng does not deviate much from the local Nairobi variety of Swahili in phonetics and grammatical structure. Instead, Sheng is characterised by a lexicon that is borrowed from English, Luo, Kikuyu among others, and also manipulated Swahili words; this lexicon is continually being replaced.

Wiese (2020) designates the African youth languages mentioned above as "multilingual mixed languages", and the European multiethnolects "majority language dialects", attributing the different outcomes to the multilingual habitus that prevails in many African locations, and the monolingual societal habitus of European countries.

Within the European youth languages, there are many differences and some commonalities. Multiethnolects in Scandinavia reportedly sound like learner versions of those languages, and indeed the first work on immigrant speech in Stockholm was motivated by the desire to show that the young people's way of speaking was a fully-fledged way of speaking, and not simply the result of incomplete acquisition of Swedish (Kotsinas, 1988). Several are described as having distinctive intonation and reduction of marked prosodic features such as Danish *stod* (Hansen & Pharao, 2010; Torgersen & Szakay, 2012). In the Germanic languages Danish, Swedish, Norwegian, German and Dutch, multiethnolects show a tendency to use V3 word order rather than the standard V2 word order (Freywald, Cornips, Ganuza, Nistov, & Opsahl, 2015). There are also levelling changes in morphosyntax, such as simplification of indefinite and definite article allomorphy in English (Cheshire et al., 2011), overgeneralisation of common gender in Dutch (Cornips, 2008), and reduction of inflectional suffixes in German (Wiese, 2009).

At the same time, there are key differences. Studies of Danish and Dutch multiethnolects report that the multiethnolect is a style that users can turn on or off, and that gets used with specific interlocutors and in specific situations (Quist, 2008; Hansen & Pharao, 2010; Nortier & Dorleijn, 2008). By contrast, MLE is described as some speakers' Labovian vernacular (Cheshire et al., 2011). Different linguistic and cultural influences are relevant to different multiethnolects: Dutch multiethnolect involves the use of a Moroccan accent (Nortier & Dorleijn, 2008); Jamaican Creole is seen as a key influence on MLE (Cheshire et al., 2011); the development of Helsinki multiethnolect is associated with Somali immigrants (Lehtonen, 2011).

Various authors agree that multiethnolects are better characterised by their function – expressing identity – than by linguistic features (Dorleijn & Nortier, 2013, p.233). A key theme across research into youth languages is that they express an oppositional identity. Kotsinas (2001) states "the features may be interpreted as markers of local identity. [...] the use of certain slang words and other linguistic elements regarded as 'incorrect', are to a certain extent used to give an impression of toughness and opposition to mainstream society." "Toughness" is in fact frequently mentioned in descriptions of the indexical associations of multiethnolects (e.g. Cheshire, 2020, p.15; Dorleijn & Nortier, 2013, p.243; Madsen, 2011b).

As suggested by Kotsinas's mention of "local identity", the locus of the youth language is the neighbourhood – hence names such as *Kiezdeutsch* ("hood German"). For example, Aarsæther, Marzo, Nistov, and Ceuleers (2015) describe how in both Genk

(Belgium) and in Oslo (Norway), the local youth vernacular has shifted from a second-order indexicality of signifying ethnicity to third-order indexicality of neighbourhood (Silverstein, 2003). In Oslo, youth from both Holmlia, a suburb, and from the city's East End, see the youth vernacular as being primarily linked to their own neighbourhood. The authors suggest that "the linguistic practices of adolescents in Holmlia and Gamle Oslo are on one level linked to their local neighbourhood, yet on another level linked to being an urban adolescent in late modernity, in a city flavoured by immigration and class stratification" (Aarsæther et al., 2015, p.267).

Finally, across various cities, the use of multiethnolect either seems to be associated more with boys than with girls, or is used differently by the different sexes. According to Cheshire et al. (2011), in London, the leaders of the changes appear to be Non-Anglo boys. Kießling and Mous (2004, p.317) state that "In all the reported cases [of youth languages in African countries], the role of boys is clearly more prominent than that of girls." Drummond (2018b, pp.194–198) posits a scale of "urban-street-style" orientation and has a boy as the archetypical high-orientation character. Meanwhile, Quist (2008) and Nortier and Dorleijn (2008) find that multiethnolect is used in different ways by girls and boys. In an ethnographic study of a Danish secondary school, Quist (2008) finds that multiethnolect features are used primarily by a group of anti-school boys who identify as "foreigners"; while the girls who tend to use multiethnolect are neither pro- nor anti-school, but favour signs of non-Danish ethnicity, e.g. listening to "ethnic pop and hip-hop".

1.1.2.1 Variety or style?

A key debate that resurfaces in discussions of multiethnolects is whether they are better conceptualised as varieties of the majority language, or as youth styles (e.g. Quist & Svendsen, 2010; Svendsen, 2015). This question is partly empirical (are multiethnolects systematic varieties, or transient youth styles?), partly methodological (do quantitative variationist, or interactional, approaches, give us greater insight into multiethnolects?), and partly political (are we disenfranchising multiethnolect speakers by describing their language as a variety/style?).

This methodological and theoretical division is first described by Quist (2008), who also shows that the two approaches are complementary. In the "structural variety" approach, the aim is to identify whether there is a new variety of a given language, and if so, what are its distinctive linguistic features that differentiate it from the standard language. "The variety approach [...] helps us to find out if there is anything linguistically systematic going on at all" (Quist, 2008, p.49). By contrast, in the stylistic-practice approach, "[s]peech [...] is looked at and analyzed within a local system of semiotic contrasts in a local community of practice" (Quist, 2008, p.49). The variety approach is a necessary precursor to the practice approach: the key linguistic variables need to be identified before we can study how they are used in interaction, or how they interact with other social symbols e.g. clothing, music preferences.

It can also be argued that the variety perspective is important in terms of legitimising the multiethnolect as a variety of the majority language. Wiese (2013) takes a "variety" approach to Kiezdeutsch in order to show that the variety has a consistent syntax, that the innovations are actually expansions of tendencies inherent in other German dialects, and consequently that the innovations reflect "language-internal motivations, rather than contact-induced effects" (p.211).

The issue, then, is the same as that found in the ethnolect research reviewed above (Section 1.1.1.1): studies following the "stylistic practice" approach must either rely on prior "variety" approaches, or their scope must be limited to surface phenomena, often transfer features. Although they avoid pre-defined social categories, they are to an extent reliant on pre-defined language varieties. Wiese (2013, pp.210–211) points out this contradiction in the practice approach: studies such as Blommaert and Rampton (2011) emphasise that named languages are ideological constructions, and advocate "deconstruction of the idea of distinct 'languages"', but then rely on naming the different source languages that are relevant in particular contexts.

1.1.2.2 Terminology

The issue of reification leads us to the question of terminology.

The chief objection against "multiethnolect" is the term's implied focus on ethnicity. As well as risking homogenising or Othering the speakers, the term can be misleading, as "there is nothing that indicates that these linguistic practices have 'only' to do with ethnicity" (Svendsen, 2015, p.8).

Rampton (2015) argues that such "styles" should be called "contemporary urban vernaculars": vernacular because they exist in opposition to a standard language; urban because they are found in cities; contemporary because this "push[es] us to consider exactly how far these styles are similar or different from the non-standard styles that pre-dated migration, a line of enquiry that is also likely to remind us that amidst all the forms identified as new, there is also often an abundance of quite traditional non-standard speech in multiethnic networks"; and because this term is more transparent than alternatives. Rampton (2015) opposes any term that contains "youth" because it should not be assumed that these ways of speaking are only used by adolescents, and in fact there needs to be more research into which features of, or in which situations and with which interlocutors, these "styles" are continued in adulthood.

Madsen (2011a) disagrees with the terms *ethnolect* and *multiethnolect* on the basis of the implied emphases both on ethnicity and on systematicity. She prefers instead the term "late modern youth styles", as "linguistic, ethnic, and cultural heterogeneity by now are inherently characteristic of late modern Western societies" and the term "late modern" "can cover the heterogeneous conditions related to these speech styles" – thus avoiding the potential homogenisation and reification that other terms bring. Conversely, Wiese (2020) emphasises the language contact aspect, suggesting the term "urban contact dialects": this designates "urban vernaculars that emerged in contexts of migration-based linguistic diversity among locally born young people, marking their speakers as belonging to a multiethnic peer group".

1.1.3 Summary

This section has outlined the broader field of research within which the current thesis is situated. Research into multiethnolects grew out of the fields of (a) L2 acquisition and (b) American research on ethnolects. Kotsinas's (1988) work was seminal in moving the study of Swedish immigrant speech away from a focus on L1-transfer errors, towards an appreciation of identity creation in L2 speech. Her work inspired similar research in other European countries, such as Denmark (Quist, 2005, 2008) and Germany (Wiese, 2009, 2013), among many others.

The tension between repertoire and lect approaches in studies of ethnolects translates into the variety vs. style debate in relation to multiethnolects (Quist, 2008). Just as in the study of ethnolects, whether the language form being studied is conceived of as a variety or a style is affected by, but also determines, the methodological approach followed, i.e. a variationist approach or a more qualitative approach.

1.2 Multicultural London English (MLE)

The study of multiethnolects in the UK began with two PhDs in the mid-2000s, S. Fox (2007) and Khan (2006).

S. Fox (2007, 2015) examined language use in a youth centre in Tower Hamlets, East London. The traditional dialect of that area is Cockney, but by the time of the study, the borough had a majority Bangladeshi population. S. Fox (2015) combined a social network analysis and Community of Practice approach to examine how innovative variants of diphthongs were diffusing through the attendees of the youth centre. S. Fox (2015) found that whereas White British girls in the study favoured traditional Cockney variants of the FACE and PRICE vowels, the Bangladeshi boys showed innovative raised and monophthongal variants. White boys who were in friendship groups with the Bangladeshi boys used the innovative variants to a lesser degree than the Bangladeshi boys. The overall trend was one of levelling, with the adolescents converging on variants that were intermediate between the most innovative forms and the conservative Cockney variants.

Khan (2006) examined the diphthongs GOAT and PRICE, and TH- and DH-fronting, in White British, Pakistani and Black Caribbean groups in Birmingham, aged 13–16, as well as older White British speakers aged 70+. Khan (2006) investigated social network, ethnic orientation and Birmingham orientation among the adolescents. Khan (2006) suggests that while the Black Caribbean adolescents led in the use of the [f] and [v] variants for (TH) and (DH), for the White British and Pakistani adolescents, use of [f] and [v] was associated with a "British" identity, as opposed to e.g. "English" or "Pakistani" orientation. The results for GOAT and PRICE also corresponded to network type and ethnic orientation.

Both S. Fox (2007, 2015) and Khan (2006) showed how within a speech community, different phonetic variants may be primarily associated with a given ethnic group, but that diverse friendship networks can lead to the diffusion of these variants to other groups, and to convergence on levelled variants. Moreover, the two studies had findings in common: in both studies, adolescents from the ethnic minority groups (Bangladeshi in Tower Hamlets, and Black Caribbean and Pakistani in Birmingham) favoured monophthongal variants for FACE/GOAT, and these same groups favoured a front open onset in PRICE, [aɪ]. This leads S. Fox, Khan, and Torgersen (2011) to suggest the term "Multicultural English" to describe the linguistic changes that appeared to be happening in the two locations (more on this in Section 1.3 below).

Meanwhile, the findings from two research projects in Greater London – Linguistic Innovators (LI) and Multicultural London English (MLE) – led Cheshire et al. (2011) to suggest that a new linguistic variety was developing in inner London, and that this variety could be described as a multiethnolect. They used as a point of comparison the work of Sebba (1993), who studied second generation Jamaicans in London in the 1980s. Sebba (1993) found that the young London Jamaicans would use a form of Jamaican Patois among themselves and in the home, but would use a form of Cockney in out-group settings. Cheshire et al. (2011) showed that whereas the speakers in Sebba (1993) had bidialectal competence in Cockney and Jamaican Patois, young people in inner-London in the mid-2000s had a vowel system that resembled the 1980s speakers' Patois vowels, rather than their London vernacular vowels.

Moreover, although Non-Anglo speakers had the most innovative vowel system, children and adolescents of both genders and of Anglo as well as Non-Anglo backgrounds participated in the same set of changes (Cheshire et al., 2011). (The authors distinguish between "Anglo" i.e. speakers whose families were White British and had been based in London for several generations, and "Non-Anglo", similar to the distinction in Scandinavian multiethnolect studies between bilingual and monolingual speakers).

Consequently, Cheshire, Fox, et al. (2013) describe MLE as "an ethnically neutral variable repertoire that contains a core of innovative phonetic, grammatical and discourse-pragmatic features". These features will be described in more detail in Chapter 5. They come from all levels of the grammar and include: unshifted, monophthongising diphthongs (Kerswill et al., 2008); TH- and DH-stopping and -fronting; a reduction in H-dropping; k-backing before non-high back vowels; slang words, largely from Jamaican Creole; *man* used as a pronoun (Cheshire, 2013); simplification of definite and indefinite article allomorphy (Cheshire et al., 2011; S. Fox, 2015). Importantly, Cheshire et al. (2011, p.153) claim that MLE may now be the new vernacular variety of English in London for many young people, according to Labov's definition: "the style which is most regular in its structure and in its relation to the evolution of the language is the vernacular, in which the minimum attention is paid to speech" (Labov, 1972, p.112).

A further important contribution of Cheshire et al. (2011) is that their study includes an apparent-time investigation of MLE in the speech of children across a variety of age groups, from 4–5 to 16–19 and young adults. The children aged 4–5 appeared to have the same vowel system as the 16–19-year-olds, and when the children's GOOSE vowel was compared statistically with the GOOSE vowel of their caregivers, no correlation was found. A discourse-pragmatic variable, the quotative *this is* + *speaker*, was also found in the speech of young children (although the younger children had less specific constraints on its use than older children and adolescents). These findings led Cheshire et al. (2011) to suggest that MLE emerged through group second language acquisition among young children (Thomason & Kaufman, 1988; Winford, 2003). In a setting such as Hackney, in inner-London, White British Londoners are no longer a majority, while many children born in the area will have caregivers who either speak minimal English, or speak it as an L2, or with a non-local accent – whether that be a regional accent from elsewhere in the British Isles, or a postcolonial variety of English such as Jamaican or Indian English. As a result, linguistic norms are diffuse (Le Page & Tabouret-Keller, 1985), and children's acquisition of English is guided by peers at nursery or at school as much as it is by their caregivers.

Other features differed between the inner London children and adolescents, namely the quotative BE LIKE and fronting of the GOOSE vowel. For these two features, the adolescents showed a significantly more advanced form than the children, indicating incrementation in the sense of Labov (2007, 2001b). Cheshire et al. (2011) explain this with reference to Milroy's (2007) "off the shelf" vs. "under the counter" distinction. "Off the shelf" changes are "relatively freely available to appropriately positioned social actors as a stylistic and social resource, regardless of the structure and location of their primary social networks" (L. Milroy, 2007, p.152, quoted in Cheshire et al., 2011, p.179), while "under the counter" features require face to face transmission and may be more linguistically complex. According to Cheshire et al. (2011), the endogenous, "under the counter" changes had been acquired by adolescents and children alike, while adolescents were leading in the "off the shelf" changes. This observation will become especially relevant when we consider the appearance of MLE changes in locations other than East London, in Section 1.3.

1.2.1 Summary

The current thesis is made possible by Cheshire et al.'s (2011) earlier work on MLE. They describe MLE as a multiethnolect, i.e. broadly the same kind of language change as has been described in other European countries (see Section 1.1 above). MLE comprises changes at different levels of the grammar and, importantly, the MLE project found that the vowel changes were already in evidence in the speech of children as young as 4–5.

1.3 The spread of MLE: the diffusion of an adolescent youth style, or similar language contact outcomes?

We have already observed that while the term MLE was first coined in reference to language change taking place in inner London, Khan (2006) had found similar sound changes in his study of different ethnic communities in Birmingham. This section reviews another empirical study with similar findings, namely a project carried out by Drummond (2016, 2018b, 2018a) in Manchester. Given the similarities across different locations, we will see that S. Fox et al. (2011) suggest the term "Multicultural English", while Drummond (2016) favours "Multicultural Urban British English". The purpose of this section is to discuss the potential causes of these similarities: do innovations diffuse between adolescents, or is it the similar situations of language diversity in all of these locations that result in similar language contact outcomes?

S. Fox et al. (2011) compare the findings of S. Fox (2007) in Tower Hamlets, London, Khan (2006) in Birmingham, and the Linguistic Innovators project in Hackney, London (Cheshire, Fox, Kerswill, & Torgersen, 2008b). The three locations are very different in that Hackney is extremely ethnically and linguistically diverse, while Tower Hamlets has a majority Sylheti Bangladeshi community. In Birmingham, meanwhile, there are significant Pakistani and Caribbean communities. Nonetheless, there are trends in the minority ethnic adolescents in all three locations to favour: [aɪ] for PRICE, rather than the Birmingham [ɔɪ] variant or the East London [ɑɪ] variant; [oː] or [ɔʊ] for GOAT, rather than the Birmingham or Cockney [ʌʊ] pronunciation; and [eː] or [eɪ] for FACE, rather than Diphthong Shifted [æɪ]. Meanwhile, White British adolescents in all three locations were more likely to favour traditional local dialect variants for these diphthongs (i.e. Cockney variants in the two London locations). Adolescents whose friendship networks included both White British and non-White British members were converging on diphthong variants that were intermediate between the most innovative and most conservative variants. S. Fox et al. (2011) name this emerging variety "Multicultural English".

Drummond (2016, 2018a, 2018b) investigated the use of MLE features among young people in Manchester. The project was intended to see if a Multicultural Manchester English could be identified, but early on, Drummond decided "it might be more useful to start thinking along the lines of a possible Multicultural Urban British English (MUBE)", defining this as "some kind of overarching variety or repertoire of shared features, with each urban centre having its own local version or sub-variety" – analogous to how (S. Fox et al., 2011) describe Multicultural English as being the same overarching phenomenon but with social and regional variation. The following features are identified by Drummond (2016) as part of an emerging MUBE, because they had previously been identified as features of MLE, were in use among Manchester adolescents, and are not part of the traditional local dialects of either London or Manchester: fronting and monophthongisation of PRICE; extreme fronting of GOOSE; word-initial TH- and DH-stopping; use of the pragmatic marker *you get me?*; and the use of Jamaican slang words.

There are different possible explanations for the similarities found between youth language in London, Manchester and Birmingham, and the purpose of the subsections below is to explore these alternatives further. The first possibility is that features diffuse between cities as part of an adolescent youth style; the second is that the same language contact outcomes have occurred in different locations. Section 1.3.1 considers the former while Section 1.3.2 discusses the latter.

1.3.1 Possibility 1: Adolescent youth style

This possibility is the one favoured implicitly by Drummond (2018b) and to some extent goes hand in hand with the "stylistic practice" approach. Drummond (2018b) links the use of MUBE features to orientation to an "urban/street/grime" way of life: those who show high orientation can be expected to use MUBE features, to be involved in making grime or rap music, and typically to be male; those who show low orientation are likely to use only traditional Manchester dialect features, to prefer hip-hop or R&B music, and to be female. In particular, Drummond's (2018a) variationist analysis of TH-stopping shows that whether an individual habitually raps as part of their everyday behaviour is a statistically significant predictor of their rates of [t] for (TH). While Drummond (2018b) does not state this explicitly, it is implied that MUBE features are adopted in an agentive way as part of the adolescents' identity creation.

It has long been known that adolescents are highly linguistically innovative (see, for example, Tagliamonte, 2016b), and that this is because adolescence as a life stage is an important nexus of identity formation. Kirkham and Moore (2013) argue that the sociolinguistic study of adolescents is valuable for two reasons: firstly, because of the adolescent peaks in changes in progress, meaning that "studies of adolescents provide the latest insights into processes of variation and change" (p.280); but secondly because of the social context of adolescence, making it "the perfect context in which to test the limits of the relationship between language and social categories, and language and social meaning" (p.280).

The adolescent peak refers to the peak visible on a graph of the rates of use of an incoming linguistic change by different age groups. Labov's (2007, 2001b) model of transmission, incrementation and diffusion attributes different roles to different age groups in advancing language change. In particular, according to Labov's model, changes in progress are expected to show an "adolescent peak" in apparent-time data, with preadolescents and young adults showing slightly lower rates of the incoming variant, compared to adolescents. The drop in the use of innovative forms among postadolescents is supposed to happen because in this model, it is assumed that there is a critical period for language acquisition, and that speakers' phonologies stabilise at around age 17. It is not clear that the adolescent peak would necessarily be applicable to the changes in progress in London, Manchester and so on, as the adolescent peak is part of Labov's (2001b) model of the transmission of sound changes, which in turn is contingent on a circumscribed, homogenous speech community.

However, Kirkham & Moore's (2013) second point, about the social context of adolescence, is highly applicable to teenagers in London, Manchester and Birmingham. Eckert (1998) describes adolescence as a kind of limbo, where children are expected to distance themselves from family influences and form friendship networks exclusively among peers: "people who are in fact becoming adult are normatively denied adult roles, and isolated from the adult sphere in institutions of secondary education" (Eckert, 1998, p.162). Because of the isolation from adjacent age cohorts, and the pressure to form friendship groups, the US high school "serves as a hothouse for the construction of identities" (Eckert, 1998, p.163). These intense processes of identity formation are what cause adolescents' propensity for linguistic invention.

Drummond's (2018b) study is by no means the first to observe adolescents' integration of non-local features into their own local accent, nor is it the first to link this phenomenon to media consumption. A. Williams and Kerswill (1999) report on a dialect levelling study that compared adolescents and older people in Milton Keynes, Reading and Hull. Reading and Milton Keynes are both in the Southeast of England, while Hull is in the Northeast. In Milton Keynes and Reading, there are signs of dialect levelling, and children rejecting the more-marked vowel variants among those used by their caregivers' generation in favour of non-regional, RP-like forms. In Hull, where social mobility is limited, the adolescents maintain conservative local vowel variants. However, the adolescents in all three towns are similar in their use of t-glottaling and TH- and DH-fronting. A. Williams and Kerswill (1999) suggest that these consonantal features, plus labio-dental /r/, constitute a set of "youth norms" that identify young people in all three towns with a wider youth culture. The authors write that "by adopting non-standard southern features, the young Hull speakers are able to signal their identification with the peer group and youth culture, while at the same time retaining their strong links with both their social class and their region of origin". The spread of these features is attributed to the Hull adolescents' regular exposure to southern English accents via TV and radio.

Similarly, other studies have also posited a link between adolescents' media consumption and their adoption of innovative, sometimes non-local, linguistic features. Tagliamonte (2016b) suggests a link between the characters of *Friends*' use of intensifiers and ongoing change in the intensifier system in the language of teenagers in the UK and North America see also (see also Tagliamonte & Roberts, 2005). Stuart-Smith et al. (2007) link Glasgow adolescents' adoption of TH-fronting to the popularity of the TV show *East Enders*.

Such findings go against the grain of what had previously been thought about the diffusion of sociolinguistic variation. Sociolinguists have traditionally believed that face-to-face contact is a necessary condition for the spread of linguistic changes (e.g. Trudgill, 1986). Trudgill's stance is in line with knowledge from language acquisition more broadly: in a comparison of children who were exposed to an L2 via social interaction, versus those who were exposed via audio and audiovisual recordings, it was found that those infants who experienced social interaction showed phonetic learning, but that those who were exposed to recordings only did not show any effect of exposure to the L2 (Kuhl, Tsao, & Liu, 2003) Yet findings such as those listed above show the need to consider adolescents' music and television preferences when analysing their language.

1.3.2 Possibility 2: similar language contact outcomes in different speech communities

The second possibility is that similar kinds of language contact have resulted in similar outcomes in different locations. In order to explore this possibility, this section will first lay out the model of the Feature Pool (Mufwene, 2003) as an explanation for the innovations seen in MLE, before moving on to discuss the possibility of similar language contact outcomes occurring in different speech communities.

The principle of the Feature Pool is that in any speech community (not just multilingual ones), speakers are coming into contact with one another, and their idiolects form a "pool" of available linguistic forms. In the formation of a creole or koine, these input idiolects may come from different communal languages or dialects. As the speakers interact with one another, selections are made from the pool of available forms and in this way, a new communal language is formed. The Feature Pool as Mufwene (2003) describes it thus appears to function at multiple levels: within the mind of the multidialectal/multilingual speaker, who is selecting between the communicative forms they have at their disposal; and at an abstract communal level.

The concept of the Feature Pool is useful to Cheshire et al. (2011) because they are describing a contact situation where direct transfer from any one source language is unlikely. The Feature Pool is named as such because it can be thought of as containing not just surface linguistic forms, but their underlying structure, e.g. grammatical forms, or syntactic or phonological features. The Feature Pool is based on an evolutionary metaphor: the features in the pool are supposed to be in competition with one another for selection. Hence, those that are reinforced by others are more likely candidates for selection. In this way, the pool is the locus of "blending inheritance", whereby features that are similar reinforce one another (Mufwene, 2003): if, for example, many of the languages in contact have distinct singular and plural second person pronouns, it is likely that the emergent mixed language will also have singular and plural second person pronouns. Importantly, this means that multilingual settings can favour outcomes that already existed in input dialects of the majority language: "In some cases they simply favored an option that was already available in some of the metropolitan varieties but was statistically too insignificant to produce the same output under different ecological conditions" (Mufwene, 2003, p.6). This closely aligns with the arguments put forward in Wiese (2009, 2013).

Mufwene (2003) calls the Feature Pool the "Complementary Hypothesis", because it resolves contradictions between substratist and universalist accounts of creole development. The universalist position is that the common features shared between different creoles arise from principles of Universal Grammar; the substratist position is that these commonalities arise from shared properties of the substrate languages that were in contact with colonial languages (Holm, 2000). Under Mufwene's Feature Pool hypothesis, both of these accounts are correct, as substrate languages form the input to the Feature Pool, and principles of Universal Grammar play a role in determining the output.

While language contact is discussed in an abstract way by Cheshire et al. (2011), its consequences are discussed more concretely by Cheshire, Adger, and Fox (2013), with reference to the relative pronoun *who* developing topic-marking functionality among young people in inner London. Across varieties of English in the UK, there is a general trend towards reduction of the number of different relative pronouns available, with young people favouring *that*. Cheshire, Adger, and Fox (2013) find that in inner London, unlike in outer-London and other areas of the UK, the increase in *that* does not correspond to a decrease in use of *who*. Through close analysis of how young Londoners use different relative pronouns, Cheshire, Adger, and Fox (2013) are able to suggest that the retention of *who* is related to its developing a topic-marking function. They explain this with reference to the heritage languages of their informants: many of the minority languages spoken in Hackney have a topic-marking feature, usually a discourse particle. English does not have such a feature, and so *who* is co-opted to provide one.

1.3.3 Summary

In sum, under the Feature Pool model, situations of linguistic superdiversity may be expected to produce similar outcomes in different speech communities, even when slightly different sets of heritage languages are involved in these separate communities. Features that are marked or cross-linguistically rare will lose out to less-marked competitors in the Feature Pool; meanwhile, if there are features that are common to several of the input languages, these will reinforce one another in the Feature Pool and be realised in some way in the output contact language.

This makes it extremely pertinent to compare the similarities and differences between forms of a multiethnolect, such as MLE, in different locations, as S. Fox et al. (2011) and Drummond (2018b) have sought to do. Comparing versions of a multiethnolect of the same national language (English) that have emerged in different geographical locations, where the social contexts differ and the range of heritage languages spoken in each community differs, can give us insight into the relative roles played by social factors on the one hand, and language universals on the other.

However, making this comparison based only on adolescent speech data produces a confound: as we have seen, adolescents are extremely linguistically innovative. They are liable to adopt and increment supralocal language changes, and this may occur even if face-to-face contact between adolescents in different cities is limited.

In the current study, in order to be able to separate language features that are part of an adolescent style, vs. those that appear to be endogenous to the local community, the speech of children will be compared with that of adolescents. The following section elaborates more on the importance of children to research on MLE.

1.4 The role of children in the development of MLE

We have said that in order to control for the possibility of MLE/MUBE spreading as part of an adolescent speech style, vs. language contact having similar outcomes in different locations, we need to compare the speech of children and adolescents.

Children are in fact seen as playing a critical role in the development of MLE. As we have seen, Cheshire et al. (2011) suggest that in highly multilingual communities, children orient to peers as their target in language acquisition earlier than in monolingual communities.

As a starting point, Section 1.4.1 first reviews the literature on the role played by children in the transmission of sociolinguistic variables in monolingual communities, where children acquire sociolinguistic variation from their caregivers and then participate in incrementation up until adolescence.

A detailed review of Kerswill & Williams' (2000a) study of Milton Keynes is then given in Section 1.4.2. This is firstly because Kerswill and Williams (2000a) was a seminal work in highlighting children's role in new dialect formation. Secondly, it is an important point of comparison with the MLE project: Kerswill and Williams (2000a) found that children aged 4 produced the vocalic variants that were favoured by their caregivers, regardless of whether their caregivers were local or non-local; while Cheshire et al. (2011) found that even children aged 4 had converged on the vowels used by their peers, rather than using their caregivers' non-local vowel variants.

In the interests of giving as thorough a picture as possible of what we might find from the children in the current study, the next sections review literature related to the findings of Kerswill and Williams (2000a) and Cheshire et al. (2011) as regards children's acquisition of sociolinguistic variation. Section 1.4.3 reviews findings from other studies of children whose caregivers were non-local. Finally, Section 1.4.4 reviews perspectives on how starting school or nursery affects children's sociolinguistic development.

1.4.1 Children in (monolingual) sociolinguistics

As touched upon above, the traditional view of children's role in continuing sociolinguistic change is that they acquire sociolinguistic variation from their primary caregiver (Labov, 2001b). Children must somehow be able to detect from the input they receive whether each sociolinguistic variable is an example of stable variation, or change in progress, and, if it is the latter, the direction of change, so that they are able to continue the change in the same direction (Labov, 2001b, pp.427–429). This process, by which the child acquires a variable's rate of use from the caregiver and then increases this rate as they grow up, is incrementation. Incrementation continues up until the child's phonology stabilises at around age 17 (Labov, 2001b, p.455), giving rise to the adolescent peak in apparent time, as mentioned above (Section 1.3.1).

It is known that children acquire sociolinguistic variation at the same time as they acquire non-variable parts of the grammar of their L1. While Labov (2013) states that "The acquisition of sociolinguistic variation begins fairly early in the third year", Chambers (2003, p.174) claims "there are no studies indicating a time gap between the acquisition of grammatical competence and the development of sociolinguistic competence"; whichever of these is correct, it appears that children begin acquiring variable rules as soon as their developing grammar permits them to do so. For example, Labov (1989) found that while a 4-year-old had acquired some of the social and linguistic constraints governing variation in (TD) clusters, his 2-year-old brother had not, because at that stage he had insufficient mastery of consonant clusters. Children's acquisition of sociolinguistic variation begins early but is constrained by their acquisition of non-variable rules.

As to the mechanism of transmission, Labov (2001b) states that children learn from their caregivers which variants are associated with formal vs. informal contexts. At a later stage, they learn to map this pattern of style-shifting onto social stratification: they learn that the variant associated with formal contexts is the overt prestige variant, while the variant associated with informal situations is the overtly stigmatised, covert prestige variant. Substantial evidence of children's ability to match their caregivers' style-shifting behaviours has been provided by Smith and Durham (2019): across a range of variables, caregivers were shown to use the dialect variant less in situations of teaching and discipline, and more in situations of play and everyday routine. For those variables for which their caregiver showed style-shifting, the children closely matched their caregivers' rates of use of the dialect variants in the different situations.

To summarise, caregiver input is as crucial to children's acquisition of sociolinguistic variation as it is to children's acquisition of non-variable rules of their L1 grammar. However, what is not clear from the summary above is how changes get transmitted and incremented when caregiver input does not provide a source of community variation – i.e. when the child's caregiver is not from the local community, and so there is a mismatch between the language heard in the home and the language heard outside in the community. Sections 1.4.2 and 1.4.3 review studies that deal with this eventuality.

1.4.2 Comparison with Milton Keynes

Kerswill and Williams (2000a) is a seminal study of how children adapt and converge to community language norms – indeed, contribute to the creation of these norms – when their caregivers are not from the local area. As such, it is also an important point of comparison for Cheshire et al. (2011). Cheshire et al. (2011) compare their findings for the vowels and especially GOOSE with the findings from Milton Keynes (Kerswill & Williams, 2000a): whereas in Milton Keynes, the 4-year-olds had adopted their caregivers' GOAT variant, in Cheshire et al. (2011), there is no correlation between the caregiver GOOSE variants and those of the 4–5-year-olds. According to Cheshire et al. (2011, p.171), this "suggests that children in multilingual communities in London attend to the speech of their peers at a younger age than in monolingual communities like Milton Keynes".

Kerswill and Williams (2000a) describe the emergence of a new dialect in Milton

Keynes, a "New Town" in the South of England. Kerswill and Williams (2000a) recorded 48 children who had been born in Milton Keynes, aged 4, 8 and 12, and the main caregiver of each child. Families who were actually from the local area were in a minority, and consequently there was "a catastrophic, creole-like discontinuity of dialect transmission" (Kerswill & Williams, 2000a, p.100); it was also found that many young people from Milton Keynes were unable to recognise the local accent of the area (Kerswill & Williams, 2000b). But as 94% of the caregivers in the sample were native speakers of a dialect of British English, there was no urgent need for a lingua franca. Instead, Kerswill and Williams (2000a) are concerned with investigating what forms koineisation takes (which variants end up in the koine), how quickly focusing (Le Page & Tabouret-Keller, 1985) takes place, and what the roles are of different age groups.

According to Kerswill and Williams (2000a), their evidence points to focusing taking place even in the first generation of children growing up in Milton Keynes – in contrast to Britain's (1997) suggestion that focusing takes place in the third generation. The children aged 8 and 12 showed different rates of use for several variables compared to the caregivers, including those mothers who were born in the region. The children favoured [f] and [v] for /θ/ and /ð/, and they favoured the RP variant, [aʊ], for MOUTH, even though the majority form among the caregivers was [æʊ] – Kerswill and Williams (2000a) describe this as a "strategy of neutrality". Similarly, although the caregivers' native dialects had a wide variety of GOAT variants (e.g. [oː, oʊ, ʌʊ, æʏ, æɪ]), the 8- and 12-year-olds had converged on [æʏ].

The key finding from Kerswill and Williams (2000a) is that the 4-year-olds appeared to have acquired the dialect of their principal caregiver, while the 8- and 12-year-olds (a) favoured quite different variants from their caregivers, and (b) showed relatively low levels of intragroup variation, compared to the level of variability seen among the caregivers. They suggest that the critical age for second dialect acquisition is between 4 and 8. This is seen in the apparent-time results across the different variables, but also in the real-time results for one boy, James, whose parents had moved from Scotland. When recorded in 1991, at age 4, he showed largely Scots phonology, but by the time he was recorded again at age 6, he had acquired the accent of his peers.

Kerswill and Williams (2000a) explain this finding in terms of orientation to the peer group vs. orientation to caregivers: at the age of 4, children are strongly attached to their caregiver, but as they progress through school and move towards adolescence, the peer group becomes increasingly important. It is through the peer group that linguistic norms are established, acquired and maintained. This is supported by the findings from the social network analysis in Kerswill and Williams (2000a), in particular for the GOAT variable: "The main factor is the child's orientation toward the peer group. All the high scorers (including the 8-year-olds not discussed here) are very well integrated into a (mainly school-centered) group of friends; they are sociable and are often cited as friends by other children. By contrast, the low scorers are somewhat distanced from their peers" (p.94).

In sum, for the Milton Keynes children, there appeared to be a turning point somewhere between the ages of 4 and 8. The 4-year-olds with non-local caregivers favoured their caregivers' preferred variants for GOAT, while the 8-year-olds favoured the vowel variants used by their peer group rather than the variants used by their caregivers. The critical difference between the findings of Kerswill and Williams (2000a) and those of Cheshire et al. (2011) is that the apparent-time results of the latter do not show a change between the ages of 4 and 8: even the 4-year-olds in inner-London were more similar linguistically to their peers than to their caregivers.

1.4.3 Other studies of children with non-local caregivers

Two studies from non-English contexts support the finding of Kerswill and Williams (2000a) that a turning point seems to occur between 4 and 8: Habib (2014) and Stanford (2008). Habib (2014) investigated four vocalic variables in 50 children aged 6–18 in Oyoun Al-Wadi, a village in Northern Syria. In this community, changes in recent decades had led to greater contact with urban Damascene Arabic, and many local men married women who came from outside the village. The children in the 6–8 age group showed the highest use of the urban vowel variants, whereas use of the rural vowel variants appeared to increase in successive age groups. The biggest increase was between the age groups 6–8 and 9–11. Stanford (2008) investigated three Sui clans of Guizhou province, China. The Sui culture involves exogamous marriage, meaning that women leave their own clan and marry into one of the other Sui clans. Children are initially exposed to their mother's dialect but are socialised to speak their father's dialect. Acoustic analysis of tone acquisition divided the children into two groups, those aged 3–5 and those aged 9–12. The children aged 3–5 used their patrilectal form for one tone, but for another tone in the system, they varied between patrilectal and matrilectal forms. The 9–12 year olds, meanwhile, used an exclusively patrilectal tone system.

Elsewhere, it has been found that children with non-local caregivers might be able to acquire community linguistic variation even at 4 years old, depending on the complexity of the variable. Roberts (1997b) compared two children aged 3;4 to 4;11 for whom at least one parent was not local to the Philadelphia fieldsite with their peers who had local caregivers. Mike, whose parents had both moved to the US from Italy as adults, did not differ from his peers with regard to the "simple phonetic variable" of (aw)-fronting. However, Mike had not acquired the raising of (ey) in checked syllables, unlike his peers, and did not show distinct tense and lax variants of short-*a*. Gia's father was Philadelphian but her mother "had moved frequently". Gia, like Mike (and all of the other children in the sample) had acquired (aw)-fronting. Her checked (ey) tokens were raised compared to non-checked (ey), but the two variants were not as distinct in her speech as they were in the speech of her native Philadelphian peers. Gia did not appear to have acquired short-*a* tensing in the same way as her peers: she showed some intermediate tokens, and some lax tokens before /f/, an environment that would strongly predict tensing in the Philadelphian system. Gia's greater success compared to Mike is attributed to her having extended family in South Philadelphia, having had a local babysitter, and having been cared for a lot of the time by her father rather than by her mother.

A related issue is second dialect acquisition: what happens when the child is born in one place and migrates to another community? This question is addressed by Tagliamonte and Molfenter (2007), who present a longitudinal study over 6 years of child second dialect acquisition. The subjects of the study are the first author's three children who were born in Canada and moved to York, UK, when all three were under 5 years old and the youngest was only 19 months old. Tagliamonte and Molfenter (2007) investigate the children's use of intervocalic flapped /t/, the North American form, vs. the local British variants of alveolar /t/ and the glottal stop. This study provides important evidence of how early sociolinguistic variation is acquired: evidence from overgeneralisation (e.g. [spaɪtə] for *spider*) shows the youngest child, who was 19 months old at the time of the family's migration, hypercorrecting intervocalic /d/ to [t], indicating that this child had acquired both the rule that flaps /t/ intervocalically, and that this rule does not apply in the British dialect: "in producing these tokens she is applying a first dialect rule to the second dialect." All three children showed a sharp increase in their rates of use of the British variants after they began school at age 4 (cf. Nardy, Chevrot, & Barbu, 2014). The youngest child initially showed the same rates of British variants as her older sister, who was already at school in York when the study commenced; this youngest child then decreased her rates of the British variants; and increased them again when her brother also started school in the UK. This study thus shows that siblings are important in the process of second dialect acquisition in ways analogous to the role played by siblings in bilingual language acquisition (e.g. Silva-Corvalán, 2014).

In multilingual contexts such as London, where many children's caregivers have an L1 other than English, children are exposed not only to their caregivers' L1 speech, but also to accented L2 input. This raises the question of when and how children identify their peers' speech as their target in acquisition, rather than their caregivers' L1-accented input. As we have seen, Cheshire et al. (2011) found that in their sample, the 4–5-year-olds seemed to be acquiring GOOSE-fronting, regardless of whether their caregivers also fronted GOOSE or pronounced it as a back vowel. In other words, those children oriented to their peers as their target in the acquisition of GOOSE, rather than their caregivers.

To account for children's ability to acquire the accent of their peers, rather than their caregivers, Chambers (2002, pp.121–122) hypothesised that immigrant children who receive L2-accented input from caregivers are able to "filter out" their caregivers' non-native accent:

> Ethan and the others come equipped with an innate filter so that when he hears his mother say "cherry" with tap /r/, he hears it as retroflex and pronounces it that way. When he hears his father say a word like "cell" with the tonic vowel pronounced [eː], he hears the vowel as [ɛ], and says it like that

According to Chambers (2002, p.122), rather than having to un-learn their caregivers' non-native accent, children "simply fail to hear [it]".

This account has since been called into question, for example by Khattab (2007, 2013). Khattab studied three English-dominant bilingual children from a Lebanese family living in Yorkshire, aged 5, 7 and 10. In the monolingual English recording sessions, the bilingual children were shown to favour the same set of variants used by their monolingual English friends, and to avoid the accented variants that their caregivers sometimes produced (Khattab, 2007). The children also used innovative, fronted pronunciations of the GOAT diphthong, meaning that they were participating in language change in progress in the local speech community (Khattab, 2007). In the Arabic-language recordings with their caregiver, the children switched to English occasionally, and Khattab (2013) analyses the use of English vs. Arabic phonetics in these switches. Switches to English often occurred when the child did not know the correct word for something in Arabic, and using Arabic phonetics in English could be a "compromise strategy", complying with the mother's encouragement to use Arabic but also using the child's preferred language, English. Meanwhile, use of Yorkshire-accented English could be a divergence strategy to show disagreement (Khattab, 2013). The data showed that the Arabic-accented variants present in the caregiver input must have been stored in the children's repertoires, and that these were used in an addressee- and context-appropriate way.

To sum up, while the findings of Habib (2014) and Stanford (2008), like Kerswill and Williams (2000a), found a key shift in children's productions from matching their caregivers to matching their peers at around age 8, others have somewhat different findings. Roberts (1997b) found that individual children's success in acquiring Philadelphian vocalic variables was linked to (a) the complexity of the variable and (b) the child's exposure to local adults. Tagliamonte and Molfenter (2007) present real-time evidence that the process of second-dialect acquisition begins as early as 19 months, and is affected by input from siblings and the key milestone of starting school. Khattab (2007) found that even the 5-year-old bilingual in her study was acquiring local changes in progress in the community.

However, it should be borne in mind that the studies summarised in this section differ from Kerswill and Williams (2000a) in that all of them present cases where there are relatively focused sociolinguistic norms, as opposed to Milton Keynes (and London; Cheshire et al., 2011) where linguistic norms were diffuse. Khattab's (2007) study is the most similar in this respect, in that the children were exposed to both local and middle-class English accents in their community.

Following from the insights in Tagliamonte and Molfenter (2007), the next section focuses on two key influences on children's language development from outside the home: starting school or nursery, and the peer group formed there.

1.4.4 Perspectives on the influence of school and the peer group

Nardy et al. (2014) examined sociolinguistic convergence of 4–5 year old children over their first year of kindergarten in Grenoble, France. They looked at three variables: variable liaison; optional deletion of /r/ in postconsonantal word-final position; optional deletion of /l/ in words such as *ils, elles*. The children were recorded at two time points, roughly one year apart. At the first time point, the children could be categorised as being more or less socially integrated, and those that were more integrated had higher rates of the non-standard variants for the three variables. At the second time point, the children could no longer be categorised as more or less integrated. Those children who had previously had highest rates of the non-standard variants had decreased; and those who previously had lowest rates of the non-standard variants had increased their rates. The authors argue that this result is evidence of sociolinguistic convergence.

Research in bilingualism has produced analogous results. K. McCarthy et al. (2014) compared Sylheti-L1 children's perception and production of English voicing contrasts at 52 months when they were in nursery (after an average of 7 months' exposure to English), and one year later, when they were in school. These children were growing up in London, where English is the dominant community language. At the first time point, the bilingual children differed from their monolingual peers in the production of voiced stops and in the perception of the /k – g/ contrast. When tested again one year later, the bilinguals did not perform significantly differently from monolinguals on either the perception or production tasks. This study demonstrates the rapidity with which children are able to adjust their L2 categories to match those of their monolingual peers. This, along with studies such as Nardy et al. (2014), is indicative of the importance of children's early interactions with peers in the first years of school or nursery.

Nance (2020) investigated the production of Gaelic laterals and stops by children aged 7–11 in Gaelic-medium education. Of the 18 children, 9 had little to no exposure to Gaelic in the home, while the remaining 9 had exposure to Gaelic via parents or grandparents. Home language was not found to have an effect on the production of either stops or laterals – indicating that the exposure to Gaelic in the school environment levelled out any differences that may have arisen from the different home language environments of the children. This is in line with other studies of adolescents acquiring a minority language (Mayr, Morris, Mennen, & Williams, 2017; Morris, 2017; Nance, 2015).

The studies reviewed above showed the rapid influence of the peer group on children's language. However, entering school can also be a standardising influence. The following two studies of bidialectal children suggest that starting school coincides with a shift in the child's production towards the standard language and away from the first-acquired dialect.

Van Hofwegen and Wolfram (2010) present a longitudinal study of children's acquisition of African American English from age 48 months to school grade 10 (age 15–16). They employ two measures of the children's dialect usage: the token-based Dialect Density Measure – number of vernacular features per communication unit or per words per speaker – and the type-based Vernacular Diversity Index, i.e. how many different types of features are represented in a given speaker's speech sample, as well as a variationist analysis. They identify two trajectories of AAE usage in terms of the Dialect Density Measure. Both of these involve a decrease in the number of features per utterance between 48 months and grade 1; then, the "rollercoaster" trajectory involves an increase in features per utterance between grades 4 and 6, but a drop after grade 6; while the curvilinear trajectory involves a steady increase in features per utterance from grade 4 to grade 10. The results of the token-based and type-based analyses are found to support each other: "the most vernacular children will also use the most varied AAE feature types and vice versa" (Van Hofwegen & Wolfram, 2010, p.441). They put the decline in AAE use in grades 1 to 4 down to "corrective effect of early school socialization in Standard American English" (p.448). They also note that "Different linguistic variables may show varied trajectories of change over the lifespan of speakers; some of these show similarities with the trajectories indicated by the dialect density measures, but not all" (Van Hofwegen & Wolfram, 2010, p.449; cf. Smith, Durham, & Richards, 2013; Smith & Durham, 2019). These authors also discuss the relationship between the acquisition of sociolinguistic variation, and variability due to developmental factors in children's speech: high rates of copula absence at 48 months are attributed to "convergence between a developmental structure and a dialect feature", as copula absence is a general feature of the early stages of child acquisition of English, not specific to children acquiring AAE (Van Hofwegen & Wolfram, 2010).

Youssef (1991) presents an in-depth study of one child acquiring both Trinidian Creole and Standard English. The child was recorded with a variety of addressees – some speaking primarily SE, some speaking primarily TC, and some a mix – between the ages of 2;7 and 4;4. A number of findings are relevant in the present context. Like Van Hofwegen and Wolfram (2010), Youssef (1991) found that vernacular use declined when the child entered school. At the beginning of the study, the TC verbal system was dominant, and at the final time point, the child again had a higher proportion of TC forms compared to SE, but at the time that the child was in a linguistically prescriptive pre-school, certain TC forms disappeared from his speech altogether. Secondly, Youssef (1991) notes that for several of the verbal features discussed, when the form first appeared in the child's production, it was highly differentiated according to addressee and became less stylistically differentiated with time. Certain forms were also subject to semantic or topic-related constraints: SE past tense marking was used only in storytelling contexts until age 4;4; when *will* first emerged at age 2;10, up until 3;3 it was used solely in argumentative contexts; the child showed a semantic distinction between TC *go* for the real future vs. *goin to* for imagined future, then at 4;4 between SE *will* for real future vs. *goin to* for imagined future, substituting *will* for *go* in those semantic contexts.

1.4.5 Summary

We have already said that it is important to compare the speech of children with the speech of adolescents in order to establish how far a multiethnolect is an adolescent style. However, children are in their own right critical to the study of multiethnolects. Kerswill and Williams (2000a) had already suggested that children played an important role in the emergence of a new dialect in Milton Keynes; Cheshire et al. (2011) suggest that children play a similar role in the development of MLE in London, and that their participation actually begins at an earlier age than might be expected in monolingual communities.

1.5 Approach in the current study and research questions

To summarise what we have seen so far:

- The starting point for this project is recent research on multiethnolects in Europe (Clyne, 2000). Variationist sociolinguistics in the 20th century often focused on well-defined speech communities, comprising individuals whose families had lived in the area for generations (e.g. Labov, 1966). Increasingly, researchers have begun to pay attention to the contributions made by multilinguals and immigrants – mobility – in language change (Horvath & Sankoff, 1987; Kotsinas, 1988; Rampton, 1995). Arguably, there has been a switch from privileging "sedentarism" to fetishising "nomadism" in variationist sociolinguistics (Britain, 2016). In various European countries, much recent research has focused on the innovative speech practices associated with young people in multiethnic, multilingual friendship groups (for an overview, see Cheshire et al., 2015). Multiethnolects are argued to be a particular type of contact language (Dorleijn & Nortier, 2013).

- A seminal study by Cheshire et al. (2011) indicated the existence of one such youth variety in London. They conducted two research projects: a comparison of speech in Hackney (inner London) and Havering (outer London) – the Linguistic Innovators project; and a study of children, adolescents and young adults from across North East London – the Multicultural London English project. Their results indicated various kinds of innovation associated with Hackney and North East London: for example, extreme advances in supralocal changes such as GOOSE-fronting; innovation in the diphthong system, and rejection of the traditional Cockney accent; the incorporation of Cockney features such as TH-fronting and l-vocalisation, as well as ethnically marked DH- and TH-stopping; simplification of definite and indefinite article allomorphy (Cheshire et al., 2011). Some features are common to other dialects, some represent amplification of a tendency found elsewhere, and some, such as the *man* pronoun, /k/-backing and the quotative form *this is* + *speaker*, appear to be innovations unique to inner London (Cheshire et al., 2011; Cheshire, 2013). They described these features as comprising "an ethnically neutral variable repertoire that contains a core of innovative phonetic, grammatical and discourse-pragmatic features" (Cheshire, Fox, et al., 2013, p.65). They also claimed that for some speakers, this speech variety had become their "unmarked Labovian 'vernacular"' (Cheshire et al., 2011, p.153).

- The apparent spread of MLE to other cities raises key questions. We have seen that a comparison of reversal of Diphthong Shift in Tower Hamlets (London), Hackney (London) and Birmingham leads S. Fox et al. (2011) to suggest the term "Multicultural English", while the interaction of MLE features and local dialect features in Manchester adolescents' speech leads Drummond (2016, 2018b) to propose the term "Multicultural Urban British English (MUBE)". Where do these similarities come from? Are the same changes spontaneously occurring in different locations, or is there a supralocal youth language?

- We saw in Sections 1.3.1 and 1.3.2 that both of these are possible. The peculiar social context of adolescence makes adolescents especially linguistically innovative, and even without face-to-face contact, they are capable of contributing to the diffusion of supralocal language changes. At the same time, the linguistic diversity found in cities such as London, Manchester and Birmingham makes it entirely possible that similar changes in English could take place in all of those locations, without diffusion necessarily being a factor.

Therefore, the current project aims to investigate MLE in a different part of London – most previous research having focused on areas of East London, e.g. S. Fox (2007, 2015); Cheshire et al. (2011); Ilbury (2020); Gates (2019). This project will be the first substantive investigation of MLE in an area of London outside North East London. By comparing a new part of London with previous studies of MLE in East London, we will have some evidence (a) as to whether MLE is used in other parts of London, and (b) which features have spread. As we have seen, S. Fox et al. (2011) present similarities in the FACE, PRICE and GOAT diphthongs across Tower Hamlets, Hackney and Birmingham, while Drummond (2016) only mentions PRICE fronting as a feature of MUBE in Manchester.

This also allows us to probe the influence of the local Feature Pool on MLE in the chosen fieldsite. Each borough in London has quite different population demographics, meaning that different arrays of heritage languages are spoken in different local areas. This means that it is possible for different innovations to occur in different parts of London, due to the different sets of languages in contact.

However, we have established that local language diversity is not the only reason why we might find MLE features in the fieldsite: adolescents are capable of diffusing MLE features to their local community. To help us distinguish between these two possibilities, children are included in the study. Children are presumed to be less mobile than adolescents (cf. Kerswill, 1996; Kerswill & Williams, 2000a), and so features found in their speech are presumed to arise purely from the input they receive from caregivers, teachers and siblings, and (group second language) contact with their peers. By contrast, adolescents are likely to mix with peers from a wider geographical area, and are likely to consume a wider range of media (music, TV, internet), enabling them to acquire a diffusing, enregistered urban youth speech style.

Examining children's acquisition of MLE is an important topic in its own right. Children are thought to play a critical role both in transmission and incrementation of sociolinguistic variation, and in the emergence of new dialects. However, most studies of multiethnolects only include adolescents. This makes it impossible to know whether these multiethnolects are truly adolescent styles or not. It also means that Cheshire et al.'s (2011, p.171) suggestion that "children in multilingual communities in London attend to the speech of their peers at a younger age than in monolingual communities like Milton Keynes and Buckie" remains uninvestigated. One of the major contributions of this thesis is its comparison of MLE variables in the speech of adolescents and children.

Therefore, the research questions for the current study are:

1. Are MLE features used by adolescents in the Ealing fieldsite? What social and linguistic constraints govern their use?

2. Do the children appear to be acquiring the same speech variety as the adolescents?
3. Is there any evidence of multiethnolect development that has not been attested in East London?

For the sake of comparability, this project will follow the variationist methodology of Cheshire et al. (2011) as closely as possible. The aim is to recruit a sample of adolescents in the chosen fieldsite, and to recruit a comparable sample of children aged < 8 from the same community. In the end, as explained in the following chapter, a fieldsite was selected in Ealing, West London.

There are a number of possible outcomes from this comparison of East and West London adolescents, and adolescents and children in Ealing. there are a number of possible outcomes from this comparison of East and West London adolescents, and adolescents and children in Ealing. If there are similarities between East London MLE and the speech of Ealing adolescents, but the Ealing children do not appear to be acquiring the same variety as the Ealing adolescents, this would suggest that the MLE features are acquired by Ealing adolescents as part of an enregistered youth language, and that this style is available to teenagers but is not acquired or contributed to by young children.

On the other hand, if similarities are found between the East London findings (Cheshire et al., 2011; S. Fox, 2015; Gates, 2019; Ilbury, 2020) and the Ealing adolescents and children, this speaks to a situation of "vernacular universals" and/or universal principles in the selection of features from the Feature Pool (see contributions in Filppula, Klemola, & Paulasto, 2009).

If there appears to be continuity between the speech varieties of Ealing children and adolescents, but they diverge from MLE as described by Cheshire et al. (2011), this could indicate that MLE is not used in Ealing, and instead a different variety is being acquired, e.g. levelling changes taking place in the Southeast; or that a different multiethnolect is developing, influenced by the particular heritage languages present in Ealing.

Finally, where there are differences between Ealing adolescents and East London adolescents, *and* between Ealing adolescents and children, this would suggest that the Ealing adolescents had innovated one or more features. Note that these possibilities are not mutually exclusive, and there may be different results for different variables, i.e. similarities between adolescents and children for one variable, but not for another.

In sum, children and adolescents are compared so that we have a way to gauge whether linguistic similarities between East London and Ealing adolescents are because situations of high linguistic diversity have led to similar outcomes (e.g. in terms of monophthongisation of the diphthongs) in these different locations, or because of diffusion of an enregistered, age-graded, adolescent speech style. If similarities are found between the children and the adolescents, this supports the view that MLE arises because of group second language contact among children. Where differences are found, this points to MLE being a youth style that is adopted in adolescence.

1.6 Outline of thesis

The rest of the thesis is structured as follows:

- Chapter 2 presents the study design: selection of fieldsite; methods for collecting adolescent data; methods for collecting child data
- Chapter 3 describes the fieldsite and participants.
- Chapter 4 gives an overview of the data, describing and giving examples of the MLE features found to be in use in Ealing.
- Chapter 5 gives background information on the diphthong variables and explains the methodology for the sociophonetic analysis.
- Chapters 6 and 7 present the results of the sociophonetic analysis: first, just the intra-adolescent results; and secondly, the results for adolescents and children.
- Chapter 8 is about the adolescents' use of the Arabic borrowing *wallah*. This has been attested in several other multiethnolects across Europe, but had not been picked up by prior sociolinguistic studies of British English. The analysis presented here treats *wallah* as one of a set of epistemic phrases, inspired by the work of Lehtonen (2015).
- Chapter 9 is the discussion and conclusion.

1.7 Chapter summary

This thesis is about MLE, a multiethnolect – although, as we have seen, that term is contentious, it is used here "with due caution" following Cheshire et al. (2011, p.153). The study of multiethnolects grew out of work on ethnolects, and some of the same issues – particularly whether structural vs. repertoire approaches are more appropriate – are inherited from that field. This means that the study of multiethnolects is also linked in important ways with research in language contact and bilingualism. Cheshire et al. (2011) find Mufwene's (2003) concept of the "Feature Pool" particularly helpful in explaining the innovations found in London. Previous research projects in London have indicated that MLE is acquired early by children, and potentially arises through group second language acquisition. This means that it is important to keep in mind the respective contributions of children and adolescents as we investigate MLE in a different area of London, Ealing.

Chapter 2

Study design & data collection

2.1 Introduction

This chapter lays out the study design for the thesis.

This project was carried out within the paradigm of variationist sociolinguistics (Tagliamonte, 2012). So far as possible, the aim was to replicate the methodology of Cheshire et al. (2011), i.e. carry out speech recordings of 16–19-year-olds and children aged <8 in a different area of London, far away from the East London boroughs where previous work on MLE has been carried out (Cheshire et al., 2011; S. Fox, 2015; Gates, 2019; Ilbury, 2020). The following chapter, Chapter 4, describes the process of selecting a fieldsite, and describes that fieldsite in more detail.

Section 2.2 sets out the key details of the study design. Section 2.3 describes the methodology for collecting speech data from adolescents. Section 2.4 gives the methods used to gather data from the child informants, including the design of the modified Diapix task (Baker & Hazan, 2011; Granlund, 2015).

2.2 Study design

2.2.1 Age ranges to be recruited

It was desirable to recruit adolescents in the age range of 16–19 years for comparability with Cheshire et al.'s (2011) Hackney study. This age range is appropriate for other reasons too. Adolescents have been central to many sociolinguistic studies (Kirkham & Moore, 2013) and "Teenagers are the innovators and movers and shakers of language change" (Tagliamonte, 2016a, p.xiv). This generation is generally acknowledged to be at the forefront of linguistic innovation.

The MLE project included children in the age groups 4–5 and 8–9 (Cheshire et al., 2011, p.157). For the current project it was decided to recruit children aged 5–7, or the year groups 1 and 2 in the British school system. This is so that children who had no exposure to English before school had had at least one year of schooling in the British system at the time of recording.

2.2.2 Selection of variables

Diphthongs were selected as the main variable for the study, as they are supposed to be one of the most salient features of MLE (Cheshire et al., 2011; Kerswill, Cheshire, Fox, & Torgersen, 2013; Kerswill et al., 2008). As we saw in Chapter 2, a key finding by Cheshire et al. (2011) was that in Hackney, children as young as 4–5 had acquired the same vowel system as adolescents: we will see if this finding can be replicated in Ealing.

At the time the study was designed, FACE and PRICE had been selected as the key variables; later, GOAT was added. This means that in the design of the modified Diapix task (see Section 2.4 below), keywords including FACE and PRICE were included in the design.

2.3 Methodolodgy: adolescent data

2.3.1 Introduction

This section will situate the methods of the current study in relation to methods used elsewhere in sociolinguistics. It reviews the sociolinguistic interview and alternative methods before describing the approach taken in the current study.

2.3.2 Data collection in sociolinguistics & the Sociolinguistic Interview

Sociolinguistic interviews were used to gather adolescent data. The sociolinguistic interview (SLI) has been the method of choice in variationist sociolinguistics since the field's inception, having been first used in the New York study of Labov (1966). The SLI is designed to minimise the "observer's paradox", and to tap the speaker's unmonitored vernacular, which is supposed to be the most automatic style in that speaker's repertoire, and therefore the most internally systematic (Labov, 1972).

However, the SLI has also been the focus of criticism over the years. Some critiques target the culturally specific nature of the SLI (Briggs, 1986). Other issues include the artificiality of the SLI as a speech event (Wolfson, 1976). Relatedly, studies over the last three decades have revealed limitations in the SLI methodology with regards to eliciting a speaker's repertoire of styles, including the vernacular (Rickford, 2014). Rickford and McNair-Knox (1994) demonstrated that one teenager's use of vernacular variables was determined by the in-group/out-group status of the interviewer. Finally, Sharma (2011) demonstrated that even in interviews where interviewee and interviewer were of the same ethnicity, the SLI failed to capture the range of the speaker's repertoire of styles.

Many sociolinguistic projects, especially in Third Wave studies, use ethnographic methods, and this necessarily leads to a different kind of interview from the traditional "one-shot" SLI (Mendoza-Denton, 2008, p.222). Kirkham (2013) used a modified version of the sociolinguistic interview and describes this as an "ethnographic interview". These interviews took place 12 months after Kirkham began fieldwork and the interview protocol contained questions about teachers the teenagers particularly liked or disliked, fights they had taken part in or witnessed, and accents.

Other studies have triangulated traditional interviews with other methods. The project by Drummond (2018b) involved placing small recorders on tables in a Pupil Referral Centre, giving them to adolescents to use to record themselves, and also leaving them with centre managers while the researchers were not at the fieldsite. They recorded "as many contexts as we practically could" (Drummond, 2018b, p.81) including classes, youth forum meetings, one-to-one conversations with the young people, between-class activities, and peer recordings. Opsahl (2009) split interviews into two sections: for the first part, there was an interviewer present with the two adolescent participants, but in the second part, the adolescents were left alone in the room.

Self-recordings have increased in popularity as an alternative or complement to the SLI (Snell, 2010; Rampton, 1995; Sharma, 2011). While they are arguably the most ecologically valid method of data collection in sociolinguistics, self-recordings have the disadvantage that only a certain amount of information can be gleaned from conversations which the analyst was not present for. In SLIs, especially if they are transcribed as soon as possible after the recording takes place, the interviewer has access to contextual information which is unavailable from self-recordings. In this respect, self-recordings can yield less information, unless they are combined with follow-up interviews. Combining self-recordings and follow-up interviews produces extremely rich data (e.g. Rampton, 1995, 2011); unfortunately, this requires extra time both for collection and for processing of the data.

2.3.3 Adapting the sociolinguistic interview design for the current study

Like other studies, the current project adapted and borrowed from the sociolinguistic interview so as to develop an interview protocol suitable for the target population.

2.3.3.1 Location

Labov (1972) states that the SLI should take place in the participant's home. This was not possible due to ethical considerations, as the sample population included under-18 year olds. Interviews were conducted in the youth club in a room which adjoined the main hall. Total privacy is not possible or desirable in a youth centre: the room had an internal glass window, so that people inside were visible from the main hall. The room was known as "the girls' room" but in practice could be used by anybody. It was relatively small (which is desirable, in order to minimise echoing; De Decker & Nycz, 2013) and contained comfy chairs, tables, a piano and some bookshelves. One possible issue with this location was that while some of the participants would regularly go to this room to spend time with their friends, others, particularly from the Studio CofP, went to this room much less frequently – and so the setting was not equally familiar to all informants.

2.3.3.2 Modules and structure

The interview protocol consisted of several modules, including: race and ethnicity, including discrimination; fights; childhood; the local area and growing up in London; music; religion and superstition; future plans; language. In the course of conducting interviews, it became apparent that many participants responded well to questions about political issues which concerned them; meanwhile eliciting personal narratives could be extremely difficult, due to reluctance to "snitch" (cf. Mendoza-Denton, 2008, p.224). In the traditional methodology of the sociolinguistic interview, personal narratives are the ideal kind of data, and "soapbox" speech is to be avoided. In the current study, soapbox speech was far easier to elicit, and I do not believe it should be excluded (cf. Eckert, 2001, pp.120-121). Moreover, these kind of big topics were more successful in eliciting involvement from both participants, whereas in some cases, one participant appeared to zone out while the other was telling a story, if the anecdote was addressed to the interviewer and was already familiar or uninteresting to the other interviewee.

A reading passage was used: *The Boy who Cried Wolf* (Deterding, 2006). This was used because it is less dry than some other widely used reading passages; it is also closer to the kind of texts which might be read aloud in school, therefore potentially eliciting formal classroom speech. Additionally, I had used this passage for my MA research, and so hoped that using it again would facilitate comparisons across datasets in the future.

2.3.3.3 Participants

The Labovian SLI is a one-to-one conversation between interviewer and interviewee. The interviews in the current study were between one interviewer and two interviewees, to facilitate comparability with the Multicultural London English corpus (Cheshire et al., 2011). This is advantageous with the population under study: a one-to-one interview might appear to be too formal a set-up, whereas the presence of a friend is intended to help both interviewees feel more comfortable.

2.3.3.4 Triangulating methods

It was decided to complement interview data with self-recordings, for reasons touched on above. Becker (2013) notes that the decision whether or not to use the sociolinguistic interview depends in part on the variable of interest: for phonological variables, the SLI is extremely useful, but for eliciting syntactic variables, it can be less helpful. Self-recordings seemed the best way to elicit the discourse variable *wallah* – which had been noticed in use among the adolescents during the participant-observation phase of fieldwork (see Chapters 4 and 9) – and others have shown this feature to be difficult to elicit in SLIs (Opsahl, 2009).

2.3.4 Adolescent recruitment

Adolescents were recruited opportunistically, with ethical approval from the Queen Mary University of London Research Ethics Committee. The information and consent form given to adolescent participants can be found in Appendix E.

I visited the youth club two or three times a week during the time when I was conducting recordings and generally interviewees were approached and interviewed on the same day (as opposed to arranging interviews in advance). Unfortunately, this introduces the bias of having a convenience sample. However, this is arguably inevitable in studies with an ethnographic component, as opposed to traditional survey studies in sociolinguistics. It is also a consequence of conducting fieldwork in a youth club, where attendance is fully voluntary, rather than a school.

Participants were offered £10 in return for participating in interviews. This is a delicate issue as rewarding participants can compromise their ability to offer properly considered and fully informed consent: it is unethical to offer participants payment if it may incentivise them to take part against their better judgment. In the end, it was decided to pay participants to thank them for giving up their time to help with the research. This seemed reasonable for various reasons. The young people taking part were aged 16 and above, and therefore could have been using this time for studying or doing a part-time job. The recordings will benefit my career and to the extent that I will profit from them, it seemed fair to give the participants recognition of their assistance in the project. In other scientific disciplines, payment for taking part in research is standard practice.

However, a few recordings into the data collection process, I adopted the policy of asking people if they would be willing to take part without mentioning the payment to them. That way, they initially agreed to take part on a purely voluntary basis, and became aware of the reward only when reading through the information sheet and consent form. Overall, I am satisfied that the reward was sufficient to be meaningful, but that participants were not taking part only for the sake of the money.

No payment was offered for creating self-recordings. This was because anyone carrying out a self-recording had to inform those around them of what they were doing, and go ahead only if their interlocutors assented to the recorder being used. There was a risk that offering payment in return for self-recordings might have incentivised participants in a way that would lead them to go against the wishes of others present for the recordings. An adapted consent form was given by the recording participant to friends present for the self-recordings, so that they could consent to their speech being transcribed, or request for their speech data to be destroyed (see Appendix F).

It has been observed that "Sociolinguists prefer to downplay their interest in language in order to elicit as much unselfconscious speech as possible" (Eckert, 2013). This was my initial approach in entering the community and beginning to conduct interviews; however, it later became apparent to that potential research participants were more suspicious of me (e.g. as being a potential undercover police officer) when they did not know the exact purpose of my presence. When I explained that the project was about youth language, this was almost always met with a positive response. Many young people in London are proud of their ability to speak "slang", and this perhaps made them more amenable to the idea of a university researcher studying their language.

2.3.4.1 Equipment

The interviews were recorded on a Zoom H4n recorder. Each participant was recorded on an audiotechnica lavalier microphone which fed into one channel of the stereo recording, recorded as a 16-bit 44.1 kHz WAV file. This stereo file could then be split into two mono .wav files during analysis, creating a separate mono recording for each participant. For the self-recordings, an H2n Zoom with lavalier mic was used. Participants were offered a bumbag to keep the mic in, but all of them opted to carry it in a pocket or in their own bag.

2.4 Methology: child data

Ideally, one could use exactly the same data collection strategy with children as with adolescents. However, the needs of these two populations are sufficiently different that this could not be the case here. While sociolinguistic interviews were used to collect adolescent speech data, the sociolinguistic interview cannot be used on children without quite significant adaptations (Roberts, 1997a).

2.4.1 General considerations for recording child data

The length and timing of recordings are key considerations when analysing child speech. With adults, it is desirable to have one long interview, on the assumption that speech becomes less monitored as they settle into the recording session. However, with children, multiple recording sessions are preferred. This is firstly because children produce less speech than an adult might in the same amount of time (Tagliamonte & Molfenter, 2007). Secondly, the recording sessions cannot be as long as they would be for adults (O'Shannessy, 2014). However, in a synchronic study, the time lapse between the recordings must also be short enough that the children's language acquisition has not progressed significantly between the first and last recordings. Roberts (1997a) "interviewed" children multiple times during a 4-month fieldwork period. Infants in a study by Starks and Bayard (2002) took part in three sessions over a 1-month period. When psycholinguistic studies aim to gather completely naturalistic infant speech, recordings of spontaneous speech typically involve recording the child for about 1 hour every 1-2 weeks over the course of a year (Tomasello & Stahl, 2004). Tomasello and Stahl (2004) estimate that this type of sampling is adequate for more frequent phenomena, but leads to problematic over- or underestimation of children's errors when it comes to the acquisition of less frequently occurring language features.

2.4.2 Child data collection in sociolinguistic research

In comparison to the number of studies of adolescents' and adults' language variation, there have been relatively few studies of children in sociolinguistics. Consequently, whereas the sociolinguistic interview stands as a tried-and-tested method for recording speech production data from adults, there is no such go-to method for collecting sociolinguistic data from children.

2.4.2.1 Interview

Some studies have used an interview-like task for collecting data from children. Roberts (1997a) created a "play-interview session" for the purposes of studying children's acquisition of the Philadelphia short-a pattern. The props used were such that children's play would probably involve a lot of speech: children were prompted to make up a story using picture-books; there was a toy telephone. And for the purposes of eliciting the key variable, short-/a/, a bag of pictures was used: children could take a picture from the bag and say what it was, then replace it (Roberts, 1997a). Stanford (2008) conducted "interviews" with the child sat on the mother's lap.

2.4.2.2 Naturalistic recordings

Payne (1976) gathered data from middle-class families in King of Prussia, Philadelphia. According to (Labov, 1989, p.89), "Payne's methods in the King of Prussia study (1976) developed considerable familiarity with the many families involved, and her interviews profited from intimate social relationships with parents and children". This involvement with the families allowed Payne to get large quantities of speech from the children (aged < 10). Similarly, Tagliamonte and Molfenter (2007) used family involvement to their advantage: the first author was able to record her children in naturalistic settings in the home. Smith et al. (2013) instructed caregivers to record their interactions with the child in a variety of routine situations, such as at mealtimes, trips in the car, walks in the park, and while playing.

2.4.2.3 Peer-to-peer speech

For the current study, it is highly desirable to record peer-peer speech: the hypothesis of Cheshire et al. (2011) is that for many children in London, other children constitute their main source of English-language input; they suggest that children turn from parental influence to peer influence in language acquisition at a younger age than has been found in more monolingual communities (Kerswill & Williams, 2000a). Spontaneous speech between peers may be recorded by means of a portable microphone or radio microphone. Snell (2010) successfully used this approach with children aged 8–9.

2.4.3 Data collection in child language acquisition

Recording methods typically used in acquisition research stand in contrast to those used in variationist sociolinguistics. In variationist work, the aim is to establish the range of linguistic variation, both within and between speakers; in acquisition research, the aim is to eliminate variability, or to elicit categorical variation, by establishing whether or not a child has acquired a particular contrast or rule (Roberts, 2002). As a consequence, many of the more popular methods used in establishing children's phonological acquisition are experimental, whereas the traditional methodology of variationist sociolinguistics – the sociolinguistic interview – is observational. For related reasons, acquisition research in phonology tends to focus on consonants, partly because of the complexity of analysing child vowels (see further Section 5.3.3), and partly because there is so much dialectal variability in the English vowel system (Core, 2012, p.80).

2.4.3.1 Co-operative tasks

Some studies have used a cooperative task to elicit peer-peer speech from children (see Kirschner and Tomasello (2010) for use of a cooperative task in a non-linguistic study of child development). Granlund (2015) made use of the Diapix task (Baker & Hazan, 2011) to gather spontaneous speech from 9-year-olds, and supplemented this with a grid task.

The Diapix task (Baker & Hazan, 2011) is essentially a spot-the-difference task. The two participants each have a picture but cannot see the other's picture. They must describe what is in their own picture and find out how it is different from their interlocutor's. The researchers can design the paired pictures in such a way as to include "keywords" that contain target variables: in Baker and Hazan's (2011) study, the authors were interested in stop contrasts such as /p–b/. According to the Baker and Hazan (2011), "both participants are encouraged to work collaboratively and to contribute equally to the conversation. This means that a roughly equal amount of speech should be produced from each participant and the types of conversations that are elicited are closer to natural communication than the 'instruction giving/receiving' feature of the map task" (p.763).

Figure 2.1: Baker and Hazan's (2011) Diapix scene Park A

2.4.4 Child recordings in the current study

For the current project, it was desirable to elicit both speech that was as naturalistic as possible, so as to capture as great a range of the children's sociolinguistic variation as possible, but also to include more controlled elicitation tasks, to ensure that the data in-

Figure 2.2: Baker and Hazan's (2011) Diapix scene Park B

cluded tokens of the variables of interest. For this reason, a modified version of the Diapix task was designed.

In total, four different types of recordings were made: self-recordings in the classroom; self-recordings during the lunch hour; a Diapix training session, using Baker and Hazan's (2011) Park picture set; and another Diapix session, using the Diapix pictures adapted for this study. These recordings took place over a 6-month period from September 2017–February 2018. The materials and procedures for these recordings are described in the following sections.

2.4.5 Materials

2.4.5.1 Designing the adapted Diapix

Certain modifications had to be made to the Diapix task to make it a viable method for this study (Baker & Hazan, 2011). Firstly, it needed to be modified to include sufficient tokens of the target variables of this study; secondly, it had to be made easy enough for 5-7 year olds.

To assess how difficult the Diapix task would be for the children, as well as to familiarise the children with the procedure, children did the Diapix task together using Baker and Hazan's (2011) pilot scene (the park scene pictured above). This was administered with no modifications to the scene. Although the task has successfully been used on children as young as 9, this stage of the recordings showed that the children found this scene too difficult to complete.

The adapted Diapix task was simpler than the original picture sets in a number of ways. The number of target words was reduced. Baker and Hazan's (2011) Diapix included 36 target words, spread across three picture sets. Different studies attempt to elicit different numbers of target words from young children. Granlund (2015) used 16 minimal pair keywords to analyse three consonant and vowel contrasts among 9-15 year old hearing-impaired and non-hearing-impaired children. McGregor, Friedman, Reilly, and Newman (2002) used 20 object words from the Snodgrass and Vanderwart (1980) picture set to elicit data from 5;4 year olds. In Dodd, Holm, Hua, and Crosbie (2003), children aged 3-6 were asked to name 30 pictures. It was judged best to adopt a conservative approach in setting the number of keywords, given that the subjects were from a multilingual population, while databases such as the Bristol norms are based on monolingual populations (Stadthagen-Gonzalez & Davis, 2006).

For this study, 16 keywords were selected (see Table 2.1): four each for each of the variables FACE and PRICE; and two each for each of the point vowels FLEECE, TRAP, LOT and FOOT.

Participants in Baker and Hazan's (2011) Diapix task were instructed to describe the scene one quadrant at a time. To facilitate this among 5–7 year olds, the task was modified by making gridlines visible on the scenes (see Figures 2.4, 2.5, 2.6, 2.7).

The number of differences was also reduced. Baker and Hazan's (2011) Diapix task requires the participants to find twelve differences. This number was reduced to ten differences, and the differences were spread as evenly as possible between the four quadrants. These differences were also made relatively easy to spot. PhD students piloted the task (see Figure 2.3) and any differences that they took longer than 5 minutes to spot were modified to make them easier for children to spot and describe.

Finally, in the adapted Diapix, no writing was visible in the images. The original picture sets had writing visible in most of the scenes, e.g. on signs and in speech bubbles. This was removed because the participants in the current study were at the early stages of learning to read, and so differences in literacy levels between the child participants could have influenced the results.

2.4.5.2 Target words

The following factors were taken into account when selecting target words for the modified Diapix task.

Age of acquisition Each word's age of acquisition (AoA) was checked in the Bristol norms and Kuperman databases (Stadthagen-Gonzalez & Davis, 2006; Kuperman,

Figure 2.3: PhD students piloting the Diapix task

Stadthagen-Gonzalez, & Brysbaert, 2012). Words were chosen with an AoA ≤ 4. Although the children in this study were aged 5–7 at the time of the Diapix task, the database norms are based on monolingual populations. It is known that while bilinguals' vocabularies typically grow at the same rate as monolingual infants, this means that at any one stage, they may have fewer words in one language than children with monolingual competence in that language (Nicoladis, 2006). For this reason, a slightly conservative cut-off was set for the age of acquisition of the target words.

Frequency, imageability and familiarity Where possible, the familiarity and imageability ratings of words were checked in the Bristol norms (Stadthagen-Gonzalez & Davis, 2006). Their frequency per million words in the British National Corpus (BNC; Davies, 2004) was checked and this was highly variable across the words chosen (see Table 2.1). In the cases of words which came up less frequently in the BNC, such as *kite*, this was not deemed a problem because of the low AoA rating and high familiarity rating of such words.

Word structure The words were ideally of a CVC structure and monosyllabic. Sometimes there was a tradeoff between word structure and another factor, e.g. imageability. The CVC structure meant that the target vowel was contained in a closed syllable, and not liable to coarticulation effects from surrounding words when produced in continuous speech. As to the consonants, plosive and fricative consonants were permitted. Clusters were dispreferred, but allowed for the sake of other factors. Words containing nasal, liquid and approximant consonants were not included. This is for the practical reason that plosives and fricatives are the most easily visible consonants on a spectrogram, and so they improve the reliability of the vowel segmentation. The latter class of consonants also have more pervasive coarticulation effects than stop and fricative consonants.

Table 2.1: No. of phonemes, no. syllables, age of acquisition, imageability, familiarity and BNC frequency for the keywords selected for the modified Diapix task. Empty cells indicate that the keyword had no rating for that factor.

Word	Vowel	Phonemes	Syllables	Age of acquisition	Imageability	Familiarity	Frequency
cake	FACE	3	1	4.4	624	594	2700
gate	FACE	3	1	4.4	632	532	3398
baby	FACE	4	2	3.84			8480
table	FACE	5	2	4.39			19128
kite	PRICE	3	1	4.1			701
five	PRICE	3	1	4.51			39453
spider	PRICE	5	2	3.43			648
tiger	PRICE	4	2	4			870
sheep	FLEECE	3	1	4.18	641	484	2942
cheese	FLEECE	4	1	4.41	592	588	2504
cat	TRAP	3	1	3.68			3788
hat	TRAP	3	1	3.33			2872
dog	LOT	3	1	3.62	636	598	7780
sock	LOT	3	1	2.94			938
football	FOOT	6	2	4.82	597	565	6536
book	FOOT	3	1	4.47	591	643	24,142

2.4.5.3 Creating the images

To create the modified Diapix task, two images from the farm selection of scenes were used. The background was kept but all objects were removed, and objects representing the target words were added one by one, and distributed as evenly as possible across the four quadrants. If object pictures were modified, they were tested on five children who were not taking part in the study, to see if they could correctly identify the target word. Images for the keywords were selected and modified as shown in Table 2.2. The adapted scenes are shown in Figures 2.4, 2.5, 2.6 and 2.7.

Table 2.2: Sources and modifications made to the images used in the modified Diapix task

Word	Source	Modifications
cake	Snodgrass and Vanderwart (1980) picture 42	Modified by adding colour and decoration
gate		Drawn in Photoshop
baby	MultiPic picture 507 (Duñabeitia et al., 2018)	
table	Diapix	
kite	Snodgrass and Vanderwart picture 129	Colour was added
five		The Arabic numeral, typed using Photoshop
spider	MultiPic picture 38; IPNP picture 416 (Szekely et al. 2004)	Modified by adding colour.
tiger	MultiPic picture 98	
sheep	Diapix	
cheese	http://pluspng.com/png-26211.html	
cat	Diapix	
hat	MultiPic picture 690	Modified by changing the colour
dog	MultiPic picture 707	
sock	MultiPic picture 416	Modified by changing the colour

football	MultiPic picture 185	
book	MultiPic picture 505	

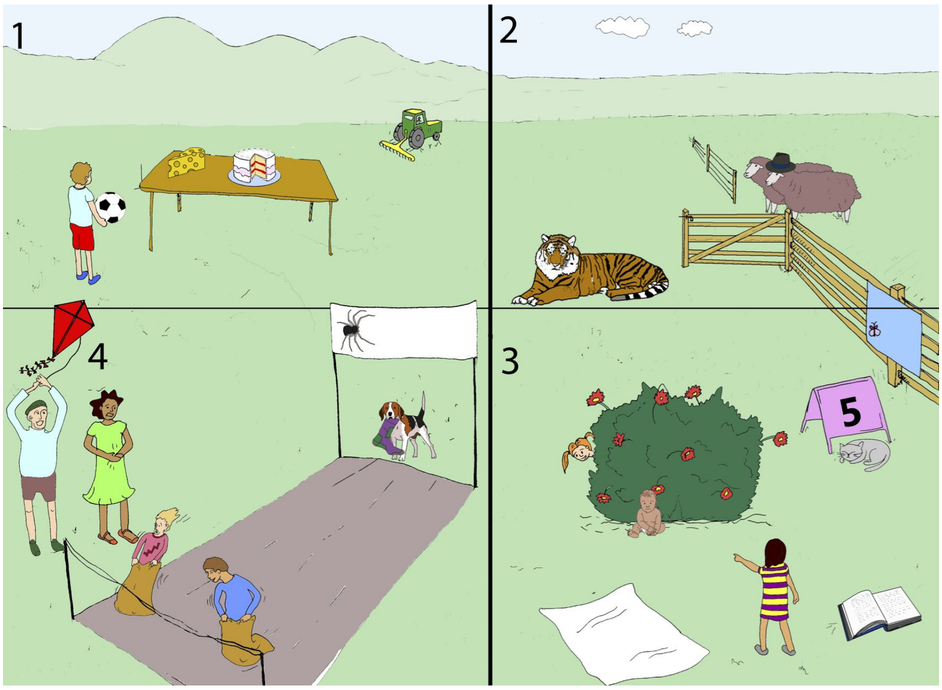


Figure 2.4: Modified Diapix: image 1A

2.4.6 Recruitment and ethical considerations

Consent was sought from parents. Parents were approached in person and contacted by letter (see Appendix C). Assent was sought from children prior to the recording by showing the child a sheet with three faces on it (representing happy, upset and angry) and asking him or her to select which face best represented how s/he felt about taking part in the recording. It was also established with the child that if they changed their mind and wanted to take the recorder off at any point, this was allowed. In particular, when children used the portable Zoom recorder during the lunch hour, I made sure they knew where to find me if they had any problems during the recording. Children were rewarded with stickers for completing the Diapix task.

2.4.7 Language background questionnaire

A questionnaire was used to establish which languages were present in each child's home environment – this can be found in Appendix D. Where possible, these questions were administered over the phone or in person (Hoff & Rumiche, 2012, p.306). Parents were

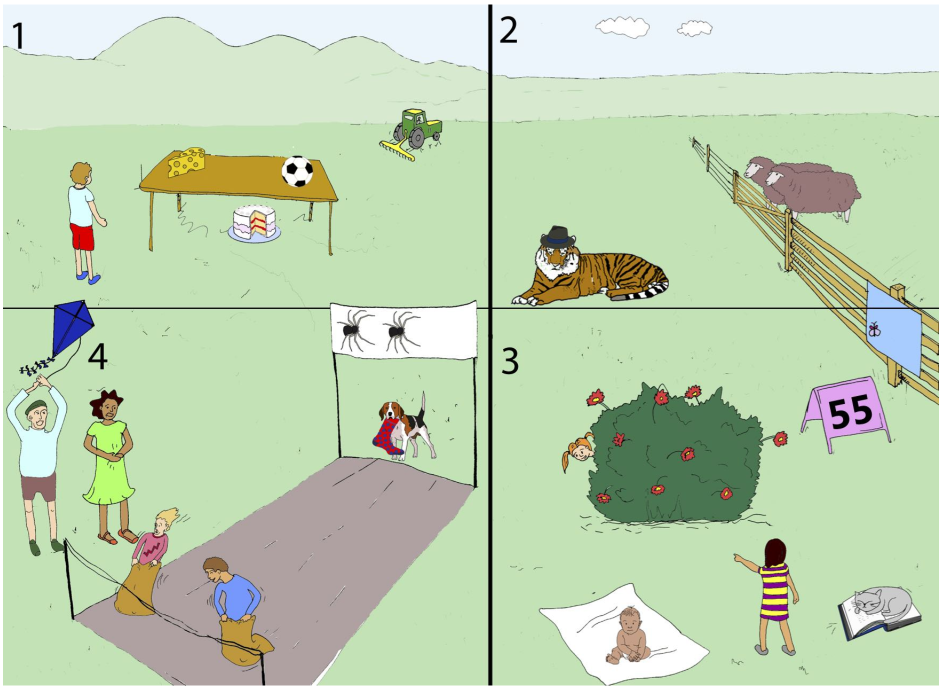


Figure 2.5: Modified Diapix: image 1B

not required to report language exposure as a percentage, though some did this anyway. The results of this questionnaire are presented and discussed in Chapter 3.

2.4.8 Recording procedures

The equipment was kept constant between adolescents and children so as to control for this factor when comparing between the two populations. In the Diapix training and modified Diapix tasks, two children at a time were recorded using a Zoom H4N mic and audiotechnica lavalier microphones. As with the adolescents, the two children were recorded into separate channels of the stereo .wav file, which could then be split into separate mono .wav files for analysis. For the self-recordings, the same Zoom H2 recorder was used as had been used with the adolescents for their self-recordings.

2.4.8.1 Diapix training and adapted Diapix

Children did the Diapix task in self-selected pairs. Almost all of the children took part in two sessions: a training session, using the Diapix pilot scene, the park scene; and a second session in which the one of the modified Diapix scenes were used. Recordings took place in an "intervention space" in the school: this was a small room between two adjoining classrooms, usually used by teaching assistants to work with one or two children at a time. These rooms were ideal for speech recordings because they were small, thus reducing echo, and the doors were heavy firedoors, providing some level of soundproofing.

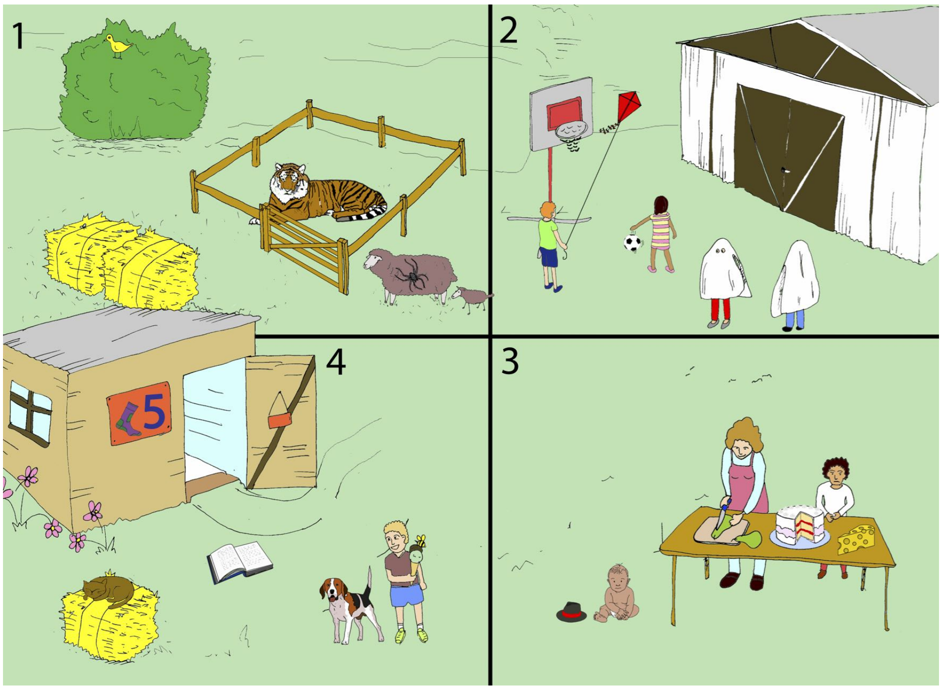


Figure 2.6: Modified Diapix: image 2A

From a safeguarding point of view, myself and the children were visible to teachers in the adjoining classrooms through small windows in the doors.

In the first session, before the Diapix training began, there was an interview-like component to the recording session. The two children were asked questions such as what they did for their last birthday, what languages they could speak, whether they had brothers and sisters, what games they liked to play in the playground or at home, and whether they had a religion. This part of the recording was intended to be as informal as possible and an opportunity for the children to chat freely.

The aim of the training session was to familiarise the children with the recording procedure and the specific rules of the Diapix game. The children sat opposite one another at the table. First, I worked through an example with them, using one of the Diapix street scenes. The children were asked if they could see anything that was different between the two pictures, and received positive feedback when they found differences. After this had been done, the children found the differences in the Diapix park scene, this time with a barrier placed between them so that they could not see each other's pictures. They were instructed to talk to each other to find six differences between their pictures. The recording equipment was covered with a bag or papers, so as not to be a distraction to the children. If they wanted to, they could stand up to look over the barrier at one another, but they were not allowed to look at each other's pictures. They were each given a pencil so that they could circle differences when they found them. A time limit was set of 10 minutes, and after this time, the barrier was removed so that the children could look at the

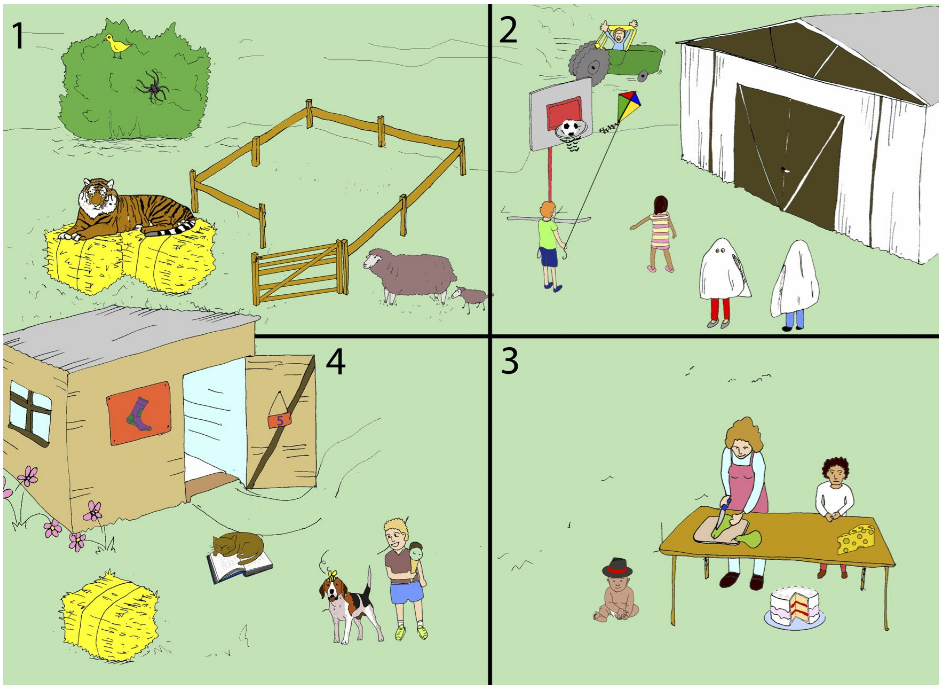


Figure 2.7: Modified Diapix: image 2B

pair of scenes and find the remaining differences together.

The second Diapix session took place within a few weeks of the first. The set-up was the same, but the scenes used were the modified Diapix scenes that had been designed for the current project. Before the task began, children were shown the Diapix pilot scene as a reminder of the activity they had completed in the previous session. They were instructed to find 10 differences. The children were given a reward grid and told that they would receive a sticker every time they found a difference between their pictures.

2.4.8.2 Self-recordings

For the spontaneous speech recordings in this study, a child-friendly bumbag was used to hold the microphone. This bag was colourful and resembled a crab, making it attractive to children and easily visible to myself and school staff.

The child would wear this bag as they went about their normal activities, either in class or during the lunch break. The child's assent having been obtained, the child was told that while they were wearing the bag, the microphone would hear what they were saying: they could do everything as normal and pretend it wasn't there. In practice, of course, many children chose to speak directly into the microphone at least some of the time (cf. Snell, 2010). I established with them where I would be during the lunch break – either at the edge of the canteen, or in a particular spot near the gate to the playground – so that they knew where to find me if they ran into any difficulties or found the microphone uncomfortable and wanted to take it off. I chose these locations so that the child would always know where to find me, but also so as to reduce my influence on the child's play and conversations as far as possible.

2.5 Summary

This chapter has laid out the study design and described the data collection procedures used in the fieldwork for this project. Data collection among the adolescent population proceeded through participant-observation, interviews and self-recordings. To gather speech data and language background information about the children, a parental questionnaire was used, and the children were recorded via self-recordings and in a more controlled setting, completing a Diapix-type task (Baker & Hazan, 2011) that had been developed specifically for this study.

The next chapter describes the fieldsite where the project was conducted. The methodology for acoustic analysis is explained in detail in Chapter 5.

Chapter 3

The fieldsite and the participants

3.1 Introduction

This chapter will describe the fieldsite and the process of carrying out fieldwork. It sets out the time frame over which fieldwork was conducted (section 3.2) and describes the process of selecting a fieldsite and then entering the community (sections 3.3 and 3.4). The theory and method of taking an "ethnographic perspective" (J. Green & Bloome, 2005) and carrying out participant-observation are described (section 3.5). In section 3.6, the social landscape of the fieldsite is described. Section 3.7 presents evidence for two distinct Communities of Practice (CofPs) existing in the youth centre.

3.2 Time span of fieldwork and different components

Table 3.1: Timeline for the components of fieldwork

Time	Activity
October 2016–December 2016	Travelling to different boroughs of London and visiting youth centres there in order to choose a fieldsite
January 2017–April 2017	Participant-observation in the selected youth centre
May 2017–September 2017	Sociolinguistic interviews and self-recordings in the youth centre
April–July 2017	Observation in the primary school
September 2017–February 2018	Recordings in the primary school

3.3 How the fieldsite was selected

It was necessary to find a fieldsite which was comparable to East London boroughs (Newham, Hackney, Tower Hamlets) where previous research on Multicultural London English had been conducted. Residents of the boroughs Hackney, Newham and Tower Hamlets have already been part of recent sociolinguistic studies (Cheshire et al., 2011; S. Fox, 2015; Gates, 2019), and so a fieldsite was sought further afield, in West London or south of the River Thames. The key criteria were linguistic and ethnic diversity, and socioeconomic deprivation – circumstances which are argued to lead to the emergence of multiethnolects (Cheshire et al., 2011, pp.152–153).

Much of London is multilingual and multiethnic. In the 2011 Census, over a third of London's residents had a country of birth outside the UK (Greater London Authority, 2012). 45 of the 50 most ethnically diverse wards in the UK are in London (Greater London Authority, 2013). Nine of the top 20 are in Newham and 6 in Brent, making these the most ethnically diverse boroughs. In terms of multilingualism, "22.1 per cent of Londoners list a language other than English as their main language, a total of 1.73 million people" (Greater London Authority, 2013, p.1). Polish is the language with the highest number of speakers, and "Bengali is the most spoken Asian language in London while Somali is the most spoken African language" (Greater London Authority, 2013, p.1). The boroughs with the highest numbers of people reporting a main language other than English were Newham (ranked 1st with 41.4% of its residents reporting a main language other than English), Brent (2nd, 37.2%), Tower Hamlets (3rd, 34.2%), Ealing (4th, 33.9%)and Westminster (5th, 30.8%) (Greater London Authority, 2013, p.5).

As to deprivation, London is a mixed bag, with rich and poor communities existing side by side. On the Index of Multiple Deprivation, the most deprived areas in London in 2015 were in Hackney, Islington, Westminster, Haringey and Tower Hamlets (Leeser, 2016). Also in 2015, the boroughs with the highest percentage of people working for less than London Living Wage per hour were Barking and Dagenham (43%), Newham (40%), Brent (39%) and Ealing (36%) (Office for National Statistics Annual Population Survey, 2015). The boroughs Barking and Dagenham, Tower Hamlets and Hackney rank the highest for average income deprivation, while Tower Hamlets, Islington, Barking & Dagenham, Hackney, Newham, Lambeth, Southwark and Lewisham rank among the top 10 boroughs in England for average deprivation affecting children (Leeser, 2016).

I decided to find a suitable fieldsite in Brent or Ealing (both West) or Lambeth (South). In October and November 2016, I took regular bus trips around these areas, usually between three and five o'clock in the afternoon, the time when British school children leave school, and I would also go into high street shops and fast food restaurants. I also contacted youth clubs in these boroughs as a way to get to know local areas. It seemed practical to base the choice of fieldsite partly on the linguistic and ethnic diversity of the borough, and partly on which youth centre would be most convenient for carrying out sociolinguistic research.

The next section will give an overview of demographic data for the borough of Ealing, where fieldwork was eventually conducted, and the section after that will describe the youth centre where fieldwork was carried out. Ealing will later be described from a more emic perspective in section 3.6 below.

3.3.1 Ealing

Ealing has the third highest population of all the London boroughs (Mangara, 2017). It is also the third most densely populated of the Outer London boroughs (Mangara, 2017). In the 2011 Census, Ealing was the 3rd most diverse borough in England and Wales. Compared to the rest of England and Wales, Ealing had the largest Polish population, the 2nd highest number of Japanese residents, 2nd highest number of Iranians, 3rd biggest population of Somali and Somalilanders, and 4th highest number of Arabs (Mangara, 2017, p.11). In addition, "Traveller groups have frequented Ealing for many years. Department of Environment, Food and Rural Affairs (DEFRA) 'caravan counts' consistently record Ealing as having one of the largest Traveller populations in the Greater London area" (Ealing Grid for Learning, 2017).

This diversity is also seen among the borough's children and young people. In 2016, "[t]he most common ethnic groups in Ealing's school population are white British (15%), Indian (14%), Eastern European (10%), Somali (8%), Pakistani (7%), Asian Other (7%), Afghan (4%), Arab Other (4%) and black Caribbean (4%)" (Mangara, 2017, p.13). In terms of languages spoken, "[a]t the time of the 2011 Census, around a third (35%) of pupils in Ealing's primary schools spoke English as a first language, whilst in the high schools the figure was 45%. Pupils in Ealing schools speak over 100 different languages and the 10 most common languages spoken are: English, Polish, Punjabi, Somali, Arabic, Urdu, Tamil, Persian/Farsi, Gujarati, Pashto/Pakhto (in order of the numbers of speakers)" (Mangara, 2017, p.14).

Different wards of the borough differ greatly in terms of wealth and deprivation. Some areas of the borough score low on measures of deprivation and high on measures of wellbeing, while the opposite is true of other areas. In the 2011 Census, "Ealing had a slightly higher percentage of homeowners, both outright and with a mortgage, than London" as well as "lower levels of social renting than London and England" (Mangara, 2017, p.10). However, the distribution of home ownership differed between different ethnic groups. "In terms of ethnicity, residents of white and Asian ethnicity were over-represented amongst people who owned their own home and under-represented among those using social rented housing [...] Conversely, residents of black ethnic origin were over-represented among people in social rented housing and under-represented amongst people who owned their own home" (Mangara, 2017, p.10).

Overall, Ealing was relatively less deprived in 2015 compared to 2010, i.e. the borough performed better relative to the other boroughs in England and Wales (Mangara, 2017, p.37). However, the 2011 Census found that Ealing was the 18th most deprived borough in the country: among the indicators of household deprivation, Ealing had 37% of households not deprived in any dimension, 35% deprived in one dimension and 28% deprived on multiple indicators (Mangara, 2017, p.35). In 2015/2016, Ealing had a rate of 17.5 homeless people per 1000, performing "worse than both London and England in both measures of homeless published in the Public Health Outcomes Framework" (Mangara, 2017, p.33). The ward where the youth centre was located was one of the poorer areas of the borough, and has significantly lower life expectancy than the national average (Mangara, 2017).

3.4 Gaining access

3.4.1 Entering the community and selecting a youth centre

It is known that "the most difficult step in any sociolinguistic study is almost certainly entering the community for the first time" (Schilling, 2013, p.177).

Many sociolinguists have accessed adolescent or child subjects via a school (e.g. Eckert, 2000; Kirkham, 2013; Snell, 2010; Mendoza-Denton, 2008; Lehtonen, 2015; Madsen, Karrebæk, & Møller, 2015; Alam, 2015), and this approach has certain advantages. Attendance is compulsory and the school will be attended by pupils who live within a geographical catchment area, meaning that the sample is likely to be representative of the local population. Also, the researcher is guaranteed to see the same group of pupils day after day, which facilitates the building of relationships. However, at the same time, as an adult, it is extremely difficult to align yourself with students, rather than being positioned by them as a member of staff, and researchers will often take pains to avoid being seen as an authority figure. For example, Lehtonen (2015) took the decision not to be given keys to the teachers' room. Mendoza-Denton (2008) chose to eat lunch in the cafeteria with students, rather than accepting an invitation to the staff room. Eckert had to stop attending students' classes, and make contact with students in other ways, because the class teachers were too prone to position Eckert as a fellow teacher (Eckert (2000), cited in Schilling (2013)). Moreover, although in the past ethnographic projects have been successfully carried out in schools, I learned from colleagues that nowadays, UK secondary school students' time is tightly circumscribed, leaving them little room to talk to each other or to a researcher during class time. Finally, schools close for the Christmas, Easter and summer holidays, as well as the half-term breaks, meaning that no fieldwork can be carried out during those gaps.

By contrast, attendance at a youth club is voluntary. This makes it difficult to be sure of following one group of participants. At the same time, however, the nature of the youth centre as a social space makes it more likely that teenagers will use vernacular speech. The ratio of social time to adult-guided activities will be much better than in school. It also reduces the risk of being positioned as an authority figure because while youthworkers are responsible adults, ultimately, their authority over children is less than that of a teacher. Finally, many youth clubs stay open over the summer, meaning that research does not have to pause during school holidays.

3.4.2 Deerpark Youth Centre

From October 2016 to December 2016 I attended three different London youth clubs. The managers of all three were sympathetic to the research aims and very helpful in sharing their knowledge of the local area. Deerpark is the pseudonym for the youth centre that was eventually chosen as a fieldsite. I chose this Ealing youth club for several reasons, after having visited others in Brent and Lambeth.

Deerpark had the highest number of open hours per week. Whereas the other clubs opened only in the evenings, Deerpark was open from 12pm until 6pm most days, and until 9pm other days. The other clubs also restricted who could attend on certain days or the type of activity on offer, e.g. having a sports night, a film night, boys' night, girls' night, mixed night etc. Related to this, Deerpark took the most hands-off approach in terms of organising activities and also discipline. Other youth centres would each night have a set of activities organised and led by youthworkers; at Deerpark, youthworkers were there for attendees to talk to if they wanted, but the youthworkers also took care to give the adolescents their own space. Sometimes activities were offered at Deerpark – such as basketball, or even archery with rubber arrows – but joining in was a matter of choice. Probably as a consequence, Deerpark was much better attended by adolescents than the other two youth centres I visited.

The dynamic and understanding between adolescents and youthworkers was very different at Deerpark: at other youth clubs, teenagers were not allowed to enter if youthworkers suspected that they had been smoking cannabis. At Deerpark, teenagers were not allowed to smoke cigarettes indoors, but could do in the outside areas. Individuals caught rolling up cannabis were simply asked to take it off the premises; youthworkers didn't do anything as drastic as banning attendees or confiscating the drugs. The priority was teenagers' safety: the youthworkers' stance was that it is better for teenagers to spend as much time as possible in a safe space such as the youth centre, rather than spending time on the streets. Members were only banned if they did something that violated the youthworkers' trust, such as vandalising the building, or if it was felt that they put other members in danger in some way.

It is known that there are drawbacks, but also a lot potentially to be gained, from entering the community via a "broker" (Schilling, 2013; Levon, 2013). Entering the community in this way was necessary for the current project, as the target population was young people (Schilling, 2013). I was fortunate in that the youth workers, and in particular the manager of the youth centre, truly acted as brokers in this project. I had been primed to avoid forming too close an alliance with the youthworkers, so as to avoid being seen as an authority figure; however, the trust between youthworkers and attendees, and the attendees' relative mistrust of strangers (e.g. in case an unfamiliar adult turns out to be plain-clothes police) meant that it was an advantage to be affiliated with youth centre staff. The manager in particular introduced me to key figures in the different peer groups at the youth centre and gave me advice on how to integrate myself, as well as assisting with recruitment for the interviews.

3.4.3 Entering a school

Once Deerpark Youth Centre was confirmed to take part in the project, primary schools in the nearby area were contacted. It was desirable to secure the participation of a school within the same borough, and ideally in the same ward, so that the adolescent and child samples would be representative of the same population. The primary school that agreed to take part was located on the same estate as the youth centre. While there are no immediate family connections between the adolescent and child samples (i.e. no siblings), many of the adolescent and child participants were neighbours. Some of the adolescent participants reported having attended this primary school as children and similarly, some of the children reported having older siblings at the high school that many of the youth centre members attended.

In the summer term, i.e. May–July 2017, I observed classes in the primary school in order to build familiarity with the children and the teachers. At this time, the child participants were in Reception and Year 1. Recordings were carried out the next academic year, September–February, when the participants were in Year 1 and Year 2.

3.5 "an ethnographic perspective"

The adoption of ethnographic methods in the current study was necessary for a few different reasons. Firstly, to make use of recent advances in sociolinguistic theory: sociolinguistic theory, as characterised by Eckert (2012), has proceeded from its beginnings in an ethnographic approach, to a "retreat from ethnography" (p.88), to a re-adoption of ethnography as a methodology. "The ethnographic studies of the second wave provided a local perspective on the findings of the survey studies of the first wave, making the connection between macrosociological categories and the more concrete local categories and configurations that give them meaning on the ground" (Eckert, 2012, p.93). According to Eckert, a local perspective is necessary in order to progress beyond the first two waves, by seeking to understand how language, speakers and social space co-construct one another.

In particular, the ethnographic method of participant-observation was necessary for the sake of building familiarity with the fieldsite and potential research participants in order to give higher quality data and increase the validity of the findings. When the researcher is an outsider, as I was, participant-observation can allow the researcher to carry out "intrusive acts of data collection"; it also "reduces the problem of reactivity-of people changing their behavior when they know that they are being studied"; it "helps you ask sensible questions" (important for conducting sociolinguistic interviews); and it "gives you an intuitive understanding of what's going on in a culture and allows you to speak with confidence about the meaning of data ... [i]t extends both the internal and the external validity of what you learn from interviewing and watching people" (Bernard, 2011, pp.265-266).

The title for this section comes from J. Green and Bloome (2005). They differentiate three different ways of incorporating ethnography into one's methodology: there is *doing ethnography*, which "involves the framing, conceptualizing, conducting, interpreting, writing, and reporting associated with a broad, in-depth, and long-term study of a social or cultural group, meeting the criteria for doing ethnography as framed within a discipline or field"; adopting an *ethnographic perspective*, by which "we mean that it is possible to take a more focused approach (i.e., do less than a comprehensive ethnography) to study particular aspects of everyday life and cultural practices of a social group. Central to an ethnographic perspective is the use of theories of culture and inquiry practices derived from anthropology or sociology to guide the research"; and finally, it is possible to *use ethnographic tools*, which involves "the use of methods and techniques usually associated with fieldwork. These methods may or may not be guided by cultural theories or questions about social life of group members" (J. Green & Bloome, 2005, p.4). I would situate the current study within the second of these options: this is not a "comprehensive ethnography", but the research is guided by "the use of theories of culture and inquiry practices derived from anthropology or sociology", and is more than just using techniques associated with fieldwork.

3.5.1 Participant-observation in the youth centre

One key ethnographic tool used in the current study is participant-observation. Participant-observation means “getting close to people and making them feel comfortable enough with your presence so that you can observe and record information about their lives” (Bernard, 2011, p.256). The data collected can include field notes, photographs, audio recordings and interviews, among others (Bernard, 2011). The current study involved the writing of field notes and following up references to music (e.g. the teenagers’ own rap videos) or popular culture that were observed in the field (Levon, 2013).

Participant observation involves a tension between immersion and distance. “Participant observation involves immersing yourself in a culture and learning to remove yourself every day from that immersion so you can intellectualize what you’ve seen and heard, put it into perspective, and write about it convincingly” (Bernard, 2011, p.258). This means that different modes of participation must be employed, and “ethnographers must routinely alternate between moments of high involvement and moments of low involvement in the activities that surround them” (Duranti, 1997, p.102). When necessary, the ethnographer gets involved with whatever activity is going on, and when possible, the ethnographer finds and occupies a “blind spot” that allows him or her to observe the activities at hand without being involved. “The underlying rationale for finding the blind spot and trying to be as unintrusive as possible is not to pretend that one is not there, but to get as close as possible to what it is like to be a marginal participant” (Duranti, 1997, p.102).

Before beginning to recruit participants for sociolinguistic interviews, I carried out approximately four months of participant-observation. In the early months (January–April 2017), I attended evening sessions, which seemed to be the busiest, as well as sometimes afternoon sessions in order to observe what went on in the recording studio – the sound engineers did shifts starting at 12pm, and young people could book sessions with an engineer to record their music. In the summer, I would attend during the daytime when the youth centre was somewhat quieter, as it was preferable to conduct interviews with as little ambient noise as possible.

While at the youth centre, my participation was often passive, especially in the beginning. Later on, I would approach individuals to talk to them. I kept field notes and early on, I took to writing these on my phone, as this made me stand out less than if I had been writing in a notebook. I hoped that this would help me blend in, as many of the attendees had their phones in their hands for the majority of the time that they were sitting or standing in one place. In particular, in the music studio, youth would enter and take a seat, and then would be using their phone until it was their turn to go into the recording booth. I would then type up the field notes at the end of each session. At other times, I participated fully: for example, by taking part in a barbeque in the summer; playing table tennis or pool with youth club attendees; occasionally baking or cooking with individuals in the kitchen.

3.5.1.1 My position as researcher

> "The strongest constraints that prevent linguists from utilizing the wealth of linguistic data with which they are surrounded are the barriers against interaction with strangers in one's own culture" (Labov, 1972, p.110)

Due to my social and ethnic background, I entered the community as an outsider. This is not unusual in linguistic and anthropological work: ethnographers are usually "wealthy and powerful individuals who have only a temporary and in many respects very limited interest in the community they study and live in. Beyond ethnographers' intentions, motivations, or awareness, there are political and global processes that enter into the relationships they establish in the field" (Duranti, 1997, p.95). This probably means that developing rapport with research subjects was harder than it would have been for an insider, as there were greater social barriers to be overcome.

In terms of my own background, I grew up in Cheltenham and Durham in the UK, and grew up in a White middle-class family. Probably as a consequence of my educational background (at grammar school and then Oxford University), I speak Southern Standard British English (SSBE). My appearance, my accent and entire habitus therefore marked me out as coming from a background of privilege. To make matters worse, an additional problem was the gentrification going on in the area (see section 3.6.1 below). At the time, being White and sounding posh aligned me with the people moving into or even buying the luxurious new flats being built nearby.

My age and gender affected the data in different ways. I began the fieldwork for the project aged 23. Similar to Mendoza-Denton, I had "status as a barely older youth at the time" (2008, p.40) – indeed, one of the interviewees in the project was two months older than me. My gender in particular will have affected how I was positioned by youthworkers and young people, as well as the nature of the recorded interviews. For example, a male colleague's gender had the apparent consequence that his interviews conducted in Hackney contained far more tokens of the *man* pronoun than mine did (Ilbury, 2020).

3.5.2 Observation in the primary school

Observation was carried out in the primary school in the summer term of the academic year 2016-2017, i.e. May–July.

This is described as "observation" rather than "participant-observation" because it was not as prolonged or as in-depth as the participant-observation in the youth centre. The purpose was to build familiarity with the children and also with the teachers, in the hopes that they could assist with recruitment of parents to give consent on behalf of their children. I spent one morning a week in the school, attending the children's classes and accompanying the children outside at breaktime. A field diary was kept to record observations from this period.

3.6 The fieldsite: Sherburn Estate

3.6.1 The estate

The youth centre where fieldwork was conducted was in the middle of what had formerly been one of London's poorest estates. At the time of fieldwork, the estate was in the process of regeneration. As part of a trend of new housing developments and gentrification across London, a partnership between Ealing council and a private development company was overseeing the transformation of Sherburn Estate into Sherburn Mews. Interviewees concurred in telling me that crime on the estate now was a fraction of what it had been even ten years previously. In the following extract, Moses presents the transformation of the estate into the Mews as broadly positive – "this is a good place now" – but also locates the agency behind this transformation as being with an anonymous "they". It is implied that the estate is not being improved, but rather being replaced by the Mews: "they took that off the map completely".

(1) Moses:
it was it was/ Sherburn was the Gaza (0.20) this isn't Sherburn Estate no more this is Sherburn Mews (0.37) this is a good place now (0.82) they didn't want Sherburn Estate on the map anymore . they took that off the map completely

Many interviewees were not optimistic about the regeneration. They claimed that the council wanted to clear them off the estate, and implied unfairness in the process of moving the old residents:

(2) Chris:
the council have to like offer you . a thing yeah .but, they'll make it like a shit deal (0.46) (RO: right) so you – then, you'd rather get the option of a like a three bedroom . house in ((area)) . or one bedroom (0.60) (RO: oh I get it) or two rooms, and a room will be boxes (0.97) (RO: yeah) (0.59) and you got like what, four three kids

Chris says that it is a legal obligation for the council to offer tenants the chance to stay on the estate but, because they would actually prefer the tenants to move away, "they'll make it like a shit deal", whereby tenants are offered a suitably-sized house in a different area, or an impractically small space in their current neighbourhood.

The incoming homeowners and renters were also described as chiefly middle class White people:

(3) "they're mainly white"

RO	((xxx)) what dyou think of the new ((people moving in))
	(2.28)
Chris	they're mainly white
	(0.14)
RO	yeah
	(2.97)
Chris	they are – I I haven't – actually not gonna lie I haven't seen anyone
Chris	of other
Chris	race
	(0.34)
Chris	part from white people

And interviewees hinted at tension between the newcomers and the old tenants, as in Example 4:

(4) Lola:
there is kinda division [...] I'm ((gonna)) say us cos, it's us (Amanda: mm) the Sherburn Estate lot (0.74) to, the people that are just moved in like in these rich houses . there is a division I feel like they just look down on us (0.81) feel like . we're just . they're just – we're just the downstairs people that we're just there . like, they look . upon us

Interviewees generally expressed nostalgia for the old estate. In Example 5, Amanda and Lola laugh as they recount how "scatty" the estate used to be. Lola says twice "it was our home".

(5) "it was just our home"

Amanda	it used to be so run down this area like if you think now it's bad
Lola	[laugh]
	(0.46)
Amanda	don't know how it was before
Amanda	((it)) was so /scatty/
Lola	/it was/ disgus- I can't believe what we lived in
	/but it was just our home/
Amanda	/me too, it was/ so normal to us like
Lola	it was our home

3.6.2 Ethicity and race

On the face of it, Deerpark youth centre was a thoroughly harmonious multiethnic community. Many of the 16 year olds, when asked about issues of race, interpreted the question as referring to the context of the youth centre. They answered that racism was never a problem, and everyone got along fine. The older adolescents who perhaps had had experience in the working world often told quite different stories, as in the following example:

(6) Hudson:
when I was going – when I went carnival, in Notting Hill, I was with a big group of friends, but . it's weird like you're gonna smell – you smell – you smell drugs all in the air, all in the air. But it's like this police officer single-handedly, out of everyone in the cr- – everyone, like every – white, black, asian, arab anyone – came up to me and said 'come with me', put me against the wall and said – and put me in handcuffs and basically said like 'we're gonna like, basically look to . arrest you' And I was – and I said – (( )) obviously I had my rights and I said like 'what for' and they said 'cos we feel like you're selling drugs', so I got searched . and everything but . at the end of the day I had nothing (RO: mm) . like I was clean, and they single-handedly picked me cos they smell [weed] in the area, and they thought 'okay look young black boy'

The "Arab boys" were recognised as a category in the youth centre. One participant, Chris, complained to me about how these boys would use the N-word:

(7) "it just shouldn't be said innit"

Chris	I don't like like w- like when
	(0.24)
Chris	like, some little arab kid says "nigga"
Chris	cos like it's a bit weird innit
	(0.06)
RO	what dyou mean, like, in what – can you give me an example
Chris	like
Chris	they'll be just be saying "my nigga" to like one of their friends and they're not, like, you can't say that
Chris	but that doesn't mean we can say it
	(0.46)
Chris	but like, it just shouldn't be said innit

Despite the apparent heterogeneity of the youth centre attendees, White minorities were notably underrepresented. When one digs a little deeper into the stories told by adolescents in the interviews, hints of racial tension in the community emerge. Ealing historically has large White Irish, traveller and Polish populations, yet none of those who took part in the study were Polish, and while there were three girls with Irish heritage, there were no boys.

Irish and traveller people had a reputation locally for being racist. For example, in Example 8, Tariq and Sami relate an incident involving "Irish boys":

(8) "three little Irish boys"

Tariq	/the last/ time I've heard
Tariq	someone say nigger, or anything that's racist to your skin colour
Tariq	was three little um Irish boys
RO	/really?/
Sami	/oh yeah/ the (( )) /I swear/
Tariq	/[*place name*]/
Tariq	three little /Irish boys/
Sami	/you know them pikies/ like, ((that's)) racist anyway
Tariq	yeah
Sami	um Irish gypsies
Tariq	Irish gypsies
RO	oh /okay/
Tariq	/like they live in/ a cara- cara park, caravan park and all that
RO	right
Tariq	them ones yeah
RO	yeah yeah
Tariq	three little ginger boys yeah
Tariq	they came up yeah
Tariq	and then they were like, "you nigger"

One particularly memorable fight in the teenagers' recent memory started as a conflict between an Arab boy and a Polish boy, which resulted in several boys being excluded from the local high school. One source of conflict in the youth centre was the presence of two younger White girls who were described by the older adolescents as being of Irish heritage. On one memorable afternoon, there was an altercation between these two sisters and an older girl, who then brought her sisters to the youth centre to fight the younger girls. Referring to what had happened, one boy said to me "What do you expect? They're from an Irish family"; meanwhile, the older girl who started the fight was served a 6-week ban from the youth centre for calling one of the White girls a "white slag".

3.6.3 The Somali community

Sherburn Estate was home to a sizeable Somali community. For example, in an interview, one girl reported "Sherburn, everyone knows is known for Somalia, we even have a road called Somali Road [...] the Somali community here's very big, very very big". Similarly, in Example 9, Amanda and Lola report that "everyone picks up the lingo".

(9) "everyone wants a bit of Somali"

Lola	everyone wants a bit of Somali, like you can hear like
Amanda	yeah
Lola	people just sitting there, they'll say something in Somali. it's just
Amanda	[laugh] yeah everyone picks up the lingo

This suggests that certain Somali phrases may be commonly enough heard that non-Somalis pick them up and use them in a situation of language crossing (Rampton, 1995). There is one small glimpse of this in the interviews:

(10) "*bahal*"

Amanda	you picked up a Somali word, what's that B word
	(1.13)
Shantel	bahal
	(0.28)
RO	what's that
	(1.00)
Amanda	means like . bug . insect like

The size and visibility of the Somali community also made stereotypes and jokes about Somalians available to the teenagers. In the following example (11), Tariq and Sami explain how *Somali* is available as a "cuss". This might involve calling someone "Abdi", "fat forehead" or simply saying "shut up you Somalian" – Sami offers "shut up you're Algerian" to show how calling someone Algerian does not work as a slur, whereas "shut up you Somalian" rolls off the tongue.

(11) "there's a lot of Somali jokes"

Sami	there's a lot of Somali jokes though I won't lie to you
Tariq	ye:ah
Sami	(( ))
Sami	fat foreheads and like
Tariq	there's a lot of Somali jokes
Sami	like have have you ever heard someone say=

Tariq	Abdi as well
Sami	=“shut up you’re Algerian”
Sami	/like people use/ the word Somalian as a cuss
Tariq	/have you ever heard of –/
RO	/(( )) can you give me an example/
Tariq	/legit.
Tariq	like for example you say/
Sami	/“shut up you Somalian”/
Tariq	“shut up you Somalian” or
Tariq	“shut up Abdi”

3.6.4 Postcode wars

In the media in recent years, there has been much hype about London’s “postcode wars”. One key point of difference between the CofPs described below is their respective orientations to being “active” on the streets, including involvement in postcode wars. I would like to stress that I am not claiming that anyone who participated in this study is involved in postcode wars or so-called gangs (the two things often overlap in the newspaper reports) – rather, the focus is on a cultural orientation (Drummond, 2018b; cf. Mendoza-Denton, 2008). The shape gangs take in the popular imagination and the reality of who can be convicted for gang activity are quite different: “definitions of gangs by the police and by teachers expand in widening circles, and throw a blanket of suspicion over minority youth who have very little to do with the criminally propelled concerns under the traditional definitions of gangs” (Mendoza-Denton, 2008, p.79); under the Serious Crime Act 2015, to be convicted as being part of an “organised crime group”, one needs to be in a group of minimally three people which “has as its purpose, or as one of its purposes, the carrying on of criminal activities” (Home Office, 2015).

The following excerpt shows the effect of “gang” labelling as a problem, encouraging immature youth to embrace the identity projected onto them:

(12) Moses:

and then because they’ve labelled you as a gang (0.25) and, you’re young and you’re active (0.22) you decide . “well cool. they’ve called me a gang (Daniel: [laughs]) (0.12) let’s have a name for this gang” (1.64) and then yeah (0.32) you get a name for the gang, and normally from then . there’s another gang with another name . that, have just established themself as well . as a gang (0.12) and it’s like “who are they . fuck them, dadadada” sorry I swore you can ex that out . and it’s just like yeah boomboomboom and then next minute you know you’re in (0.12) gang rival beef or whatever and it's just because (Daniel: mm) (0.64) you knew some people who knew some people who knew some people . and you was out every day tryna find something fun to do (0.33) and fun coulda been anything when you was younger and and kind of more immature (0.22) it coulda been, running from the police coulda been fun (0.29) it coulda been (1.42) playing, tag in a high school (0.24) like

The locus of turf wars in London is the postcode, and street violence is usually referred to as "postcode wars", a term which the interviewees also used. They described the importance of knowing which postcodes were rivals and which were allies, and how this might affect journeys they themselves wanted to make within London. A problem many young men face is that you can be identified as a probable affiliate of the local gang simply because of where you live. Hence the question "Where are you from" is not one to be answered lightly. In Example 13, Raphael explains the manoeuvres he would make to avoid being caught in a "mad area at the wrong time".

(13) Raphael:
say if there's a party and like . uh a area like in south like say [...] and you're coming back say especially like . when you're by yourself . you get me, ((you xx)) if you know you're in a (0.80) mad area at the wrong time . just gotta ((boy)) (1.05) get me that's what I'm saying when I when (0.47) (xx) when I'm going to . but with me it's like when I go to parties like (0.65) that are in east london or south london, yeah (0.54) and I'm not I'm not gonna be coming home, I'll either go to . Tony's house, cos he lives in south innit (0.34) so I'll either go back to his house or when I'm in east I'll go back to my brother's house or something, I'm not gonna . try waste my time . and and move around at a risky time and come home, you get me

Denzel recounts similar experiences in Example 14:

(14) "you're telling me I can't come to a certain area"

Denzel	anywhere. 's no – no one can tell me where I cannot go
Kai	/yeah/
Denzel	/I can't/ go
Denzel	cos I just look at you like rah you don't own this country like
Kai	/yeah [laughs]/
Denzel	/you – you come on like/
Denzel	I'm black we're both black and you're /telling me I can't/ come to a certain area

3.6.5 Music

The relationship between music and youth culture in London is not straightforward. In recent times, music has been blamed for inciting violence among young people. In April 2018, then Home Secretary Amber Rudd stated "gangs are posting videos and music online that document, encourage and glamorise violence and goad and threaten others. ... today I am calling on [social media companies] to review their terms and conditions and make it clear that that they will not host any content linked to gangs or gang violence" (Rudd, 2018). Following this, in June 2018, the West London drill group 1011 was banned from posting music videos online without police permission (Cobain, 2018).

During fieldwork, I saw two sides to this story – chiefly as represented by the two Communities of Practice (described below, section 3.7). Many of the young people who I will describe as the Youthclub CofP in the next section would listen to music by self-proclaimed London gangs such as the 410s and Harlem Spartans, rapping along with their lyrics. On the other hand, a theme among youth who came to the youth club to record their own music (the Studio CofP) was that music had been a saving grace in their lives, removing them from the problems of street violence. Either explicitly or implicitly, these youth position trouble on the streets in opposition with a focus on making music (see further Examples 23 and 22).

The recording studio existed in the youth centre because of the positive effect involvement in music could have in young people's lives. One of the few areas where the youth centre enforced a strict disciplinary policy was that of music production. As part of their job, the sound engineers would refuse to record a track if they considered the lyrics too violent, or if a direct threat to a rival group were included in the song. For example, in my field notes from December 2016:

> "[*Engineer*] telling one of the boys that he wasn't willing to put a gunshot sound effect on the track, because having that on the track implies that the boy is going to shoot someone"

3.7 Communities of Practice in the youth club

"A community of practice is an aggregate of people who come together around mutual engagement in an endeavor. Ways of doing things, ways of talking, beliefs, values, power relations – in short, practices – emerge in the course of this mutual endeavor" (Eckert & McConnell-Ginet, 1992, p.464). As the participants engage with each other in these activities, this comes to define their relationships to one another, their personae and their world views, as well as reproducing and potentially redefining the norms of the activity itself. The concept of Community of Practice is useful for capturing the dialectic between language as a shared set of norms and the speakers who speak that language, seeing the two as mutually constitutive: the speakers acquire the norms of language use, but themselves are constantly reinforcing and/or shifting those norms as they use language. It also speaks to a similar dialectic between language and cultural practice, with language a component of culture, but also the means by which culture is realised and potentially altered. “The value of the construct community of practice is in the focus it affords on the mutually constitutive nature of individual, group, activity, and meaning” (Eckert, 2000, p.34). The Communities of Practice perspective has been used productively in various sociolinguistic studies (e.g. S. Fox, 2015; Eckert, 2000; Kirkham, 2013; Alam, 2015).

During my time at the youth centre, I noticed two distinct purposes for attending. Some youth attended the youth centre in order to socialise with friends, while others attended in order to use the music studio and record their own music – typically (but not always) rap. These gave rise to two distinct CofPs within the overarching CofP that was Deerpark. The properties of these two sub-CofPs are summarised in Table 3.2. The traits of these two CofPs are not mutually exclusive. So for example, while the Studio CofP were passionate about music, the Youthclub CofP would play music on the loudspeakers in the main area and rap along. The members of the Studio CofP would sometimes use the main youth centre space and play table tennis, pool or Xbox. The difference is in the primary purpose for attending: the Youthclub CofP members attended primarily to see their friends, while the Studio CofP members attended primarily for the sake of using the studio.

It is a feature of CofPs that any one speaker will be concurrently a participant in different, sometimes intersecting CofPs: “These communities of practice may be more or less overlapping, more or less interacting, more or less consonant ... The individual’s identity emerges in the process of articulation and resolution of participation in all of these communities of practice, and each community of practice’s identity emerges through its participants’ joint engagement in this process” (Eckert, 2000, p.36). So for example, in my data, while all of the youth were participants in the overarching Deerpark CofP, some were involved in the smaller CofP within that that I have called the Youthclub CofP. Of those Youthclub CofP members, some were also Muslims, which meant together, they were participants in a different CofP with a different set of shared practices, such as fasting during Ramadan. The way in which these young people resolved their participation in these latter two CofPs, which in some of their norms contradict one another, played a key role in determining the shared norms of the Youthclub CofP.

To this extent, categorising individuals as members of CofPs is a reification, but a useful heuristic for the purposes of this study. Eckert (2000, p.33) says of the theoretical concept of the speech community, “The claim that the social unit that defines one’s sociolinguistic sample constitutes a speech community, then, is above all a way of placing the study itself rather than the speakers." Arguably this is true to a certain extent of CofPs as well: the definition of distinct CofPs is a simplification that belies the complexity of real people's social lives, but is helpful to the purposes of the study.

The key characteristics of the two CofPs are summarise in the table below, and elaborated in the following sections.

Table 3.2: Summary of the distinguishing features of the two CofPs

	Youthclub	Studio
Gender	Boys and girls	Boys
Age	16–17	17 and over
Ethnicities	Arab, Somali, Black African, Black British, Black Caribbean, White British, White Irish	Black British, Black, African
Residence	Local/West London	Local and non-local
Religion	Muslims, Christians	
Primary activities	Smoking outdoors, hanging out around the youth centre, playing Xbox, table tennis, pool	Recording music in the studio
Street orientation	"active"	Orient away from "road life"
Music	Grime, drill – local influences	Hip-hop, rap – American influences

3.7.1 Youthclub CofP

The Youthclub CofP is how I have categorised the young people who came to the youth centre primarily to socialise. In terms of social spaces and activites, they would hang out in the main hall of the youth centre, in the kitchen, in the office with youth workers, or outside at the back, where they were allowed to smoke. The boys would sometimes ride their bikes or Boris bikes to the youth centre. Smoking cigarettes and smoking weed carried covert prestige for the teenagers and sometimes in winter, the teenagers would test the boundaries of what they were and weren't allowed to do by sitting next to an open door to smoke. Within this CofP there were to an extent smaller groups of friends, but not polarised the way opposing cliques may be in school – all of these adolescents had chosen to spend their free time in the youth centre.

Some of the members knew each other by virtue of having grown up together on Sherburn Estate. When asked how they knew each other, many would say that you just recognise faces that are familiar from the area – you come to know one another by being in the same neighbourhood. Many of them attended the local secondary school. They were aged 16–18 at the time of the recording – many of them took part in the interviews relatively soon after their 16th birthday.

One key feature that differentiates the Youthclub CofP from the Studio CofP is their orientation towards being "active" on the streets. This goes hand-in-hand with the local outlook of the Youthclub CofP members, with their social lives being tied to their area, and many expressing a sense of loyalty to this area (cf. Brent, 2012). Some of the members of this CofP positively oriented towards being part of a group with a "resistant identity" in much the fashion described by Moses above (Example 12). This was alluded to by Sqara as follows (Example 15).

(15) Sqara:
like, if someone disrespect us (0.31) . (( )) be disrespecting my area because . I'm living here . someone come and disrespect my area start breaking stuff breaking that breaking that (0.14) of course he(('s)) disrespecting me (0.15) . cos I live in that area (RO: mm) (0.05) he doesn't live in that area. Why he's doing this (0.94) and . and sometimes, this, lead to argument or (0.10) or fight

Of course, different individuals had different degrees or kinds of participation in one or the other CofP. While some members of the Youthclub CofP were described as being "active", others (particularly girls) discussed street life in negative terms. In the extract below, Chantelle and Sarah negatively evaluate a group of boys who they describe as "the young gees" as being "too active" (Example 16).

(16) "the young gees"

RO	what like cliques are there
	[...]
Chantelle	oh clicks
Chantelle	oh yeah there's the young gees
Sarah	yeah you got the the young gees /and then you/
Chantelle	/so like CB and that/
RO	okay
Sarah	and then you've got like, X12s which is like MW and that
RO	yeah so what's the difference between
Chantelle	the young gees are /just them/
Sarah	/but/
Sarah	the the young gees are the people /((that)) they/ they look up to these ones

Chantelle	/they're more active/
	[...]
Chantelle	and you have their their their associates like their girlfriends (Sarah: yeah) and female acquaintances
	[...]
RO	um yeah the girls, so the girls who're equi- what dyou mean by equivalent to boys dyou mean
Chantelle	so like
	(0.25)
Chantelle	we've got girls like [*name*] and Jessica
RO	mm
Chantelle	and they're just they /they ride/ for them boys so like
Sarah	/they're they're/
Sarah	they're they'll be equivalent to the young gees /cos that's who/ they chill with
Chantelle	/yeah/
RO	mm
	(0.50)
Chantelle	and then you've got like I bounce between them but I don't really like being with young gees cos they like
Sarah	yeah
Chantelle	they they're too active
Sarah	yeah

Chantelle and Sarah describe younger boys and their girlfriends as being "too active". While they look on this younger group with apparent disapproval, these peer groups are still broadly in the same social networks, and Chantelle states that she will "bounce between them".

Finally, religion was a salient aspect of identity for the Youthclub CofP. As can be seen from Table 3.3, many of these participants were Muslim. This also meant that several of them took part in interviews while fasting for Ramadan, and so religion became a key topic in their interviews. For some more than others, this entailed difficulties in balancing their social world with their religious values. At the time, several Muslim boys were in relationships with non-Muslim girls. This could be a source of friction if both parties in the relationship wholeheartedly believed in the teachings of their own religion. In particular, this sometimes led to clashes over how the girls were expected to behave and dress, with the girls being used to behaving in one way and the boys expecting something different. In Examples 17 and 18, a Muslim boy and a Catholic girl explain the problems of inter-religious relationships.

(17) "she's not of the same culture"

RO	d'your parents know about your girlfriend now
ZR	yeah
	(2.68)
RO	what do they what do they think
	(2.05)
ZR	think
	(1.93)
ZR	they think it's silly, waste of my time
	(1.05)
ZR	she's not of the same culture
	(0.74)
ZR	they think, it'll bring problems in the future
	(1.38)
RO	like what kind of problems
	(1.94)
ZR	problems between different cultures yknow like she's
	(0.74)
ZR	from a dif- very different culture ((of)) mine
	(1.03)
ZR	my one's more strict
	(0.67)
ZR	yknow
ZR	then you gotta come to debate
ZR	what religion the kids they're gonna be
	(0.70)
ZR	how your wife will
	(1.57)
ZR	present herself in front of the public
	(0.67)
ZR	as you know, muslims like to cover up

(18) "keep the religion within the family"

Jessica	he said – I'm not talking about now, but I'm saying if in the future we were still together and we had kids or whatever
Jessica	he's saying they would bor- be born into islam
RO	right
Jessica	and that's not

	(0.71)
Jessica	it's the ma- for religion, it's the male ((is)) dominant
Jessica	that's why the girl –
Lucy	so the children would be muslim
Jessica	that's why the girl's meant to marry someone who's catholic to
Jessica	keep the religion within the family

Table 3.3: Information on Youthclub CofP members. Residence: [1] = same postcode as the youth centre; [2] = Northwest London; [3] = London but outside Northwest London. Empty cells = information not given

Pseudonym	Sex	Age	Grew up in	Current residence	Ethnicity	Religion	Parents' place of birth
MW	M	17	London	2			Uganda
Tariq	M	16	London	2	Somali, Black	Muslim	Somalia
Sami	M	16	London	2	African, Arab	Muslim	Algeria
Joe	M	16	Lebanon	2	Arab	Muslim	Lebanon
Jessica	F	16	London	1	White British, Irish	Catholic	London
Lucy	F	16	London	2	White British	Catholic	London
Chris	F	16	London	1	Black African		Liberia
Shantel	F	16	London	1	White Irish		London
Sarah	F	17	London	1	Black Caribbean, British		London
Chantelle	F	17	London	1	Black Caribbean, British		Jamaica

Table 3.3 – *Continued from previous page*

Pseudonym	Sex	Age	Grew up in	Current residence	Ethnicity	Religion	Parents' place of birth
Khaled	M	16	Lebanon, London	2	Arab	Muslim	Jordan
Ahmed	M	16	Kenya			Muslim	Kenya
Sqara	M	17	Syria		Arab	Muslim	Syria
CB	M	16		2		Muslim	Iran, Afghanistan
Amanda	F	17	London	1	Somali	Muslim	Somalia
Lola	F	17	London	1	Black African	Muslim	Djibouti
Ali	M	16	London	2	Black British African	Muslim	Somalia
Khadir	M	17	London	1	Black African	Muslim	Somalia
ZR	M	16	London	1	Arab	Muslim	Libya, Lebanon
Ibrahim	M	16	London	2	Black	Muslim	Somalia
Omar	M	16	London	2		Muslim	UK

3.7.2 Studio CofP

One of the first things to note is that the Studio CofP members are all male. Girls did occasionally use the studio, and some girls from the Youthclub CofP would come into studio to socialise with Studio CofP boys. But the vast majority of young people using the studio to record their rap songs were male.

The members of the Studio CofP are also older than the members of the Youthclub CofP: the Youthclub CofP were mostly aged 16 and attending secondary school, while the Studio CofP members were older than 16, some even older than 18, and attended college or had already completed college and had full-time jobs. The greater degree of independence that comes with this life stage perhaps explains why some of them travelled from so much further afield than their younger counterparts. The greater maturity of the Studio CofP members was evident in the interviews – for example, they were more likely to take control of the conversation.

The Studio CofP come together around the activity of recording and producing music, typically rap. In Example 19, Kai and Denzel explain that while they do socialise at Deerpark, that is not their reason for attending. Kai explains that he likes to arrive early in order to get things done.

(19) “I’m not really coming here to socialise”

Kai	I get here early and do my thing you know I’m saying
RO	/mm/
Denzel	/yeah/
Kai	/I’m/ not I’m not really coming here to like
	(0.12)
Denzel	socialise
Kai	yeah, like I mean /I mean I mean/
Denzel	/like I don’t/ –
	(0.10)
Denzel	/yeah I know what you mean/
Kai	/like like I’ll/ socialise innit
	(0.12)
Kai	but it’s – I’m not really – that’s not why I’ve come
	(0.34)
Kai	here to so- I haven’t come here to so- I come here to get stuff done innit

In Example 20, this point of view is corroborated by Shantel from the Youthclub CofP: the people who use the studio get on fine with everyone else, but don’t come to the youth centre to socialise.

(20) “there’s people that go ((in)) the studio”

Shantel	like there’s people that go ((in)) the studio
Shantel	and then when like
Shantel	older people come, you just like they . they just all get along really
	(0.39)
Shantel	now
	(1.01)

RO	dyou say dyou think there's a difference between people who use the studio and like . people who /come (( ))/
Shantel	/yeah because/
Shantel	people that use the studio in here, like
Shantel	they don't really socialise with, the people that out here?

Studio CofP members were not as locally-bound as the Youthclub CofP. Whereas all of the members of the Youthclub CofP lived in West London if not the immediate local area, three members of the Studio CofP came from further afield – each of them commuted to the youth centre from a different area of South London. Similarly, their own social networks expand further than just the local area, as explained by Tony and Raphael. To put this excerpt in context, Stratford is on the far East side of London and takes around an hour to travel to by public transport from the youth centre where these young men were interviewed – I was living in Stratford while conducting fieldwork and knew only too well how onerous the journey could be, hence my surprise that these young men were also making the same commute.

(21) "we're not all in the same area"

Raphael	if I go stratford, cos stratford (it's) not even bad, like I will go stratford cos that's where
Raphael	most of . us guys link up to like
Raphael	chill and everything innit
Tony	and go parties like
Raphael	((n)) /go parties/
RO	/((xxx)) stratford?/
Tony	/that's the meet-/
Tony	yeah that's the meet-up spot /because/
Raphael	/that's where the meet-up spot ((xxx))/
Tony	/(like)/
Tony	we're we're ((a)) group of friends innit (0.44)
Tony	but, we're not all in the same area
Raphael	yeah /((xx)) we ((got)) different areas/
Tony	/so like I'm from south/
Tony	he's from west
Tony	they're from east, some are from north
Tony	northwest yeah so we have to meet up innit
RO	how dyou know each other though
Tony	on . social media innit

Excerpts such as these show how these youth have different horizons in comparison to the Youthclub CofP: whereas the "young gees" in the Youthclub CofP strongly affiliate with their local area, the members of the Studio CofP see themselves as being part of London communities.

Related to this, in contrast to some members of the Youthclub CofP (cf. 15), members of the Studio CofP typically distanced themselves from postcode wars and street activity.

(22) Tony:
but that's ((w-)) that's not that's not what I like to focus on I don't like to think about . the gangs and things (0.18) (RO: yeah) I focus on myself, on music . and . my family, friends . social events

Some of them portrayed involvement in street fighting as a past they had left behind.

(23) Kai:
I had a knife because . I was I was young (0.37) and (0.55) still young now innit but like (0.32) I was young like like like my mindset was totally different to the way that it is now . and . I just had problems innit and then yeah . that's that's that's why I carried a knife

The two CofPs also exhibit different music tastes. The Youthclub CofP would listen to music produced by local London grime groups and by London gangs, such as the 410s and Harlem Spartans. These groups produce a type of drill music that has become contentious, as it is claimed that it can incite violence (see section 3.6.5 above). By contrast, the members of the Studio CofP tended to cite more international (e.g. American) music and older rappers as their musical role models.

Table 3.4: Information on Studio CofP members. Residence: [1] = same postcode as the youth centre; [2] = Northwest London; [3] = London but outside Northwest London. Empty cells = information not given

Pseudonym	Sex	Age	Grew up in	Current residence	Ethnicity	Religion	Parents' place of birth
SD	M	20	London	1		Sikh	London
GW	M	20	London	1	Black British		Grenada

Table 3.4 – *Continued from previous page*

Pseudonym	Sex	Age	Grew up in	Current residence	Ethnicity	Religion	Parents' place of birth
Daniel	M	20	Ghana, Ivory Coast	1	African		Ghana
Moses	M	24	London	1			
Matisse	M	19	London	1	Black		Congo
Hudson	M	19	London	1	Black, British		UK
Kai	M	17	London	3			India, Barbados
Denzel	M	18	London	3	Black		Portugal, Angola
Raphael	M	18	London	1			
Tony	M	18	London	3	Black		

3.7.3 CofPs: summary

Here I have presented evidence to support the existence of two distinct, if at times overlapping, CofPs within the Deerpark youth centre. The different kinds of activities that the two communities engage in are important because they connect these young people to wider communities of practice that exist beyond the confines of the youth centre – some of which could be described as Imagined Communities (Anderson, 2006). The Youthclub CofP connects to the Sherburn Somali community, the Imagined Community of Islam, and – at a cultural level – the world of being "active" on the streets. Meanwhile the Studio CofP are more self-evidently engaged in a joint endeavour, i.e. the production of rap and hip-hop music; this in turn connects them to the wider community of everyone participating in those music genres. The Studio CofP members are also older than the Youthclub CofP members. We can expect that these different cultural orientations will lead to differences between the two CofPs in the uptake and use of vernacular features.

3.8 Child participants

21 children were originally recruited. Of these, 3 were excluded from the analysis because they were later discovered to have a speech or hearing issue. A further 4 are not included in the analysis because they left the school before taking part in all of the recording types, or because the sound quality in their recordings was not sufficient for acoustic analysis.

Information about the remaining 14 child participants is given in Tables 3.5 and 3.6. As far as possible, the child's language background is described in the caregiver's own words. In the table, "caregiver 1" represents the respondent – not necessarily the primary caregiver. Not all caregivers responded to all questions, and M4's caregivers did not respond at all.

The children were aged between 5;5 and 7;0 in November 2017, meaning no child was older than 7;3 by the time of the last recording in February 2018. There are two sets of siblings in the data: F1 is the younger sister of M5, and M1 is the younger brother of M6. Note that none of the children comes from a monolingual English background, meaning that this sample is comparable to the "Non-Anglos" in Cheshire et al. (2011).

F1 and M5 F1 and M5 are reported to be acquiring English and Somali, and to be English-dominant. Their mother comes from Mogadishu and their father from Somaliland, regions which have different dialects of Somali (Saeed, 1999) – their caregivers reportedly speak Somali to one another, but is not clear which dialect or whether they use multiple dialects. F1 and M5 are reported to receive input in both Somali and English from their mother, and English from their father; both respond in English.

M5 is the oldest sibling in the family. We know that older siblings typically achieve a higher level of proficiency in the heritage language, while younger siblings receive less input in the heritage language and achieve lower levels of proficiency (e.g. Silva-Corvalán, 2014).

M1 and M6 Less information is available for M1 and M6. Like M5, M6 is the oldest child in the family. Like F1 and M5, their mother comes from Mogadishu, though we do not know which region of Somalia their father was born in.

F3 F3 is the only child in the sample not born in the UK. She is a sequential bilingual, having arrived in the UK only two years before, and having had no exposure to English before starting school. Her caregivers report a monolingual Brazilian Portuguese environment at home.

F4 F4 is the youngest of her siblings, and her mother reports that she is English-dominant. It is possible that she has some exposure to non-London accents of English, as her father lives in Bristol, where she visits him regularly.

M3 M3's home environment is reportedly monolingual Somali and he is an only child.

F7 F7's mother emigrated from Yemen, and reports a bilingual Arabic and English environment at home. It is reported that F7 has sisters, but not whether they are older or younger.

F8 F8's mother reports that her child is acquiring both Urdu and English. English is her dominant language.

F9 F9's mother was born in Kuwait but has also lived in Syria, and speaks both dialects of Arabic; she has lived half of her life in the UK. F9's parents have an explicit "Arabic in the home" policy and send their daughter to Arabic Saturday school (like M7, below). Consequently, F9's mother reports that her daughter's Arabic is very good, although English is still her dominant language. F9 has one older brother, aged 8, and two younger brothers.

F10 lives with her mother and grandmother, who both emigrated from Kenya. Both Swahili and English are in use in the home, though F10 is dominant in English and has limited active competence in Swahili. F10 is the oldest sibling – she has one younger sister.

M7 M7's background is similar to F9's: his mother was born in Kuwait, and he attends Arabic Saturday school. M7's mother reports that both mother and father speak to M7 in Arabic, and he responds in Arabic and English. He is the oldest sibling, having one younger sister.

M8 M8's caregivers were born in Sri Lanka. His mother reports that she uses Tamil and English to speak to her son, and he replies in Tamil; M8's father addresses M8 in Tamil and English, and M8 replies in Tamil and English. M8 is the oldest sibling.

Table 3.5: Information about child participants part 1: place of birth, school year, age in November 2017, information about caregiver 1. Empty cells = information not given

Child	Place of birth	School year	Age in November 2017	Caregiver1 relationship to child	Caregiver1 place of birth	Caregiver1 arrived in UK	Caregiver1 languages	Caregiver1 speaks to child in	Child speaks to Caregiver1 in
F1	London	1	5;6	mother	Mogadishu, Somalia	2013	Somali, English, a little bit of Arabic	70% English, 30% Somali	English
M1	London	1	5;5	mother	Mogadishu, Somalia	2010			
F3	Brazil	1	5;10	father	Brazil	2015	Portuguese, English	Portuguese	Portuguese
F4	London	1		mother	Mogadishu, Somalia	1998	English, Somali, Arabic	Mostly English	English
M3	London	1	5;7	father	Somalia	2003	Somali, English	Somali	Somali

Table 3.5 – *Continued from previous page*

Child	Place of birth	School year	Age in November 2017	Caregiver1 relationship to child	Caregiver1 place of birth	Caregiver1 arrived in UK	Caregiver1 languages	Caregiver1 speaks to child in	Child speaks to Caregiver1 in
M4		1							
M5	London	2	6;7	mother	Mogadishu, Somalia	2013	Somali, English, a little bit of Arabic	70% English, 30% Somali	English
M6	London	2	6;8	mother	Mogadishu, Somalia	2010			
F7	London	2	7;0	mother	Aden, Yemen	2006	Arabic, English	Arabic, English	Arabic, English
F8	London	2		mother					

Table 3.5 – *Continued from previous page*

Child	Place of birth	School year	Age in November 2017	Caregiver1 relationship to child	Caregiver1 place of birth	Caregiver1 arrived in UK	Caregiver1 languages	Caregiver1 speaks to child in	Child speaks to Caregiver1 in
F9	UK	2	6;4	mother	Kuwait	2005	Arabic (Kuwaiti dialect and Syrian dialect), English	Arabic	Arabic
F10	London	2	7;0	mother	Kenya	1997	Swahili, English	English and Swahili	Mostly English, can use a few words of Swahili
M7	London	2	6;7	mother	Kuwait	2010	English and Arabic	Arabic	English and Arabic
M8	UK	2		mother	Sri Lanka	2005	Tamil and English	Tamil and English	Tamil

Table 3.6: Information about child participants part 2: information about caregiver 2, others in household, and whether the child knew English before starting school. Empty cells = information not given

Child	Caregiver2 relation to child	Caregiver2 place of birth	Caregiver2 languages	Language(s) caregivers speak to each other in	Caregiver2 arrival in UK	Caregiver2 speaks to child in	Child speaks to Caregiver2 in	Languages child knows	Did child know English before starting school?	Others in household
F1	father	Somaliland, Somalia	English, Somali	Somali	1992	English	English	English. Understands Somali, can speak some words, not fluent		Mother, father, M5, 2 younger siblings
M1	father	Somalia		Somali						Mother, father, M6, other siblings aged 3 and 5
F3	mother	Brazil	Portuguese	Portuguese	2015	Portuguese	Portuguese	Portuguese, English	No	Mother, father

Table 3.6 – *Continued from previous page*

Child	Caregiver2 relation to child	Caregiver2 place of birth	Caregiver2 languages	Language(s) caregivers speak to each other in	Caregiver2 arrival in UK	Caregiver2 speaks to child in	Child speaks to Caregiver2 in	Languages child knows	Did child know English before starting school?	Others in household
F4	father	Yemen	English, Arabic, Somali			Somali	English	English, some Somali	NA	Mother, 3 brothers, 1 sister. F4 is the youngest.
M3	mother	Somalia	Somali and English			Somali		Somali and English		Mother, father
M4										
M5	father	Somaliland, Somalia		Somali	1992	English	English	English and Somali – similar to his sister		Mother, father, F1, two younger siblings

Table 3.6 – *Continued from previous page*

Child	Caregiver2 relation to child	Caregiver2 place of birth	Caregiver2 languages	Language(s) caregivers speak to each other in	Caregiver2 arrival in UK	Caregiver2 speaks to child in	Child speaks to Caregiver2 in	Languages child knows	Did child know English before starting school?	Others in household
M6	father	Somalia		Somali						Mother, father, M1, siblings aged 3 and 5
F7	father							Arabic, English	Yes	Mother, father, sisters
F8								Urdu, English. Mainly speaks English at home.	Yes	
F9	father	Kuwait	Arabic, English	Arabic	1994	Arabic	Arabic	English is easier for her but her Arabic is strong. Attends Arabic Saturday school		Mother, father, brothers aged 3, 4 and 8

Table 3.6 – *Continued from previous page*

Child	Caregiver2 relation to child	Caregiver2 place of birth	Caregiver2 languages	Language(s) caregivers speak to each other in	Caregiver2 arrival in UK	Caregiver2 speaks to child in	Child speaks to Caregiver2 in	Languages child knows	Did child know English before starting school?	Others in household
F10	grand-mother	Kenya	English, Swahili	Swahili				English, a few words of Swahili		Mother, grand-mother, one younger sister
M7	father	Iraq	Arabic, English		2001	Arabic	Arabic and English	Arabic and English (attends Arabic Saturday school)		Mother, father, younger sister (aged 5)
M8	father	Sri Lanka	Tamil, English			Tamil, English	Tamil, English	Tamil, English		Mother, father, younger brother (aged 4.5)

3.9 Chapter summary

This chapter has described the fieldsite – an area of Ealing, West London. The adolescent participants were recruited via a youth centre, and the children were recruited via a school on the same estate as the youth centre. Participant observation using an ethnographic perspective was carried out in the youth centre, leading to the identification of two CofPs among the adolescents: those whose purpose in coming to the youth centre was to socialise with friends; and those whose time in the youth centre revolved around making music in the studio.

A sample of 14 children – 7 boys and 7 girls – was recruited, aged between 5;5 and 7;3 when recorded. None of them has a monolingual English home environment, making them comparable to Cheshire et al.'s "Non-Anglo" sample. Their heritage languages include Arabic, Somali, Urdu, Tamil, Swahili, Brazilian Portuguese.

Chapter 4

MLE features in use in Ealing

4.1 Introduction

This chapter has three aims. First, to establish which known MLE features there is any evidence of in the current data. Second, to give the reader a feel for the data, as well as an introduction to the key features of MLE. Third, to help future researchers: this chapter takes the form of a list of features found in the data, so that researchers working on Multicultural English elsewhere in the UK have a quick point of reference for youth language in Ealing. This chapter is intended as an overview of all the MLE features found in the data, while Chapters 7–9 analyse the target variables for this thesis in detail.

The chapter is organised according to levels of the grammar. Section 4.2 will describe the key phonetic features of MLE. Subsection 4.2.1 then looks in some detail at which of these features are and are not used by an adolescent boy, an adolescent girl, and a child. There are then sections treating lexis (4.3), morphosyntactic features (4.4) and discourse-pragmatic features (4.5). For each linguistic feature, there is a brief review of MLE research on that feature, and examples from the current dataset, if that feature was found in the Ealing data.

4.2 Phonetic features

This section will introduce the key phonetic features of MLE and briefly review existing literature on these features. In Subsection 4.2.1, I will provide select excerpts from two adolescent participants and one child participant that seem to exemplify high and low users of MLE phonetic features.

The phonetic features of MLE include:

- **Reversal of Diphthong Shift and monophthongisation of the diphthongs.** These features will be analysed in detail in subsequent chapters.

- **GOOSE-fronting**. This is described by Cheshire et al. (2011, p.156) as a "global" feature. The vowel charts from the Hackney adolescents in Cheshire et al. (2011) show a GOOSE vowel that is similar in frontness and height to KIT.

- **k-backing**. Backing of /k/ is an inner-London innovation identified by Cheshire et al. (2008b). It can occur when /k/ is in word-initial position before a non-high back vowel, e.g. *car, cousin*. Although back and front allophones of /k/ are found in English already, the backing of /k/ found in Hackney was more extreme, e.g. to an uvular place of articulation, [q] (Kerswill, 2014, p.433). Like TH-fronting and DH-stopping, k-backing was found to a greater extent among young speakers with multiethnic networks in Hackney, compared to those with largely Anglo networks (Cheshire et al., 2008b, p.7). Cheshire et al. (2008b, p.7) suggest that k-backing "is arguably an ethnic marker" because of "its lower frequency in largely monoethnic Havering".

- **Reinstatement of /h/**. Historically, Cockney is an h-dropping accent (Wells, 1982a). One attested feature of MLE is "full reinstatement of /h/ in lexical words and stressed pronouns" (Kerswill, 2014, p.433).

- **Syllable-timed rhythm**. On the continuum from more syllable-timed languages to more stress-timed languages, English is seen as being more stress-timed. However, one innovation associated with MLE is the use of a more syllable-timed rhythm. Comparing young and old speakers from Hackney and Havering, Torgersen and Szakay (2012) found that the Non-Anglo young speakers in Hackney had the most syllable-timed rhythm, and the speakers in Havering had the most stress-timed rhythm. Torgersen and Szakay (2012) suggest that monophthongisation of the diphthongs, and less vowel reduction, led to reduced durational differences between stressed and unstressed vowels in the speech of young Non-Anglo Hackney speakers. They found differences in the durations of the FACE and GOAT diphthongs between Hackney and Havering young speakers, and found that the young male speakers (both Anglo and Non-Anglo) tended to have a longer schwa duration compared to the other groups of speakers (Torgersen & Szakay, 2012, pp.833–836).

- **l-vocalisation**. English /l/ has a "clear" allophone, [l], when it occurs in syllable onset, and a dark "allophone", [ɫ], when it occurs in syllable coda. L-vocalisation is when coda /l/ is further darkened until it becomes vowel-like, [w]. This is a traditional feature of Cockney speech (Mott, 2012). It is still found in MLE, but is also common across dialects of English (Cardoso, Levon, Sharma, Watt, & Ye, 2019).

- **DH-stopping and DH-fronting; TH-stopping and TH-fronting**. The abbreviation (DH) refers to the voiced interdental fricative /ð/. DH-stopping is the realisation of (DH) as [d] and can occur word-initially, word-medially or word-finally (Drummond, 2018b, p.229). Historically, this was a feature of Cockney (Wells, 1982a, p.329) but nowadays, it is much more associated with AAVE and also Jamaican Creole (Drummond, 2018b, pp.229–230). In the word-medial and word-final contexts, DH-stopping is in competition with DH-fronting, i.e. the realisation of (DH) as [v]; but in the word-initial context, only DH-stopping can occur (Wells, 1982a, p.328). Similarly, (TH) refers to the voiceless interdental fricative /θ/. TH-stopping is the realisation of (TH) as [t] and can, like DH-stopping, occur word-initially, word-medially or word-finally; however, TH-stopping is generally less common than DH-stopping (Drummond, 2018b, pp.229–230). TH-fronting, i.e. the realisation of (TH) as [f], can also occur word-initially, word-medially or word-finally. TH-fronting and DH-fronting are strongly associated with Cockney English (Wells, 1982a, pp.328–329). In MLE, TH-fronting and DH-stopping are found, while TH-stopping and DH-fronting still occur, but at a lower rate (Drummond, 2018a; Cardoso et al., 2019).

MLE has been described as a "repertoire of features" and there are indications that the use of different combinations of features from the repertoire can affect whether a speaker is perceived positively or negatively (Cardoso et al., 2019). Some of the features, such as t-glottaling and TH-fronting, are common across vernacular varieties of British English; while some features are more specific to London, more ethnically marked or are potentially more innovative, such as DH-stopping and k-backing. In a study that compared listeners' perceptions of a number of different UK accents, Cardoso et al. (2019) found that the two speakers who had been judged by the authors to be MLE speakers elicited different evaluations from listeners in terms of professionalism, expertise, hireability and likeability. This was found to correlate with the two speakers' rates of use of different MLE features: the two speakers had similar levels of FOOT-fronting, l-vocalisation, TH-fronting and DH-fronting, which are all found in other accents besides MLE; meanwhile Mark, who tended to be rated higher, had higher rates of FOOT-fronting, which, like GOOSE-fronting, is a supralocal feature; while Eric, who tended to get lower ratings, had higher rates of k-backing, DH-stopping, FOOT-backing and GOAT-backing.

4.2.1 Examples of the MLE repertoire in use

In this section, I will present three excerpts which seem a good introduction to the dataset. The first comes from Kai, who is a member of the Studio CofP, highly focused on music production, who could be situated at the high orientation end of Drummond's "'urban/street-style' orientation scale" (Drummond, 2018b, pp.195–197). Kai's personal orientation and the fact that this excerpt comes from a self-recording make it a good example of high usage of MLE features. Kai shows robust use of a number of MLE phonetic features (and those from other levels of the grammar), including those identified as more salient by Cardoso et al. (2019). The second excerpt comes from Chantelle, another of the adolescents, who could be situated at the lower end of Drummond's orientation scale. Her excerpt also comes from an interview: because of this and because of her personal orientation to "urban/street-style", it is to be expected that she shows lower use of MLE features compared to Kai. She uses MLE features that are identified by Cardoso et al. (2019) as more generic, but avoids those features that seem more marked, e.g. DH-stopping. Finally, we will look at a speech extract from a Y1 boy, M4. He uses some MLE features variably, does not use others, and also uses features that are not found in the adolescent data, such as clear /l/ in syllable coda and trilled /r/.

Excerpt 24 comes from a self-recording made by Kai, while he was talking to friends in the studio. The excerpt shows him using features that are more generally vernacular as well as features that are more specific to MLE. In terms of general vernacular features, Kai shows l-vocalisation in *until* (l.5), *visuals* (l.13), *involved* (l.8); and intervocalic t-glottaling in *certain* (l.3), *started* (l.8), *getting* (l.8), as well as categorical glottaling of word-final /t/. But he also exhibits several more salient features: backing of /k/ in *cos* (l.1), *cuz* (l.3); monophthongisation of FACE in *way* (l.3), and PRICE in *like* (l.3, 4, 10); backing of GOAT in *only*, *videos*; and DH-stopping in *that* (l.1, 2, 3, 5, 6), *the* (l.12, 13).

(24) "man never even used to look into it like that"

01	Kai	cos before yeah I knew that like fam (0.31)
		[k̘]os before yeah [ɐ] kn[y] [d]a[ʔ] l[a(ɪ)][ʔ] fam (0.31)
02	Kai	obviously I knew that it (()) –
		obviously [ɐ] kn[y] [d]a[ʔ] i[ʔ] (()) –
03	Kai	it had to be made a certain way but man never even used to=
		i[ʔ] had to be m[eɪ]de a cer[ʔ]ain w[e(ɪ)] bu[ʔ] man never even used to
		=look into it like that cuz (0.39)
		look into i[ʔ] l[a(ɪ)][ʔ] [d]a[ʔ] [k̘]uz (0.39)
04	Kai	like (0.69)
		l[aɪ][ʔ] (0.69)
05	Kai	until man actually mentioned it yeah that he –
		unti[w] man actually mentioned i[ʔ] yeah [d]a[ʔ] he –
06	Kai	that he's a cameraman and that (0.98)
		[d]a[ʔ] he's a cameraman and [∅]a[ʔ] (0.98)
07	Kai	and (0.05)

08	Kai	started getting man involved sta[ʔ]ed ge[ʔ]ing man invo[w]ved
09	Kai	that – [ð]a[ʔ]
10	Kai	now I've only been watching the videos, like now [a]ve [oʊ]nly been watching [ð]e vide[oʊ]s, l[a(ɪ)][ʔ]
11	Kai	js watching (0.33)
12	Kai	the um [d]e um
13	Kai	the visuals innit (0.36) [d]e visua[w]s inni[ʔ] (0.36)
14	Kai	you get me [(ə)jɛʔ mi]

Contrast this with the following excerpt below from Chantelle. Impressionistically, there is an audible difference between the two, in that Chantelle's speech sounds more stress-timed, while Kai's speech sounds more syllable-timed – though this cannot be confirmed without a speech rhythm analysis, which would be beyond the scope of this chapter. Where they are not reduced, Chantelle's FACE, PRICE and GOAT vowels are largely diphthongal. Her PRICE onset is back relative to Kai's – [ɐ] rather than [a]. Similarly, her GOAT diphthong tends to take the SSBE form, [əʊ]. She l-vocalises and t-glottals near categorically – t-glottaling occurs even word-medially, e.g. in *started* and *water*. Chantelle does not DH-stop at all in this excerpt: whereas Kai pronounces [d] on function words such as *the* and *there*, Chantelle more often deletes the /ð/. In *further* (l.6), Chantelle fronts (DH) to [v]. In sum, Chantelle uses features that are vernacular, but does not use features such as DH-stopping that are more marked.

(25) Margate beach

1	Chantelle	>>we were going<< swimming >>in (()) << deep >> end and there was this<< one girl that couldn't ↑swim, >>we were g[ə]ing<< swimming >>in (()) << deep >> end and [∅]ere was this<< one g[ɜw] [∅]a[ʔ] couldn'[ʔ] ↑swim, (0.69)
2	Chantelle	>>and she goes<<, can I – can – ↓like, can you take me like >>and she g[ə]s<<, can [ɐɪ] – can – ↓l[ɐɪ][ʔ], can you t[eɪ][ʔ] me l[ɐɪ][ʔ] (0.53)
3	Chantelle	not (.) into the deep but no[ʔ] (.) into [ð]e dee[ʔ] bu[ʔ]

		(0.79)
4	Chantelle	>>and I was like<< yeah
		>>and [ɐ] was l[ɐ][ʔ]<< yeah
		(0.47)
5	Chantelle	so I got there and I started telling her like
		s[əʊ] [ɐ] go[ʔ] [ð]ere and [ɐ] star[ʔ]ed telling her l[ɐ(ɪ)][ʔ]
		(1.18)
6	Chantelle	if I have you >>on my<< back when I go further down =
		= I’m not gonna be able ((to)) touch the ↑floor,
		if [ɐ] [h]ave you >>on m[ə]<< back when [ɐɪ] g[əʊ] fur[v]er down
		[ə]’m no[ʔ] gonna be [eɪ]b[w̩] ((to)) touch [∅]e ↑floor,
		(0.66)
7	Chantelle	>>’n’ she goes<< “no let’s go let’s go”
		>>’n’ she g[ə]s<< “n[əʊ] le[s] g[əʊ] le[s] g[əʊ]”
		(0.56)
8	Chantelle	so, she js pushing me bare and she js (0.64)
9	Chantelle	tagging along (1.08)
10	Chantelle	and then um (0.78)
		and [∅]en um (0.78)
11	Chantelle	she js started panicking
		she js star[ʔ]ed panicking
12	Chantelle	hh (0.72)
13	Chantelle	>>and she goes<< “>>oh my gosh<< I’m gonna die”
		>>and she g[ə]s<< “>>oh m[ə] gosh<< [ə]’m gonna d[ɐɪ]”
		(0.47)
14	Chantelle	so she js kicking me yeah and I –
		s[əʊ] she js kicking me yeah and [ɐɪ] –
15	Chantelle	she js kicking me under water
		she js kicking me und[ə] wa[ʔ][ɐ]
		(0.58)
16	Chantelle	and, (0.14)
17	Chantelle	she j(h)s (0.60)
18	Chantelle	>>she ((ws)) js<< tryna kill me and th(h)en
		>>she ((ws)) js<< tr[ɐɪ]na ki[w] me and [∅](h)en
		(0.45)
19	Chantelle	Sarah’s like “help h(h)(h)elp”
		Sarah’s [∅][ɐ(ɪ)][ʔ] “[h]e[w]p [h](h)(h)e[w]p”
		(0.64)

20 Chantelle and she's scre(h)aming hh hh

M4's speech exhibits the kind of syllable-timed rhythm described by Torgersen and Szakay (2012), showing almost no reduction of unstressed vowels – hence, for example, GOOSE-fronting is clearly audible in *to* (l.1, 3, 12), because M4 does not reduce this vowel. In terms of segmental features, M4 DH-stops variably throughout, most often word-initially in function words (e.g. *the* (l.5, l.9), *that* (l.12)), but also word-medially in *godmother* (l.6); at the same time, the fricative variant [ð] can be heard clearly in all other instances of these words. There are only two tokens of (TH), and the first of these, in *thingy*, is realised as the interdental fricative [θ], but the second token is fronted in l.5. In l.5, M4 is checking the interviewer's understanding of a new referent, "the car thingy", and so it seems likely that DH-stopping, trilled /r/ and TH-fronting all occur here because his attention is momentarily diverted from the story in this way (Sharma & McCarthy, 2018; Matras, 2009).

Throughout, M4's GOOSE is extremely front, and possibly involves some unrounding – M4's pronunciation of *shoes* (l.7) sounds like *she's*. Other vowels are more variable. FACE, PRICE and GOAT all vary both in their onsets and in how monophthongal/diphthongal they are. Mostly, the onset of PRICE is front, [a], and the onset of FACE is close-mid, [e]. The onset of M4's GOAT is quite different from Kai's, as it is variably central–front, from [ɘ] in l.5,l. 12, to [ɵ] in l.9; GOAT-backing is not in evidence at all.

Similarly, there is no evidence of k-backing here: /k/ is realised as velar, and not backed, in *car* (l.5) and *come* (l.17). There is one instance of l-vocalisation, *beautiful* (l.7), but on the word *ball*, a clear /l/ is audible at the end of the word (l.15, 16). Unfortunately, M4's caregiver did not provide information on his language background, so we have no way of knowing whether the clear /l/ in syllable coda is due to L1 influence. Nor can we be sure of the source of M4's trilled /r/ in *car* (l.5) and *broke* (l.9). Finally, /t/ is variable: M4 seems to t-glottal word-finally on function words, e.g. *but* (l.10), *that* (l.12), *it* (l.12), but [t] is audible in *fits* (l.12) and *foot* (l.12), as well as in the word-medial contexts *getting* (l.2), *beautiful* (l.7), *waiting* (l.18).

Thus, although impressionistically, M4's speech sounds closer to MLE than that of some of his peers, at the segmental level, only some MLE features are present, and others either have not been acquired or are avoided by M4 when he is in the classroom/recording situation.

(26) M4's Cinderella story

01 M4 there's four mice (0.26) um came to rescue
[ð]ere's four m[aɪ]ce (0.26) um c[e]me t[y] [r]esc[jy]
(1.08)

02 M4 one um was (0.36) getting the key,

		one um was (0.36) ge[t]ing [ð]e key,
03	M4	the other one (0.34) was um (0.46) ch- – (0.35) the –
		was trying to
		[dð]e o[dð]er one (0.34) was um (0.46) ch- – (0.35) [ð]e –
		was tr[ʌɪ]ing t[y]
		(1.50)
04	M4	unlock um (0.69) the thingy
		unlock um (0.69) [ð]e [θ]ingy
		(0.31)
05	M4	d- – (0.19) d'you know the ↑<u>car</u> thingy, (0.22) um on – on a door,
		d- – (0.19) d'y[y] kn[ɘ] [d]e [k]a[r] [f]ingy (0.22) um on – on a door,
		[...]
06	M4	um her fairy godmother came (0.29) and
		um her fai[ɹ]y godmo[d]er c[ẽ(ɪ)] (0.29) and
		(0.94)
07	M4	and made her (0.19) beautiful dress (0.20) with blue sh- – (0.25)=
		=shiny shoes
		and m[e]de her (0.19) beau[t]ifu[w] d[ɹ]ess (0.20) wi[ð] bl[y] sh- – (0.25)
		sh[a̞ɪ]ny sh[y]s
		(0.68)
08	M4	(( )) made out of glass
		(( )) m[e]de ou[ʔ] of glass
		(0.22)
09	M4	but then (0.29) the shoe (0.50) broke?
		bu[ʔ] [ð]en (0.29) [d]e sh[y] (0.50) b[r][ɵ]ke?
		(0.26)
10	M4	and then (0.21) but cinderella (0.21) cinderella has found another one
		and [ð]en (0.21) bu[ʔ] cinderella (0.21) cinderella [h]as found
		ano[ð]er one
		(0.35)
11	M4	and then they use it
		and [ð]en [ð]ey use i[ʔ]
		(0.47)
12	M4	but she (0.22) she know that it belongs to her
		(0.19)
		because it fits her [foot]
		because i[ʔ] fi[t̚]s her foo[t]
13	Rosie	[mhm]

14	M4	and then they came um (0.16) and then (0.28) she um
		and [ð]en [d][eɪ] c[e]me um (0.16) and [ð]en (0.28) she um
15	M4	the man that was the – ((was)) taking ((the/a)) ball
		[ð]e man [d]a[ʔ] was [ð]e – ((was)) t[e]king ((the/a)) ba[l]
		(0.17)
16	M4	um (.) the guards um (0.92) taking the ball, (0.17) he said
		um (.) [ð]e guards um (0.92) t[eɪ]king [ð]e ba[l], (0.17) he said
		(0.20)
17	M4	come here young lady
		[k]ome here young l[e(ɪ)]dy
		(0.45)
18	M4	you – you're – the prince is waiting for you
		you – you're – [ð]e prince is w[ɛː][t]ing for y[y]

4.3 Lexis

In an analysis of media treatment of MLE, Kerswill (2014, p.433) notes that in general, "the only features which are referred to are slang terms, most of which are believed by the writers to be of Jamaican origin. Where whole utterances are represented, they are in Standard English with a heavy use of slang. Pronunciation seems never to be commented upon". Kerswill (2014) also finds that in the MLE corpus, most people have no label for their own language, and most refer to it as "slang". Kerswill (2014, p.433) gives the following examples of slang words: "*blood* (friend), *cuss* (defame), *ends* (place of residence), *mandem* (Creole plural), *rude*, *safe*, *tief* (steal), *man* (as address term), *man* (as indefinite pronoun)".

Drummond (2018b) presents various lexical items as examples of Multicultural Urban British English – his data are from Manchester, but the claim is that MUBE is a youth language register shared by adolescents across Britain. The slang listed in Drummond (2018b, pp.215–221) includes: *allow it*; *bait*; *bare*; *dead*; *dutty*; *live*; *long*; *macca*; *man, mans, mandem*; *peak*; *rah* (see Section 4.5.2 below); *racist*; *stush*; *sick*; *time*. Of these, *live*, *macca*, *racist* and *stush* did not appear in the current data, or were not used in the way described by Drummond.

In a corpus-driven comparison of the MLE and LI data with the earlier COLT database, G. Andersen (2016) found 390 tokens of *bruv*, 172 tokens of *blud* and 10 tokens of *bro* in LIC/MLE and none in COLT, suggesting that *blud* and *bro* are more recent developments.

Slang, and especially teenage slang, is known to be transitory, with words being replaced as soon as they lose their novelty (Tagliamonte, 2016b). In Example 27, CB describes how people who “are not about it” will punctuate their speech with *fam*. He also lists *blud, gee* and *bro* as address terms that come and go.

(27) CB discussing address terms

01	CB	((well)) the people that are not about it (0.54)
02	CB	they – they’re always like yeah **fam**, no **fam**, cool **fam** (0.25)
03	CB	they((re)) always saying that
04	RO	okay
05	CB	[((they like))]
06	RO	[that’s not –]
07	CB	it’s just little things you can tell
08	RO	how how would you – how would be the like correct way to use it (0.25)
09	CB	**fam fam**’s out of (.) time now (0.43)
10	RO	what’s that mean
11	CB	no one uses **fam** no more, [like **blud**]
12	RO	[oh okay] (0.33)
13	CB	**blud** used to be, the slang word? (0.38)
14	CB	um back in the day, then it went **fam**
15	CB	then it went **gee**
16	CB	and, **bro**

4.4 Morphosyntactic features

4.4.1 Loss of definite and indefinite article allomorphy

One feature associated with MLE is the simplification of definite and indefinite article allomorphy (Cheshire et al., 2011). This involves using *a* [ə + ʔ] instead of *an*, and *the* [ðə + ʔ] instead of [ði], before a word-initial vowel. The discussion here is limited to the indefinite article, because *a* and *an* were searchable in the transcripts. Examples of this feature in the adolescent data are shown in (28).

(28) a. Kai: ei ma- ey man’s **a arsenal** supporter too innit

b. SD: Well anyway me and G’s in **a interview** right now innit

From listening to the data, it seems that *a* is preferred over *an* by the children. There are numerous examples of *a* used in this way, some of which are shown in (29). All of the tokens of *an* that were found in the child data are listed in (30). Six of the children use *an*, and most of these use it only once.

(29) a. M4: cos that one has **a arrow** (.) but this one doesn't have **a arrow**

b. F1: **a owl**

c. M3: I got **a orange** one

d. M6: and the only time I've been in **a airport** was when I was (.) when my mum was –

e. F3: you have to ask me **a easy** question

f. M5: the girl who have **a orange** thingy

(30) a. F10: can you see (.) **an orange** bench?

b. F8: do you have only a dog next to a boy eating **an ice cream**

c. M3: and a hat, **an ice cube**

d. M1: an umbrella (.) **an elephant**

e. M1: 's not (.) **an animal**

f. M1: **an ankylosaurus**

g. F9: **a – an elephant**

h. F3: that seems like **an eye**

4.4.2 *was/were* variation

Throughout the English-speaking world, there are two dominant patterns of variation in past tense BE (Cheshire & Fox, 2009; Cheshire et al., 2011):

1. Levelling to *was* in all contexts, regardless of number, person and polarity
2. Levelling to *was* in contexts of positive polarity and *weren't* in contexts of negative polarity

Pattern (1) is more common across the English-speaking world, but pattern (2) tends to be more common in Britain (Cheshire & Fox, 2009). Khan (2006) and Tagliamonte (1998) found that in Birmingham and York respectively, there was a decrease in levelled *was* and an increase in levelled *weren't*. While the 16–19-year-olds in Havering did not use non-standard *wasn't* at all, both patterns were found among Hackney 16–19-year-olds (Cheshire & Fox, 2009; Cheshire et al., 2011). It is not possible to say which pattern is preferred by the adolescents and children in the current data, as this would require an accountable variationist analysis. Instead, this section is limited to stating which non-standard forms are and are not found in the current data.

Was in standard *were* contexts occurs in the data, as shown in (31). But the data also appears to hold tokens of non-standard *were*. In the *Linguistic Innovators* corpus, Cheshire and Fox (2009) find no tokens of non-standard *were* in standard *was* contexts. In the examples in (32), no [z] is audible but at the same time, the vowel is a schwa, rather than the NURSE vowel. The fast rate of speech and the fact that these tokens are phonetically reduced makes it difficult to be sure that these are utterances of *were*, rather than *was* with the final /z/ deleted.

(31) a. GW: if **you was** in the same room as a nit- like someone who was taking that

b. Moses: and it's just because (0.64) you knew some people who knew some people who knew some people (.) and **you was** out every day tryna find something fun to do (0.33) and fun coulda been anything when **you was** younger and and kind of more immature

c. Chris: and he knew **we was** walking up the road already, cos that's the time school comes out

(32) a. Sami: I [wə] just sitting there looking at them innit

b. Ali: so the normal Ali he [wə] the quiet one

c. Kai: it [wə] just loads, loads, like loads of work innit yeah

In contexts of negative polarity, *wasn't* and *weren't* both seem to occur. Examples of *wasn't* in standard *weren't* contexts are given in (33), while examples of *weren't* in standard *wasn't* contexts are given in (34).

(33) a. Kai: **we wasn't** focusing on visuals like that

b. Lola: some – **the things** I did **wasn't** cos (.) ah I just felt like doing it, I was always like provoked

(34) a. Matisse: nah I **weren't** here either

b. Kai: it **weren't** him

c. Khadir: even when I **weren't** allowed in I would still come in

d. Sarah: no that **weren't** year seven

In the child data, there are very few tokens of past tense BE – for the majority of the time during the recording sessions, the children were solving the Diapix task, which was not designed to elicit utterances in the past tense. Those children who do use past tense BE tend to use the standard forms, as in (35). There are no tokens of non-standard *were*, *weren't* or *wasn't* in the child data, meaning that even if an accountable analysis were carried out, we would not know whether the children show a preference for pattern (1) or pattern (2). The tokens of non-standard *was*, as in (36), are very few. Two instances of self-correction suggest that non-standard *was* may be some children's vernacular form in the sense of being their automatic way of speaking. In Example 37, F10 begins saying "was going", using *was* with the plural subject "me and Yeta" (l.1). She apparently self-corrects, as she continues in l.5 with "we were going". Similarly, in Example 38, M4 says "they was" twice in l.3 and l.4 before substituting the standard form, "they were", in l.5 and l.6.

(35) a. F3: I **was** right next to my family cos they **were** waiting in the airport

b. F3: we **were** born in Brazil

(36) a. M1: I thought – I thought **we was** gonna wear it

b. F1: d'you remember **we was** drawing something?

(37) F10 *was/were* self-correction

01	F10	y'know me and Yeta
02	F10	**was go-** – uh –
03	F10	we played football and then
04	Rosie	[mm]
05	F10	**[we were] going** for a ha- –
06	F10	for so many stickers we (.) found seven

(38) M4 *was/were* self-correction

01	M4	only those two girls went to the ball
02	M4	but cinderella didn't went to the ball because um (0.98)
03	M4	**they was** (0.35)
04	M4	**they was**
05	M4	**they were** so um
06	M4	**they were** so rude

4.4.3 Bare NPS

According to Cheshire et al. (2015, p.11), a feature more frequent in inner London speech than in young people's informal speech elsewhere is bare NPs. This involves the deletion of the preposition *to*, especially after the verbs GO and COME, as in *I'm going school* and the examples in 39 (see also Cheshire et al., 2008b).

(39) a. CB: my family from Iran go ∅ Turkey (.) and I go ∅ Turkey from here

b. Omar: we go ∅ school. we just finished school ((though)) innit

c. Sarah: it's when you're dreaming and in the dream you go ∅ toilet

There are many examples of absence of *to* in the child data. Select examples are given in Table 4.1 below. The table provides examples of utterances from the same child both with and without the preposition *to*, to show that this form is variable in their speech.

Table 4.1: Examples of null preposition *to* in the child data

Participant	With *to*	Without *to*
F1	but my brother went **to** the beach but he's crying (.) he went **to** the beach a long time ago	no I'm not going ∅ year two [...] I'm just going ∅ ark school
F3	when we have to go (.) **to** school	our mums are like "we have to go ∅ work [...] we have to get ∅ work"
M4	he went **to** the wrong um (.) um country	I can go ∅ my class by myself

4.4.4 Conjoined verbs without *and*

Cheshire, Adger, and Fox (2013, p.54) note that in LIC, adolescents in Hackney but not in Havering would use conjoined verbs without *and* (see also Cheshire et al., 2008b). In the adolescent data in the current project, this mostly occurred with the verb *try*, as in (40).

(40) a. Omar: but now if you do it to the younger people, they will they will try ∅ fight you

b. CB: oh, I thought you said "I'll try ∅ sell white"

c. Chris: and then they handcuff me (0.53) then they try ∅ get me up

In the child data, there are some examples of absence of infinitival *to*, as in (41). However, these are not contexts where it would be possible to join the verbs with *and*. Therefore these seem to be instances of developmental rather than sociolinguistic variation; it will be interesting to see whether when this generation of children reach adolescence, they expand the context for conjoining verbs to include verbs that would usually be joined with *to* and not *and*.

(41) a. F6: Can this pretend ∅ be super class?

b. F1: His birthday was (.) April but he wanted ∅ change it to November

4.4.5 The *man* pronoun

Cheshire (2013) identified *man* as a new pronoun in MLE. This can be an indefinite pronoun, or take on first-person singular (as in the examples in 42) or plural, or second person singular or third person singular reference.

(42) a. I don't really mind how . how my girl looks if she looks decent yeah and there's one bit of her face that just looks mashed yeah . I don't care it's her personality **man's** looking at (Alex, LI corpus)

b. before I got arrested **man** paid for my own ticket to go Jamaica you know . but I've never paid to go on no holiday before this time I paid (Dexter, MLE corpus)

Example 43, from the current data, neatly demonstrates *man* being used as both first-person singular and third-person singular pronoun. This is taken from a self-recording made by Kai and he is in the music studio, in conversation with a friend. He explains how he never paid much attention to music videos until a mutual friend, who was a cameraman, piqued Kai's interest in this area. In l.3, *man* functions as first-person singular pronoun: Kai explains that he knew videos "had to be made a certain way" but that he "never even used to look into it like that". Cheshire has suggested that referring to oneself as *man* can function as a negative politeness strategy by distancing the speaker from what s/he is saying (Cheshire, 2013, p.622); similarly, in this instance, *man* perhaps contributes to distancing Kai from his former less enlightened self in this story. In l.5, *man* refers to the friend, the one who got Kai interested in camera work, but in l.8, *man* refers to Kai again and this time functions as an object pronoun.

(43) "man never even used to look into it like that"

01	Kai	cos before yeah I knew that like fam (0.31)
02	Kai	obviously I knew that it ((it)) –
03	Kai	it had to be made a certain way but **man** never even used to look into it like that cuz (0.39)
04	Kai	like (0.69)
05	Kai	until **man** actually mentioned it yeah that he –
06	Kai	that he's a cameraman and that (0.98)
07	Kai	and (0.05)
08	Kai	started getting **man** involved
09	Kai	that –
10	Kai	now I've only been watching the videos, like

There are various other instances of the *man* pronoun in the adolescent data, as in (44) below. (44 a) is worth noting because it has a first person referent where the speaker is female, indicating semantic bleaching. (44 b) and (c) have first person referents, while (d) has a third person referent.

(44) a. Sarah: I don't know if **man** can swim to save someone like that

b. L: anyway, **man**'s Jewish, so shalom brother

c. Sqara:yeah and they get angry and they put me in detention (0.17) they give me behaviour point they give me that (0.45) but (0.09) **man** don't care

d. L: **man** said madagascar (0.13) madagascar's not even a country

There are no examples of *man* used as a pronoun in the child data.

4.4.6 *Why ... for* question frame

Why... for is a question frame first noted in studies of MLE, but also occurring in outer London speech (Brookes, Hall, Cheshire, & Adger, 2017). According to Cheshire et al. (2015, p.15), "*Why...for* occurs mainly in confrontational or argumentative contexts, perhaps because the framing reinforces the pragmatic force of the question". There are only three tokens found in the adolescent data, and none in the child data. Of the tokens listed in (45), the one in (a) occurred in side-chat between participants, and Amanda was not actually wearing a microphone at the time. In (b), the first token occurs in a reported speech context, and the second is reported thought.

(45) a. Amanda: **why's** he crying **for** I don't understand

b. Matisse: "she's like oh, like, what's happened, like, did - like **why**'s he here **for**, what" (0.44) and I said "I don't know like, he's th- he thinks I've stolen something, I haven't stolen anything" (0.46) and then they they've tr- they've searched me or whatever (0.37) and I haven't got anything (0.37) and then (0.14) they like – they look like the idiots innit (0.29) but it's like (0.55) ( ) from when that's happened, I used to th- I used to think I used to wonder why did it happen (0.67) like, why, **why** did it happen **for** [Hudson: mm] dyou know what I'm saying

4.5 Discourse-pragmatic features

4.5.1 Pragmatic markers

Torgersen et al. (2011) define Pragmatic Markers (PMs) as linguistic items that "express the relation or relevance of an utterance to the preceding utterance or to the context." They are an open class and they sit outside of, or are loosely attached to, the syntactic structure. Torgersen et al. (2011) compare the use of 9 pragmatic markers between the Linguistic Innovators Corpus and the Corpus of London Teenage Speech: the invariant tags *right*, *innit*, *ok*, and *yeah*; *you know*; the extended Pragmatic Markers *you know what I mean*, *if you know what I mean*, *do you know what I'm saying*; and finally *you get me*. They found that *you get me*, *(do you) know what I mean* and *(do) you know what I'm saying* showed both higher rates of occurrence per million words, and a higher proportion of speakers using them, in LIC; *ok*, *right*, *yeah* and *you know* showed higher rates of use and a higher proportion of users in COLT; while *innit* and *if you know what I mean* appeared stable between the two corpora. The PM with the highest "spread" (i.e. the proportion of speakers who used it, even if only once) in both corpora was *innit*.

4.5.1.1 *you get me*

Torgersen et al. (2017) analysed the marker *you get me* in the Linguistic Innovators and Multicultural London English corpora. They identified the following functions: an agreement marker that the interlocutor responds to either by giving a backchannel or by responding in a separate turn; where no response is audible, and the speaker continues by giving an explanation; or as a stand-alone response to signal agreement. In the MLE corpus, *you get me* was not used by 4- or 8-year-olds at all, and only 3 out of 27 12-year-olds used it. In terms of position, in each corpus, it was found that most instances of *you get me* occurred turn-medially (61% and 75%), and the second most common turn position was turn-final (29% and 18%). In both corpora, the rate of turn-initial and stand-alone uses of *you get me* was around 5%. When the utterance position was analysed, the authors found that in both corpora, the position of *you get me* was utterance-final (i.e. at the right periphery) almost 80% of the time. This means that in the LI and MLE corpora, *you get me* was most frequently used turn-medially but scoping backwards over the preceding utterance. In both corpora, most tokens of *you get me* did not elicit a response (rate of response was 14% and 7%). This led the authors to suggest that "whatever the speaker's intended pragmatic function, speaker- or hearer-oriented, it is usually treated by hearers as the former".

Of the functions identified by Torgersen et al. (2017), *you get me* sometimes elicits a back-channel as a response, as in (46).

(46) a. Raphael: we either running away or whatever [Tony: yeah] now you won't do anything back, [**you get me**]
Rosie: [mhm]

b. Kai: but then people from over here they might not, **you get me**
Rosie: yeah

The use of *you get me* preceding an explanation seems commonplace, and can co-occur with *so* in this context, as in (47).

(47) a. GW: even if you got a tracker on your car and your car gets robbed, you need to have a police report anyway (.) so (.) **you get me**, it's not really (w-) much way you can get around it

In the current data, *you get me* as a response is also found, as in Example 48

(48) *you get me* as a response

1	Tony	[when it's a white] person, 's crazy (0.50)
2	Raphael	[yeah like –]
3	Tony	[got mental] illness or something (0.23)
4	Tony	or he didn't mean to do it (0.87)
5	Raphael	[**get me**.]
6	Tony	[((s js))]

There are also instances where *you get me* is uttered by the speaker, rather than the hearer, and seems to acknowledge the hearer's response, as in Example 49.

(49) *you get me* as a response

1	Raphael	that are actually mental in their head and don't really give a
2	Raphael	toss about anyone really
3	Tony	yeah
4	Raphael	**you get me**

4.5.1.2 *innit*

It was mentioned above that *innit* was the only Pragmatic Marker of the nine studied by Torgersen et al. (2011) that remained stable between LIC and COLT, in terms of occurrence per million words and the proportion of speakers using it. In fact, out of all 9 Pragmatic Markers in the study, *innit* had the highest proportion of speakers using it in both corpora. Yet this does not mean that the functions of *innit* have remained stable, and indeed H. Pichler (2016c) argues that in LIC, *innit* shows a number of functions that were not found in COLT. H. Pichler (2016c) identifies three kinds of context for *innit*: when it occurs at the right clausal periphery and takes leftward scope over the preceding clause, this is described as "canonical position"; *innit* appearing at the start of a turn, with scope over a previous speaker's turn, is described as "follow-up", based on G. Andersen (2001); and all other occurrences are described as "non-canonical". The first two, canonical and follow-up contexts, are found in both COLT and LIC, whereas non-canonical positions are found only in LIC (H. Pichler, 2016c).

In the current data there are numerous examples of canonical *innit*, as in 50 below.

(50) a. Sami: I'm always here **innit**

b. Toni: So you just gotta be ready for it **innit**

c. GW: It was blazing hot **innit**

d. SD: Well anyway me and G's in a interview right now **innit**

e. GW: Communities are supposed to stick together **innit**

Follow-up uses of *innit* "generally signal either surprise at or alignment with previous speakers' propositions" (H. Pichler, 2016c, p.64). This is exemplified in Example 51. The interview took place the day after the Grenfell Tower fire and in this example, Amanda is relating how a friend knew someone who lived in the block, and tried to call him but got no answer. At the conclusion of her story (l.3–4), the implication is that the boy who lived in the block died in the fire. Shantel offers the assessment "THAt's deep" in l.5 and Amanda expresses agreement with what Shantel has said by using *innit* in l.6. In the current data, *innit* used in this way usually has emphasis on the first syllable and a rise in pitch on the second. The second syllable seems to be elongated to accommodate this rise.

(51) Follow-up *innit*

1	Amanda	she don't find him, she called him everything,
2	Amanda	's phone just ringing ringing,
3	Amanda	she hasn't heard from him. (0.74)
4	Amanda	°so° (0.15)
5	Shantel	THAt's deep=
6	Amanda	=**inni:t?**

The occurrences of *innit* described by H. Pichler (2016c) as non-canonical include: *innit* appended to a subordinate clause, rather than to a main clause; after a formulaic construction, e.g. *I know, innit*; at the left periphery (LP) of a main clause; after a lone noun phrase (NP) or prepositional phrase (PP); and after a left-dislocated or subject pronoun, e.g. *one person, innit, was a bit weak* (example from H. Pichler, 2016c, p.66).

Example 52 is a good example of *innit* appended to a subordinate clause. The way *innit* is used in this excerpt is actually also a good example of how H. Pichler (2016c) claims *innit* can function when appended to NPs and PPs in narratives, even though in the current example, *innit* in l.7 is appended to a subordinate clause. According to H. Pichler (2016c), phrasal *innit* can be used when introducing a new referent or subject NP into a narrative, or with a scene-setting lone PP. *Innit* can be used in this way "to help speakers monitor participants' acceptance of unexpected or sudden but discursively important changes in referent or narrative setting" (H. Pichler, 2016c, p.74). In Example 52, the interviewer has justasked Denzel how he got into making music, and he says that he remembers being into music when he was 3 or 4. By way of explanation, he offers a small story (Bamberg & Georgakopoulou, 2008), saying that his father used to play Tupac songs during bathtime. *Innit* is used in l.1, after the first mention of "Tupac songs", which are a key feature of the narrative, as the story is about how Tupac became an early inspiration to Denzel. It seems likely that *yeah* and *innit* are used in l.1 to highlight the introduction of two referents: Denzel's dad, and Tupac. The second occurrence of *innit* is at the end of a scene-setting subordinate clause in l.7. The importance of the bathroom setting is suggested by the way Denzel restarts this utterance three times in l.6–7, and by the repetition "I used to have a bath and that" in l.8; as such, *innit* highlights the bathtime setting as an important detail of the story.

(52) "When I was in the bathroom, innit"

01	Denzel	my dad yeah used to play like Tupac songs **innit** (0.53)
02	Denzel	like
03	Denzel	he was a – huge fan of Tupac
04	Denzel	and (0.26)
05	Denzel	obviously, (0.43)
06	Denzel	when Tupac used to –
07	Denzel	I was – ((>>at a time<<)) when I was in ((a)) bathroom **innit**,
08	Denzel	>>I used to have a<< bath and that,
09	Denzel	every time my dad used to (0.17)
10	Denzel	play Tupac songs (0.43)
11	Denzel	I used to um,
12	Denzel	block out (0.39)
13	Denzel	the lyrics (0.12)
14	Denzel	of Tupac (0.45)
15	Denzel	rapping, (0.18)
16	Denzel	an:d, just, mimic the beat

H. Pichler (2016c) gives several examples of phrasal *innit* being used to establish new referents as they are introduced into a narrative. In the examples she gives, the other participants in the conversation were present for the events of the narrative, so they can be expected to retrieve these referents. H. Pichler (2016c, p.74) writes that in all of the left-dislocated and lone NPs and PPs in the LIC data, "NEG-TAGS mark referents that presumably are believed by the speaker to be: (i) identifiable by (some) co-participants against their common ground;9 and (ii) inferable by (some) co-participants from the surrounding discourse context". However, in my data, *innit* is often used this way when addressing the interviewer, to whom the entire story is new information – as in Example 52. It may be the case that even in LIC, *innit* could mark referrents that were not known to interlocutors, and it just happens to be the case that the only tokens of phrasal *innit* in LIC happen to occur in stories that the interlocutors have prior knowledge of. Alternatively, this may be a more recent development, i.e. it may be the case that at the time when LIC was collected, *innit* marked referrents of which the interlocutors had shared knowledge, but its possible contexts for use have since expanded, such that it can now be used to mark referrents that the speaker expects to be received as new information.

In the current data there are also many examples of *innit* attached to a lone NP, as in 53 below.

(53) a. Omar: Obviously as a parent, **innit**, as a mother (.) gotta respect your mum

b. ZR: Yeah that's it. Just favours **innit**. Connections.

c. Daniel: the chicken **innit**

d. Denzel: I've known someone (.) I can use someone as example, like um ((*clicks tongue*)) Abracadabra **innit**. Ab- Abra from (.) Abs °**innit**°

Example 54 is a good example of *innit* used at the left periphery. According to H. Pichler (2016c, p.68), "In LIC, *innit* and a handful of other NEG-TAG variants are occasionally recruited to the LP to perform functions closely related to their prototypical RP functions, the main difference being that agreement is invited in an anticipatory rather than retrospective manner". In l.3, *innit* is positioned at the left periphery of the IP and scopes rightward over the proposition *it's bare fun*. At this point in the interview, Amanda and Shantel, the other participant in the interview, were explaining the concept of Snapchat streaks to the interviewer. A "streak" is when two friends have sent each other "snaps" (images) on the app Snapchat for 3 or more consecutive days in a row; these girls had each achieved streaks of hundreds of days. In this example, Amanda is trying to convey the appeal of Snapchat streaks to the interviewer and in l.3, she turns to Shantel, apparently seeking agreement with the assessment "it's bare fun".

(54) "Innit it's bare fun"

1 Amanda it's bare like –

2 Amanda I think it's bare fun
3 Amanda [*turning to Shantel*] **innit** it's bare fun
4 Amanda I find streaks (.) so fun like

(H. Pichler, 2016c) writes that "LP neg-tags uniformly seek interlocutor attention and corroboration of following propositions" and indeed some of the other instances of LP *innit* in the current data co-occur with other ways of seeking attention, such as the examples in 55. In 55 (a), *innit* co-occurs with the directive "listen", while in 55 (b), Chantelle addresses her interlocutor, Shantel, by name.

(55) a. Amanda: but listen, **innit** this is a mess?
b. Chantelle: Shantel **innit** this song's a banger

The current data also offer contexts for *innit* that are not described in H. Pichler (2016c). There are a number of instances of *innit* attached to a directive, either at the left periphery, as in 56 (a), or at the right periphery, as in 56 (b). Example 57 provides one such instance in context. In l.1 the interviewer (Rosie) tries to get the interviewee (MW) to expand on something he has just said. This is apparently at odds with MW's expectations for the interview format, which involves the interviewer asking questions and him giving answers. The use of *innit* appended to "ask me questions" in l.4 could potentially mark the activity of asking questions as something that should be obvious to both interviewer and interviewee. In Example 58, *innit* appends a request (to get Daniel some barbequed chicken) and could potentially have a positive politeness function; but also it could signal a referent that should be readily accessible to the interlocutor, as Moses is at that moment leaving the interview to get some chicken for himself.

(56) a. Amanda: **innit** move
b. Shantel: remember you have that on **innit**

(57) "Ask me questions innit"
1 Rosie can you tell me more about that
2 MW um:
3 MW >>what dyou mean<<
4 MW like ask me questions **innit**

(58) "Buss me one innit"
01 Moses I'm gonna go and check for this chicken I'll be back in
[two minutes]

02	Daniel	[aight but] –
03	Rosie	[okay yeah] if it's – if it's rea[dy then]=
04	Moses	[yeah]
05	Rosie	=((you c-)) cos [((xxx))]
06	Daniel	[eh yeah] b- buss me one **innit**?
07	Daniel	>>I want one as well<<

Finally, the current data offer examples of *innit* appended to a question, as in (59).

(59) a. Moses: Why are you so in man's business **innit**

b. Raphael: Bruv where does the mic come from **innit**

c. Kai: And what's the writing **innit** what's it say

4.5.1.3 Pragmatic markers in the child data

In the child data, there are no tokens of *innit* nor of *you get me*. However, the children do sometimes use *yeah* as an invariant tag, as in (60). *Yeah* is an older pragmatic marker found in London teenage speech (e.g. G. Andersen, 2001, pp.42–43). Example 61 shows it functioning in the way described by Andersen, "as a device for checking that the preceding [...] noun phrase refers to a mutually manifest concept".

(60) a. F10: s- – uh so .hh so we had these sheets and **yeah** .hh we had to – uh t- – you – .hh tell each other what you had on the shee- –.hh what (.) we didn't have

(61) "A billygoat, yeah?"

1	F4	um (1.00)
2	F4	d'you have (0.58)
3	F4	a (0.14)
4	F4	billygoat **yeah¿** (0.50)
5	F3	mhm (0.25)
6	F4	um that has (0.12)
7	F4	a black hat on

4.5.2 Interjections

G. Andersen (2016, p.34) reports that the interjections *duh* and *rah* are found in the Linguistic Innovators/Multicultural London English corpus, but not at all in COLT. Only 4 tokens of *duh* appear in the LI/MLE corpus, and 11 tokens of *rah* (G. Andersen, 2016, p.35). G. Andersen (2016) quotes the following Urban Dictionary definitions of *rah*: "a sound made when something is cool or exciting"; "an expression of frustration, anger, joy or excitement". At the time of writing, the UD top definition is "slang word for when somethings bad/unbelievebale/exciteable...depending on the sentence you use it in. Pronounce it raaaaahhhh [sic]" (Urban Dictionary, 2020, *rah*). G. Andersen (2016, p.35) writes of *rah*, "The exact attitudinal function of this interjection may be hard to pin down. The examples suggest that it can mark both positive and negative evaluative attitudes [...] The definitions offered for *rah* in the UD support my own interpretations of its variable use". In my data, *rah* only rarely appears as "an expression of frustration, anger, joy or excitement", as in (62 d). Mostly it occurs at the onset of reported speech or a reported reaction, as in (62 a–c, e).

(62) a. Denzel: me I can personally say **rah** fuck the police innit

b. Denzel: you can tell by like (0.68) the scenario that you're (0.38) that we're in innit, [Kai: mm] the way that I use the slang (0.71) from the way that I use it you'll cut, **rah** yeah he said that

c. Tony: but I didn't release the E P innit cos (0.73) it was – it was too personal for (.) the world to know (0.73) but I let (.) like my friends hear it (0.68) and they was like **rah** like (0.60) what you're saying like, I'm feeling innit, I'm touching – it's touching innit [Rosie: mm] and that's – that's what I want my music to do

d. Kai: **rah** this – there's so many wires though innit? ((*laughs*))

e. Sarah: and then you wake up and you're like **rah** (.) I never went to the toilet I just pissed in my bed

4.5.3 Quotatives

Cheshire et al. (2011) report a new quotative in London: *this is* + *speaker*, e.g. "*this is me* 'I'm from Hackney"', "*this is her* 'that was my sister"'. However, no tokens of *this is* + *speaker* were found in the current data. This is perhaps unsurprising, as Ilbury (2020) also did not find any tokens in data collected at a Hackney youth centre in 2017/2018. Similarly, Drummond (2018b, p.211) reports that *this is* + *speaker* does not occur in his Manchester data, nor did he overhear it being during the fieldwork for that project.

4.5.4 *still*

One innovation associated with MLE is the use of *still* as a discourse marker (Adams & Cheshire, 2013; Kerswill et al., 2013). Conventionally in English, *still* can function as a time adverb as in (63), as a concessive adverb as in (64), or as a connective (Adams & Cheshire, 2013). In MLE, *still* may also be used as a clause-final discourse marker, either emphatic or concessive. The examples in (65) show *still* being used as a discourse marker in the current data.

(63) a. Chantelle: he's **still** got the white ball

b. Tony: I think they **still** say "I swear on my mum's life"

c. Shantel: there's **still** like problems on the streets and that

d. Chris: this is after school bear in mind, so I'm **still** in my school uniform

(64) a. Khadir: even when I weren't allowed in I would **still** come in

b. Sami: ((*giving definition of* joint enterpise)) when you're there for something but you never actually took part in it but you **still** get arrested or (0.12) excluded cos, you were there innit

c. Raphael: yeah the enemies will **still** see you as a enemy

(65) a. Tony: saw it **still**, on the social media

b. Rosie: how would you describe how you speak?
GW: normally **still**

c. GW: yeah "where you from **still**" ((it's like)) "yeah man's from down the road **still**"

d. GW: nah I'm not tryna embarrass myself **still**

e. L: I can't tell you that **still**

f. Ahmed: lemme chat to you **still**

g. Sarah: I went and I stuck to that blue church **still**

h. Ibrahim: looks hard **still**

i. Ibrahim: yeah I'm gonna get it tomorrow **still**

4.5.5 *bare*

Bare is described by (Drummond, 2018b, p.217) as meaning "very, a lot". In the current data, this can function as: an adverb modifying a verb, as in (66); an adverb modifying an adjective, as in (67); or as a quantifier, as in (68).

(66) a. Omar: ((they talk)) just talk **bare** to everyone, to anyone, they don't care

b. Chantelle: and she goes no let's go let's go, so she just pushing me **bare** and she just, tagging along

c. Chantelle: yeah cos they they just looking around **bare**, and they're just really paranoid

d. Chantelle: basically well I know basically **bare** from per- personal experience like

(67) a. Sarah: my mouth is **bare** dry

b. Sarah: with our generation I think it's . **bare** close to home

c. Chantelle: no that was actually **bare** fun

(68) a. CB: **bare** man got locked up

b. Sarah: like **bare** of them wanted to get famous so they (were) all chatting shit and wrote a book

4.6 Summary

In summary, almost all of the features of MLE described by Cheshire et al. (2011, 2015); Kerswill et al. (2008, 2013) are found in the current data, at least in the adolescents' speech if not also in the children's speech. At least some of the adolescents in the current data use: the indefinite article *a* before a word-initial vowel; non-standard *was*, *wasn't*, *weren't* and, apparently, *were*; bare NPs e.g. *I'm going toilet*; conjoined verbs without *and*; *man* as a pronoun, usually 1st person but also for 2nd or 3rd person; the *why... for* question frame; the pragmatic markers *you get me* and *innit*; the interjection *rah*; discourse marker *still*; and the modifier *bare*.

Further, the children and adolescents alike use: *a* before a word-initial vowel; bare NPs; and non-standard *was*. Identifying whether in the child speech these are cases of developmental variation or sociolinguistic variation would be more difficult, and is unfortunately beyond the scope of this chapter, though it would certainly be a productive goal for future research (Smith & Durham, 2019, pp.59–64).

There were also notable differences between the data described here and the findings of the MLE and Linguistic Innovators projects. There was no evidence of the quotative *this is* + *speaker* in the Ealing data; in one respect, this is unsurprising, given that Drummond (2018b) and Ilbury (2020) also did not find this feature in fieldsites in Manchester and Hackney respectively. Whether the adolescents in the current data prefer pattern (1) or pattern (2) of *was/were* variation cannot be ascertained without an accountable variationist analysis, but the tokens obtained from a cursory search of the data suggest that the adolescents here use non-standard *were*, which was not found in the analysis by Cheshire and Fox (2009). Finally, for some features, such as *innit* and *you get me*, this chapter has suggested that they serve additional functions that have not been described in prior research.

Section 4.2 took a close-up look at speech extracts from three speakers, who each use a different selection of MLE phonetic features. The following chapters will analyse sociophonetic variation in the diphthongs quantitatively.

Chapter 5

The diphthong variables and acoustic analysis

5.1 Introduction

Chapter 2 described the methodology for data collection, including the choice to include diphthongs as the key sociolinguistic variables in this study. The previous chapter, Chapter 4, gave a brief insight into phonetic variation in the speech of three participants in the current study. The current chapter will explain the methodology used to analyse the diphthong variables acoustically.

5.2 Background

5.2.1 What is a diphthong?

A diphthong is a vowel that involves "a change in vowel quality during the course of the syllable" (Ladefoged, 2005, p.29). In English, depending on your accent, the vowels in words such as PRICE, MOUTH and CHOICE are likely to be diphthongs: Giegerich (1992) describes these as the "true diphthongs" of the English reference accents (i.e. RP, General American and Standard Scottish English). If you speak Southern Standard British English or General American English, the vowels in the words FACE and GOAT are also likely to be diphthongal. The opposite of a diphthong is a monophthong: this is a vowel that shows no change in quality over the course of the syllable.

In addition to PRICE, MOUTH, CHOICE, FACE and GOAT, Wells (1982b) lists the vowels in the words NEAR, SQUARE and CURE as diphthongs in Received Pronunciation (RP).

The diphthong variables in the present study are those found in the words FACE, PRICE and GOAT. More will be said about these in Section 5.2.3.

5.2.2 Lexical sets

In referring to the vowels, I will follow common practice in using lexical sets. The purpose of using lexical sets or word classes is that"[t]hey do not represent the most appropriate phonetic or phonemic notation for any one dialect; instead, they represent a framework that allows us to compare dialects" (Labov, 1994, p.164). So, for example, the word *face* may be pronounced in RP as [feɪs] or in Cockney as [faɪs]. Referring to this vowel with a lexical set, i.e. FACE, makes it clear that we are considering the vowel that is in all the words that rhyme with *face*, and is therefore more accurate than using an IPA symbol such as /eɪ/, which may not be representative of the accent(s) we are studying. I will follow the lexical sets defined in Wells (1982b).

5.2.3 The diphthong variables: FACE, PRICE and GOAT

This and the following sections will describe the English diphthongs that are the focus of this project – this review is intended to make sense of the methodology for the analysis of the diphthongs.

5.2.3.1 FACE

Wells (1982b) defines FACE as: "a front narrow closing diphthong or, less commonly, a front half-close monophthong; in either case, it is unrounded." He classifies FACE as part of "Part-system B" of the English vowels, "the traditional long vowels and diphthongs which have a front mid to close quality or (if diphthongal) endpoint" (Wells, 1982b, p.171), which includes FLEECE, FACE , PRICE and CHOICE. The monophthongal variant of FACE mentioned by Wells (1982b) is found in General American, usually in unstressed syllables, "so that *vacation* may have a monophthong in the first syllable but a diphthong in the second". In RP the monophthongal variant arises through Smoothing, i.e. when a diphthong is reduced to just its nucleus in the environment of a following vowel. This gives ['plɛːɪŋ] for *playing*. Yet the diphthongal variant is a relatively recent innovation. Modern day FACE has three separate origins in Middle English, according to Wells (1982):

1. /aː/, which shifted to [ɛː] during the Great Vowel Shift. This class included *tape, late, cake, safe, case, babe, fade, vague, age, wave, bathe, craze, name, mane, vale, change, waste; taper, bacon, nature, station, lady, raven, invasion, April; bass, gauge, jail, crepe, fete, bouquet*

2. A lower diphthong, /ɛɪ -- æɪ/, with which the words listed in (1) merged during the Long Mid Mergers. Wells (1982b) lists in this class *wait, faith, plaice, aitch,*

raid, nail, main, faint; day, play, way, grey; rein, veil, beige, feint; they, whey, obey; weigh, weight, eight, straight; reign, campaign, deign

3. A third set of words that once upon a time had /E:/. This set includes *great, steak, break, yea*.

The [aː] words shifted to [ɛː] during the Great Vowel Shift. Then, during the Long Mid Mergers of the 16th or 17th Century, these merged with words with a diphthong that was realised as [ɛɪ] or [æɪ]. At this point the words in (1) and (2) were all realised with [ɛː]. Then, during the eighteenth century, this was raised to [eː]. By 1750, in RP, FACE is supposed to have been chiefly realised as [eː], with [eɪ] existing as an allophonic variant, and [ɛː] being found in the environment of a following /r/ (Wells, 1982b, pp.211-212). The process of Long Mid Diphthonging, thought to have taken place in RP around 1800, changed [eː] to [eɪ]. Wells (1982b) describes it as "intrinsically likely" that this change happened first in free monosyllables and spread to other environments later. In many varieties of English, Long Mid Diphthonging did not occur, and so in these varieties, according to Wells (1982b), there is still a monophthongal realisation of FACE . These varieties include northern England, the US North, "the Celtic Countries", and Caribbean Englishes.

Diphthong Shift moved the nucleus of FACE to a more back and more open position. Accents which underwent Diphthong Shift, such as London Cockney, or Australian English, realise FACE as [aɪ] or [ʌɪ]. There are other possible variants: "A centring diphthong, [eə], is found in Tyneside speech, while an opening diphthong, [ɪɛ] etc., is typical of popular West Indian accents" (Wells, 1982b, p.142).

5.2.3.2 PRICE

Wells (1982b) defines PRICE in RP and GenAm as "a wide diphthong with a starting-point which is open, unrounded, and most usually centralised-front, [aɪ], though front and central variants, [aɪ -- ɑɪ], are also common within the standard accents." Wells (1982b) classifies PRICE as part of "Part-system B", which includes FLEECE, FACE , PRICE and CHOICE. According to Wells (1982b, p.171), RP has been moving from a Part-system B that has FLEECE, FACE and PRICE as front vowels, and CHOICE as relatively back, to a system that has FLEECE and FACE as front, and PRICE and CHOICE as less front. This is relevant to the present study as it suggests that in RP, the onset of PRICE has been moving backwards, to become central. PRICE derives from the Middle English close front vowel /iː/. During the Great Vowel Shift, this diphthongised to [eɪ]. The Great Vowel Shift was completed by about 1600. During the eighteenth century, this diphthong had a more open realisation, [ʌɪ], and later shifted to [aɪ]. There is some controversy as to whether the nucleus of PRICE centralised before backing, or vice versa (see Labov (1994)). Wells

(1982b) divides PRICE words into those that were followed by /x/ in Middle English, and those that were not.

There was historically overlap between the vowels in PRICE and CHOICE (Wells, 1982b). The pronunciation of CHOICE words overlapped with that of PRICE words in standard accents until the nineteenth century. There is still overlap between PRICE variants and CHOICE variants across accents. For example, the RP pronunciation of *toy* is very similar to the Cockney pronunciation of *tie*.

In terms of synchronic variation, Wells (1982b, p.149) identifies four main dimensions of variation in the PRICE vowel across accents: how front the starting point is; how open the starting point is; what the quality of the second element is; and what Wells calls the "speed" of the diphthong, here interpreted as how monophthongal/diphthongal the vowel is:

> Very back starting-points, [ɑɪ -- ɒɪ] are characteristic of the urban south of England, the southern hemisphere, and New York speech. Front starting-points, [aɪ] etc. of the north of England, and less open [ɛɪ] (in certain environments) of Tyneside and Northern Ireland. A starting-point that is not fully open, [ɐɪ -- ʌɪ -- əɪ] is typical of the rural south of England, of Barbados, and of parts of the north-eastern United States. In Canada, Virginia, and coastal South Carolina there is marked allophonic differentiation, with a narrower diphthong [ɐɪ -- ʌɪ -- əɪ] before a voiceless consonant and a wide one, [ɑɪ] etc., in all other environments. Some degree of allophonic differentiation of this kind is also found in upstate New York speech (Wells, 1982b, p.149).

Lengthening of the first element is usually associated with a reduction of the glide, so that the glide becomes [ɛ] or [ə]. Or price might just become an "outright monophthong" in such cases, e.g. [aː]. "These variants are particularly characteristic of the American south. But diphthongs with a weakened second element also occur more widely as optional variants; they are found in London speech, in Manchester and Leeds, in South Africa, in Australia, and in Jamaica" (Wells, 1982b, p.150). Scotland, Ulster and Tyneside have not one but two PRICE diphthongs, "perhaps phonologically distinct" (Wells, 1982b, p.150): these two forms are represented as [æ] and [ʌi] (Wells, 1982b, p.172). Finally, according to Wells (1982b), the PRICE vowel is subject to Smoothing in words such as *science, fire.*

5.2.3.3 GOAT

Wells (1982b) defines GOAT as being /əʊ/ in RP and /o/ in GenAm. In GenAm, it can be a "back half-close rounded monophthong or narrow closing diphthong, [o -- oʊ]"; the RP variant is a diphthong with a "mid central unrounded starting-point" similar to RP

/ɜː/, with the second element [ʊ]. This vowel occurs in both checked and free syllables. Wells (1982b, p.173) classifies GOAT in Part-system C, "the traditional long vowels and diphthongs which have a back mid to close quality or (if diphthongal) endpoint", along with GOOSE and MOUTH. Wells (1982b, p.146) observes "This vowel is particularly variable both regionally and socially, and may be found with a variety of monophthongal and diphthongal qualities ranging from [oː] to [eɤ], [ɐʊ] and [ʊɔ]". The origins of GOAT are, according to Wells (1982b):

1. Middle English /ɔː/, which was raised to [oː] in the Great Vowel Shift. These words include *toe, sole, nose*.
2. Middle English /ɔu/ which merged to goat in the process that Wells (1982b) calls the "Long Mid Mergers". These words merged with those in the former category shortly after /ɔː/ had shifted to [oː], giving a merged [oː] for both categories. These words include *tow, soul, knows*.

Next, a diphthongal variant of GOAT developed. By 1750, GOAT probably had a diphthongal allophone (Wells, 1982b, p.212). "The merged [oː] diphthongised in polite usage around 1800, giving an [oʊ] such as is still current in GenAm and may sometimes be heard in old-fashioned RP, as well as in various other accents" (Wells, 1982b, p.193). This process, which Wells (1982b) calls "Long Mid Diphthonging", applied to FACE at the same time. In many regional accents it did not occur, and such accents still have a monophthongal GOAT vowel:

> "in rural and conservative urban working-class accents of the north of England; rather more generally in Wales and Ireland; very generally in Scotland, where diphthongs may even be perceived as a mark of the anglophile; in cultivated West Indian speech, where it is often in sociolinguistic variation with a lower-prestige opening diphthong; and in the northernmost part of the midwest of the United States (Michigan, Visconsin, Minnesota), particularly in the environment of a following voiceless consonant, thus *gate* [ɡeˑt], *soap* [soˑp]; more widely in GenAm in unstressed pretonic syllables, as in the first syllables of *vacation, chaotic, donation*, and *oasis*; and lastly in Indian English and often in African and some other kinds of Third World English." (Wells, 1982b, p.211)

After Long Mid Diphthonging, GOAT had the realisation [oʊ]. At some point subsequently, the onset fronted to give [əʊ] in RP and [ʌʊ] in Diphthong-Shifted accents. This fronting of the GOAT nucleus is referred to by Wells as GOAT Advancement. This

> "has presumably been current since at least the nineteenth century, although [ɜʊ] has only quite recently (since the Second World War?) ousted [oʊ],

> or perhaps rather [oʊ], as the ideal image of a 'correct' or 'beautiful' RP goat diphthong. Some forms of RP have a further advanced variant, [eʊ]. Others retain some rounding, having a rounded mid central vocoid as the first element of a diphthong [ɵʊ]" (Wells, 1982b, p.237)

In RP, then, there is a tendency to front the onset of GOAT, and this is presented by Wells as a prestige innovation.

5.2.4 MLE diphthongs & reversal of Diphthong Shift

To understand the vowel changes that are said to be characteristic of MLE, we need to take a detour through the history of London diphthongs.

MLE is said to show reversal of Diphthong Shift (DS) (Kerswill et al., 2008). Diphthong Shift is a change that is thought to have derived Cockney-like vowels from RP-like vowels – the assumption here is that RP is more conservative and, being a prestige accent, resists change to an extent. According to Wells, Long Mid Diphthonging must have occurred first and Diphthong Shift could have occurred once this was completed (Wells, 1982b, pp.256–257). Long Mid Diphthonging occurred in RP sometime around the beginning of the 19th century, and must have occurred in "popular speech" before this (Wells, 1982b, p.257).

As pointed out by Kerswill et al. (2008), it is uncertain whether RP is indeed what the previous state of popular London English looked like before Diphthong Shift derived the Cockney vowels. They cite evidence from Britain (2005) that the earlier state of the MOUTH diphthong was [əʊ], while the RP-like variant [aʊ] was "virtually unknown" in the southeast of England at that time (Kerswill et al., 2008, p.454).

Wells (1982b) describes Diphthong Shift as a reorganisation of what he calls part-systems B and C in the British English vowel system. Part-system B comprises the vowels FLEECE, FACE, PRICE and CHOICE, while part-system C contains GOOSE, GOAT and MOUTH. In part-system B, under Diphthong Shift, the diphthong nuclei move anti-clockwise in the vowel space, as shown in Figure 5.2: FLEECE lowers, the FACE nucleus moves down and back, PRICE moves back, CHOICE moves back and up. Meanwhile the part-system C vowels move clockwise, with GOOSE being lowered, GOAT lowering, and MOUTH lowering and moving forwards. These two movements are shown in Figure 5.1. Shifted diphthongs are a feature of popular London speech, Birmingham and areas of the midlands of England, as well as Australian and New Zealand English (Wells, 1982b, p.256). Of its origins, Wells states "It is not known when the Diphthong Shift arose. Probably it originated in London; presumably it was well under way by the first half of the nineteenth century, so that early settlers took it to Australia" (Wells, 1982b, p.257).

Trudgill (2004) expands on the latter idea and argues that DS is an example of drift, i.e. an inevitable change inherent in the form of English that was transported by colonial settlers to Australia and NZ.

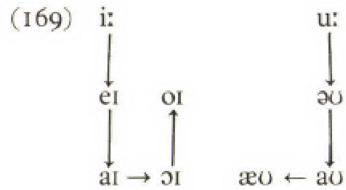


Figure 5.1: Diagram of Diphthong Shift, reproduced from Wells 1982, vol. 1, p.256

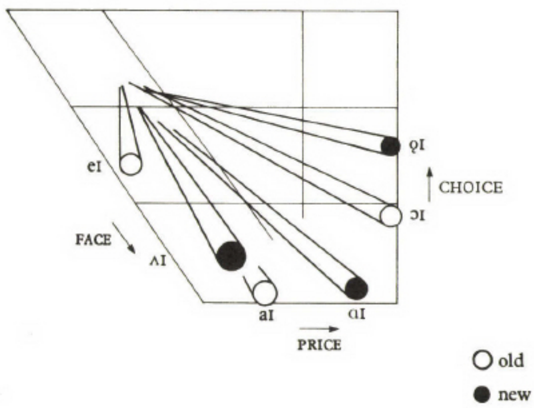


Figure 5.2: Vowel chart showing old and new values of shifted diphthongs, reproduced from Wells 1982, vol. 1, p.257

Labov (1994) sets out both principles and patterns of chain shifting: the principles are rules that apply probabilistically; the patterns are combinations of these rules that have tended to occur together historically in chain shifts in different languages. The principles of vowel shifting are as follows:

- Principle I: In chain shifts, long vowels rise
- Principle II: In chain shifts, short vowels fall and (b) the nuclei of upgliding diphthongs fall
- Principle III: In chain shifts, back vowels move to the front

Labov (1994) then expands his taxonomy of vowel systems to hypothesise a peripheral and non-peripheral vowel system, where lax vowels are non-peripheral and tense vowels are peripheral (Figure 5.3). This is motivated by evidence from chain shifts in progress from Southern US dialects (the Southern Shift) and also Cockney, which involve a movement like the one shown in Figure 5.4. According to Labov, dialects such as these are incompatible with Principle II, which states that the nuclei of upgliding diphthongs should fall – dialects such as Cockney show FACE, MOUTH and PRICE rising rather than falling. Closer acoustic inspection of these shifting systems reveals that the nucleus of MOUTH ends up fronter than the starting position of FACE, and similarly the nucleus of PRICE must end up backer than the other back vowels (Labov, 1994, p.170). This means that it is not enough to have one front position and one back position in the vowel space. Labov (1994, p.170) introduces the concept of *peripherality*, where front rounded and back non-rounded vowels are nonperipheral.

This leads to a modification of the principles to the following:

- Principle I: In chain shifts, tense nuclei rise along a peripheral track (Labov, 1994, p.176)
- Principle II: In chain shifts, lax nuclei fall along a nonperipheral track (Labov, 1994, p.176)
- Principle III: In chain shifts, tense vowels move to the front along peripheral paths, and lax vowels move to the back along nonperipheral paths (Labov, 1994, p.200)

This leads to an understanding of Diphthong Shift whereby the nucleus of MOUTH rises along the front peripheral track, the nucleus of PRICE rises along the back peripheral track, and the nuclei of FACE and GOAT fall along the front and back nonperipheral tracks respectively.

Diphthong Shift combines Labov's Patterns 3 and 4: in Pattern 3, "low vowels move up and back, while the high and mid vowels move to the front"; Pattern 4 involves the lowering of the nuclei of FLEECE, FACE and, in parallel with FACE, GOAT. At the first stage of Pattern 4, the nuclei of FACE and FLEECE become lax so that they are on the "nonperipheral track" and moving downwards. The next stage, according to Labov, depends on the behaviour of PRICE: in the Southern Shift in the Gulf states and Texas, PRICE moves forwards and becomes monophthongal; in southern England (including Cockney), PRICE moves back and up along the peripheral track (along with CHOICE) (see Figure 5.5).

Pattern 4 in the Southern United States involves the fronting of GOAT, GOOSE and FOOT (Labov, 1994, p.215). Labov (1994, p.208) mentions that in London and Norwich, the nucleus of GOAT instead falls to a low, lax position – presumably /ʌ/, as described by Wells (1982a) for Cockney.

The argument of Kerswill et al. (2008) is that East End London speech shows a diachronic shift away from Diphthong Shifted (Cockney) vowels to more RP-like vowels and diphthong onsets. They show this through comparison between older White Londoners and the 16–19 year olds in the Linguistic Innovators project. They note that this is a

(16) PERIPHERAL AND NONPERIPHERAL TRACKS

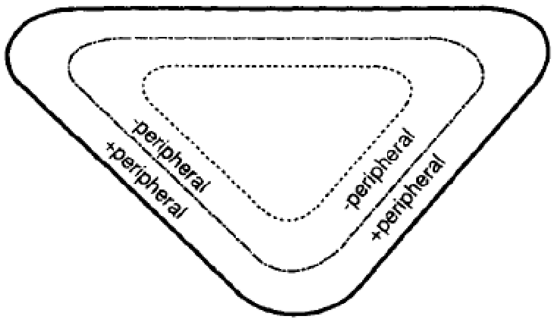


Figure 5.3: Peripheral and non-peripheral tracks of vowel shifting, reproduced from Labov 1994, p.177

(13) EXTENSION OF PATTERN 1 CHAIN SHIFT

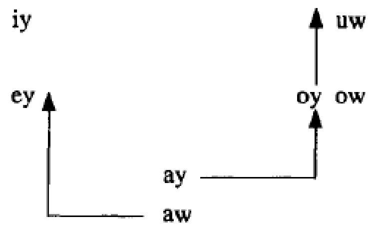


Figure 5.4: Extension of Pattern I chain shift, reproduced from Labov 1994, p.170

(34) PATTERN 4 CHAIN SHIFT

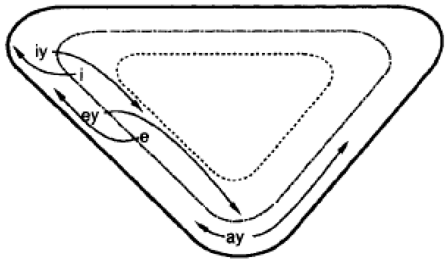


Figure 5.5: The initial movements of a Pattern 4 chain shift, reproduced from Labov (1994, p.209)

reversal of Diphthong Shift only in the sense that the changes go in the opposite direction to that described by Labov (1994) and Wells (1982b), and that they do not believe that before Cockney, popular London speech resembled modern-day RP (Kerswill et al., 2008).

Previous work had already pointed to a move away from Diphthong-Shifted vowels in Reading and Milton Keynes (Kerswill & Williams, 2005). For example, Kerswill and Williams (2000a) showed that even though a slightly fronted onset for MOUTH was common among Milton Keynes caregivers ([æʊ]), their children were converging to an RP-like variant, [aʊ]. This change cannot be levelling to the majority variant, because RP is a minority accent; it seems to be instead "a move away from individual regionally marked forms (both urban and rural) to a socially and regionally more neutral variant" (Kerswill et al., 2008, p.462).

Kerswill et al. (2008, p.482) state that their findings are in some ways similar to the findings of Kerswill and Williams (2005) in Reading and Milton Keynes, in that Hackney young people's diphthongs were more similar to RP than to Cockney. However, young Londoners had additional diphthong variants that are not found elsewhere in the southeast. The tendencies observed in the Hackney young people were as follows: back and lowering of MOUTH; fronting and/or lowering of the onset of PRICE; raising and backing of GOAT; and raising of the onset of FACE (Kerswill et al., 2008, p.483).

5.2.5 Phonetic qualities of MLE diphthongs in comparison to reference varieties

Table 5.1 summarises the qualities given by different authors for the diphthongs FACE, PRICE, GOAT and MOUTH in MLE and (1) Cockney/Popular London and (2) southeastern English, as described for Reading and Milton Keynes.

Figure 5.6 shows the vowel system that is now thought to be typical of young Londoners. It shows two values for GOAT: one of these is the fronted variant found in the southeast of England, and the other is the backed variant that Kerswill et al. (2008) found in Hackney.

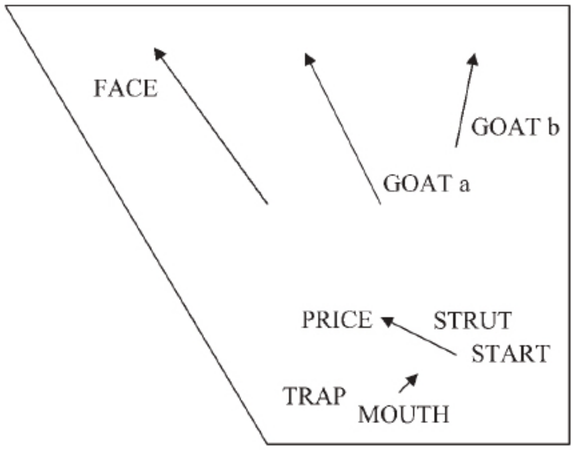


Figure 5.6: Emerging London diphthong system, reproduced from Kerswill et al. (2008, p.484)

Table 5.1: Approximate diphthong qualities for Popular London (from Wells), levelled southeastern (from Williams and Kerswill 1999), and MLE (from Kerswill et al. 2008; Fox 2015; Khan 2006)

Vowel	**Popular London**	**Levelled southeastern**	**MLE**
FACE	[ʌɪ]	[ɛi] ∼ [æi] ∼ [ɐɪ] ∼ [e̞ɪ]	[e̝ɪ ∼ e̞ɪ]
GOAT	[ʌʊ]	[əʏ]	[oː] ∼ [ɔʊ]
PRICE	[ɑɪ]	[ɑ̟ɪ]	[a(ɪ)] ∼ [ɐ(ɪ)] ∼ [æ]
MOUTH	[æʊ]	[aʊ]	[ɑʊ]

5.2.6 Linguistic conditioning

Some sociolinguistic variables show well-established language-internal constraints, whereby certain phonological and/or grammatical contexts favour the occurrence of one variant over another – for example, (ING) variation (e.g. Vaughn & Kendall, 2018; Forrest & Wolfram, 2019; Smith & Durham, 2019)); (td)-deletion (e.g. Labov, 1989; Tamminga, 2018); t-glottaling (e.g. Smith & Durham, 2019; Smith & Holmes-Elliott, 2018; Schleef, 2013). Compared to these well-known variables, relatively little is known about the linguistic conditioning that affects vowel variation in MLE.

S. Fox (2015) carried out a detailed investigation of FACE and PRICE as they were used by adolescents in East London in the early 2000s, in a youth centre in Wapping. S. Fox (2015) identified the newer variants [e̝ɪ], [eɪ] and [ɛɪ] for FACE, as opposed to the more traditional Cockney variants [æɪ] and [aɪ]. S. Fox (2015) reports that a preceding /l/ or /w/ favours the use of the newer variants, and also that occurring word-finally, or being followed by a voiceless obstruent, favours the newer variants. For PRICE, S. Fox (2015) identifies [æ], [aɪ] and [ɐɪ] as newer variants, as opposed to the older variants [ɑ̝], [ɑɪ] and [ɑː]. S. Fox (2015) reports that a preceding voiceless stop, and preceding fricatives or affricates, are the preceding environments most likely to favour the newer variants, while a preceding voiced stop is the context that favours the newer variants the least. As to following environment, following voiced fricatives or stops, and following nasals, favour the newer variants.

Some of these results were corroborated, and others not, by a recent study in an East London secondary school (Gates, 2018, 2019). Gates (2018) analysed the first elements FACE and PRICE acoustically, and modelled these separately for the girls and the boys in a year 10 (age 14–15) age group. For FACE, among the girls, it was found that a preceding nasal, and a following lateral, each significantly lowered the first element; while a preceding voiceless stop significantly raised the onset (Gates, 2018). For FACE among the boys, it was found that a preceding voiceless stop, a preceding voiced stop, and a preceding voiceless fricative, were the contexts that significantly raised the first element; meanwhile, a following nasal significantly lowered the onset (Gates, 2018). For PRICE, among the first, a preceding voiceless stop, and a following approximant, were the contexts that favoured a significantly fronter nucleus (Gates, 2018). For PRICE among the boys, a preceding lateral, and word-final occurrence, were the contexts that favoured a significantly fronter nucleus (Gates, 2018). Gates (2018) points out that for PRICE, the girls' linguistic constraints are the same as those identified by S. Fox (2015), but the boys' are not.

The aforementioned studies did not include GOAT, and as such, we cannot say what constraints may apply to GOAT backing and monophthongisation in East London. One of the contributions of the current study is to examine the effects of preceding and following environment on GOAT variation in the Ealing fieldsite.

5.2.7 Summary

- Of the variables in this study, FACE and PRICE, belong to a different sub-system from GOAT: Wells (1982b) places the former two in part-system B, while Labov (1994) refers to these as the front upgliding diphthongs; GOAT belongs to part-system C, the back upgliding diphthongs.

- Fronting and monophthongisation of PRICE, as occurs in MLE, is also an optional part of the Southern Shift, according to Labov (1994)

- The changes associated with MLE are summarised in Table 5.2

Table 5.2: Summary of MLE tendencies in terms of onset and monophthongisation in the diphthongs FACE, PRICE and GOAT

Diphthong	MLE tendency
FACE	• Closer realisation i.e. lower F1 • Monophthongisation
PRICE	• More front realisation i.e. higher F2 • Monophthongisation
GOAT	• More back realisation i.e. lower F2 (Raising of GOAT is also associated with MLE, but this analysis will focus on variation in the front/back dimension) • Monophthongisation

5.3 Acoustic analysis

5.3.1 Processing the recordings

The recordings were first transcribed at the sentence level in ELAN ("ELAN (Version 5.9)", 2020), with each speaker's utterances transcribed on a separate tier (See Figure 5.7). After orthographic transcription of each speaker's utterances was complete, the transcripts and sound files were force-aligned using FAVE (Rosenfelder et al., 2014). Tokens of the target diphthongs were then segmented in Praat (Figure 5.8) (Boersma & Weenink, 2019). It was decided to segment the vowel tokens manually, rather than using the vowel extraction capability of FAVE, to ensure accuracy, as the sound files had varying amounts of background noise and speaker overlap. Once the vowel tokens had been manually segmented, a Praat script was run to automatically measure the first and second formants of each vowel at five measurement points (see below Section 5.3.4). After this stage, the script output was visually inspected and hand-corrected as necessary. A second Praat script was run to extract the textgrid contents to a csv file for analysis in R.

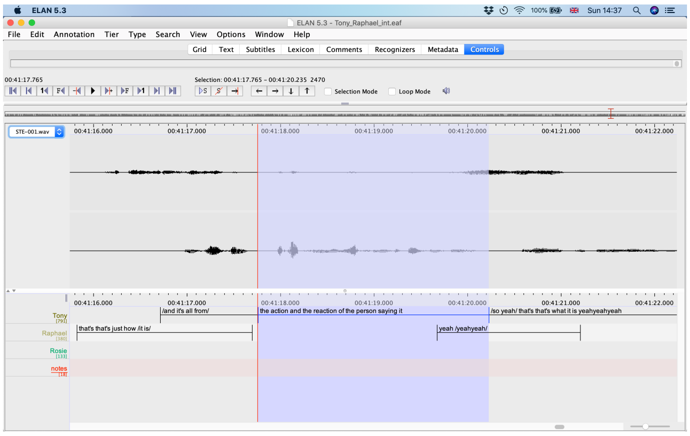


Figure 5.7: Orthographically transcribing the interview in ELAN. Here the speaker Tony's utterance *the action and the reaction of the person saying it* is highlighted. Different speakers' utterances are transcribed on different tiers.

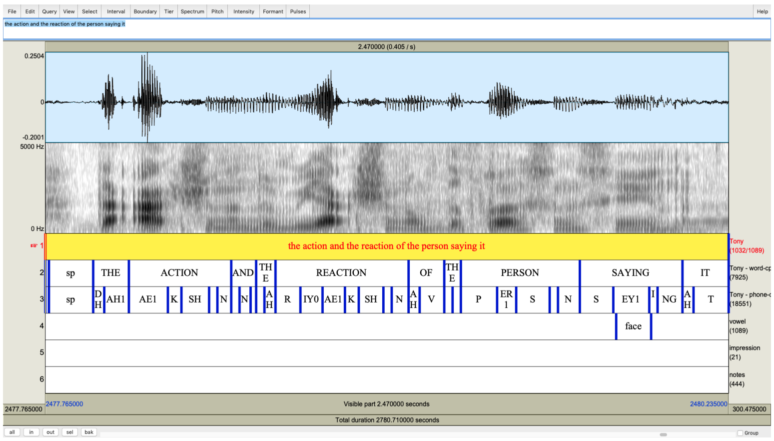


Figure 5.8: The Praat Textgrid used for segmenting tokens of the target vowels. There are tiers for: utterance; word; phone; impression of how the vowel sounds (e.g. "front" or "back", "diphthong" or "monophthong"); and notes.

All of the data was coded in ELAN for both preceding and following phonological and phonetic context (e.g. following underlying phoneme might be /t/, but the phonetic realisation might be [ʔ]).

5.3.2 Segmentation criteria

The onset of the vowel was judged to be (from Thomas (2011) unless otherwise specified):

- **Following an aspirated stop:** the onset of glottal pulsing as visible on the spectrogram and as audible
- **Following an unaspirated stop, e.g. /b/:** the onset of glottal pulsing visible immediately after the burst

- **Following pause:** at the onset of a vowel after silence, there is often a glottal stop or a brief period of creaky voice. This glottalisation would be excluded from the vowel token, and the onset of the vowel was demarcated at the onset of regular glottal pulsing.
- **After a voiceless fricative:** where the onset of voicing becomes visible, **or** where F2 becomes visible
- **After a voiced fricative:** the point at which the F2 becomes visible
- **After a preceding [w]:** the point at which it starts to sound like a [b]
- **After a preceding /r/:** the point at which it starts to sound like a [b]
- **After a preceding [j]:** the point at which it starts to sound like a [g]
- **After a preceding /l/:** an increase in amplitude evident on the waveform (Drager, 2011)
- **After a preceding nasal:** an increase in intensity, evident on the spectrogram as a sudden darkening of the formants
- **After another vowel:** judged auditorily

The offset of the vowel was judged to be:

- **Before a voiced or voiceless obstruent or before a pause:** either where F2 is no longer visible, or where periodicity ceases as visible on the waveform (Thomas, 2011)
- **Before a fricative:** the last periodic wave before the beginning of frication noise as visible on the waveform and spectrogram
- **Before a nasal:** where there is a dampening of the amplitude, as visible from the darkness of the spectrogram and the magnitude of the waves on the waveform
- **Before a lateral or approximant:** judged auditorily, and from changes visible on the spectrogram
- **Before another vowel:** judged auditorily

The boundary between two vowels was judged to be:

- The point of lowest amplitude between the two vowels (see Figure 5.9)

At both the onset and offset of the vowel, the boundary was placed on the nearest zero-crossing where the soundwave was in upward motion.

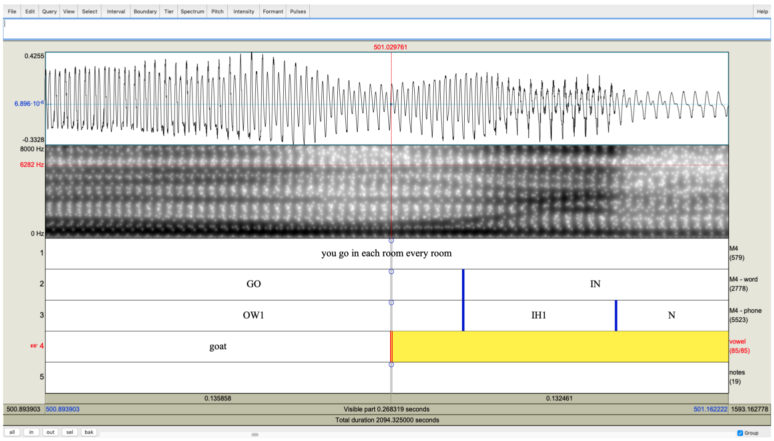
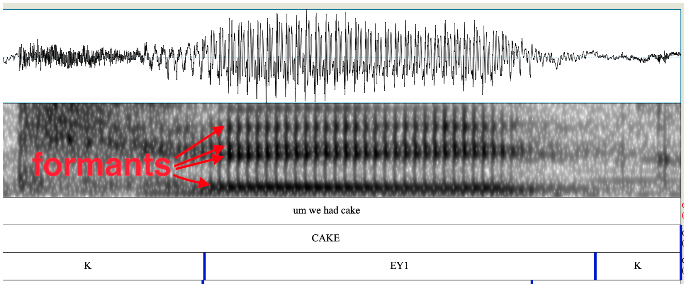


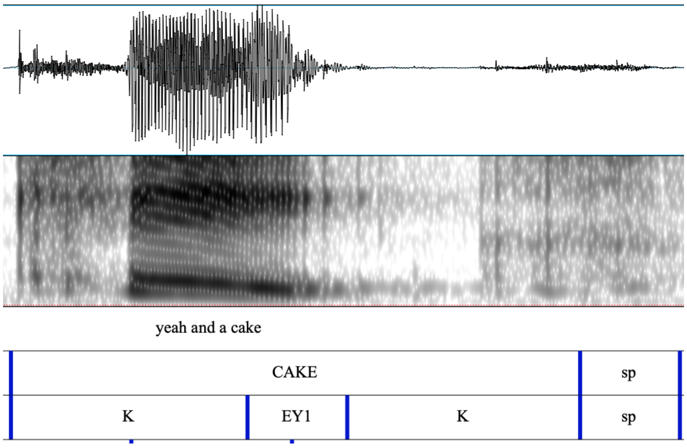
Figure 5.9: The boundary between two vowels is placed at the point between them with lowest amplitude, as visible from the waveform, and where the spectrogram is slightly faded.

5.3.3 Acoustic analysis of children's vowels

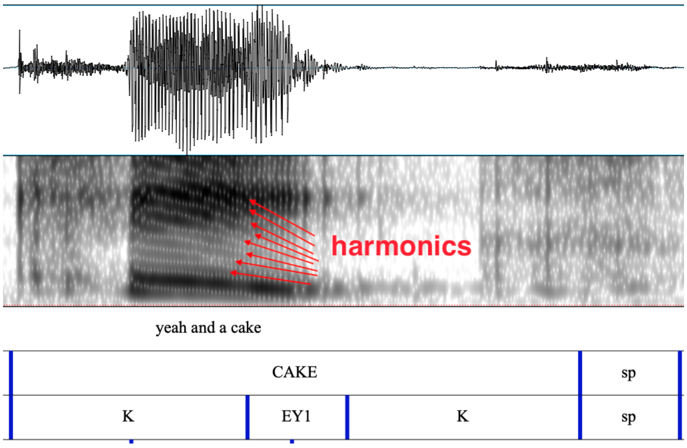
The speech production of children poses certain problems for acoustic analysis, particularly of vowels. As already mentioned above, in analyzing vowels, we examine the spectrogram for the formants that serve as indicators of vowel quality. Formants can be thought of as clusterings of harmonics, which in turn are sound waves at multiples of that speaker's fundamental frequency (Thomas, 2011, p.21). Children have much smaller vocal tracts than adults, and so their fundamental frequency tends to be much higher, and this in turn means that the harmonics will be more widely spaced apart than they would be in, for example, the voice of an adult male (Thomas, 2011, p.160). This spacing out of the harmonics can make it extremely difficult to see formants in the spectrogram of a child's voice as compared to an adult's. Compare Figure 5.10a, where the formants are clearly visible, with Figures 5.10b and 5.10c, where the harmonics are widely spaced and easy to see, but the formants are not so easy to identify. To solve this problem, the parameters of the Praat script used to measure the vowels were altered according to the ages and sexes of the speakers. By default, Praat looks for 5 formants, and expects a maximum formant of 5500Hz. 5500Hz is an appropriate maximum formant value for adult females and 5000Hz is appropriate for adult males (Boersma & Weenink, 2019). Praat suggests that a maximum formant of 8000Hz may be more appropriate for young children (Boersma & Weenink, 2019). The parameters that were used in the present study are outlined in Table 5.3.



(a) The word *cake* spoken by an adolescent male, with the spectrogram configured using Praat's default settings. The formants in the vowel are easily identifiable.



(b) A child's utterance of the word *cake*, again with the spectrogram configured using Praat's default settings.



(c) The harmonics are widely spaced and clearly visible, while the formants cannot clearly be seen.

Figure 5.10: Spectrograms showing the ready visibility of formants in an adolescent male's data using Praat's default settings, while the same settings show harmonics, rather than formants, in a child's speech

Table 5.3: Formant extraction parameters used in the Praat script to extract F1 and F2 frequencies

	Adolescent males	Adolescent females	Children
Timestep	0.0025	0.0025	0.0025
No. of formants	5	4	5.5
Max. formant frequency (Hz)	5000	5500	8000
Window length	0.025	0.015	0.01
Pre-emphasis from	50	50	50

5.3.4 Measurement points & Trajectory Length measure

As a diphthong is a vowel that involves a change in quality, the acoustic measurements of the formants need to capture this change in quality. Even monophthongs are not absolutely static (Di Paolo, Yaeger-Dror, & Beckford Wassink, 2011; R. Fox & Jacewicz, 2009). With diphthongs, then, it is doubly important to measure the formants at multiple points throughout the vowel segment.

Di Paolo et al. (2011) identify four types of approaches to deciding on measurement points. These are: the default distance approach; maximal displacement approach; proportional distance approach; and the interval approach. In the default distance approach, measurements are taken at a specified time distance into the vowel, e.g. 25ms. In the maximal displacement approach, the analyst visually identifies the peak or trough in F2 and takes formant measurements at this point. In the proportional distance approach, measurements are taken at percentages of the vowel duration, e.g. 20%, 50% and 80% points (this was the approach followed in the present study). Finally, in the interval approach, measurements are taken at regular intervals of time, e.g. every 10ms, throughout the vowel segment.

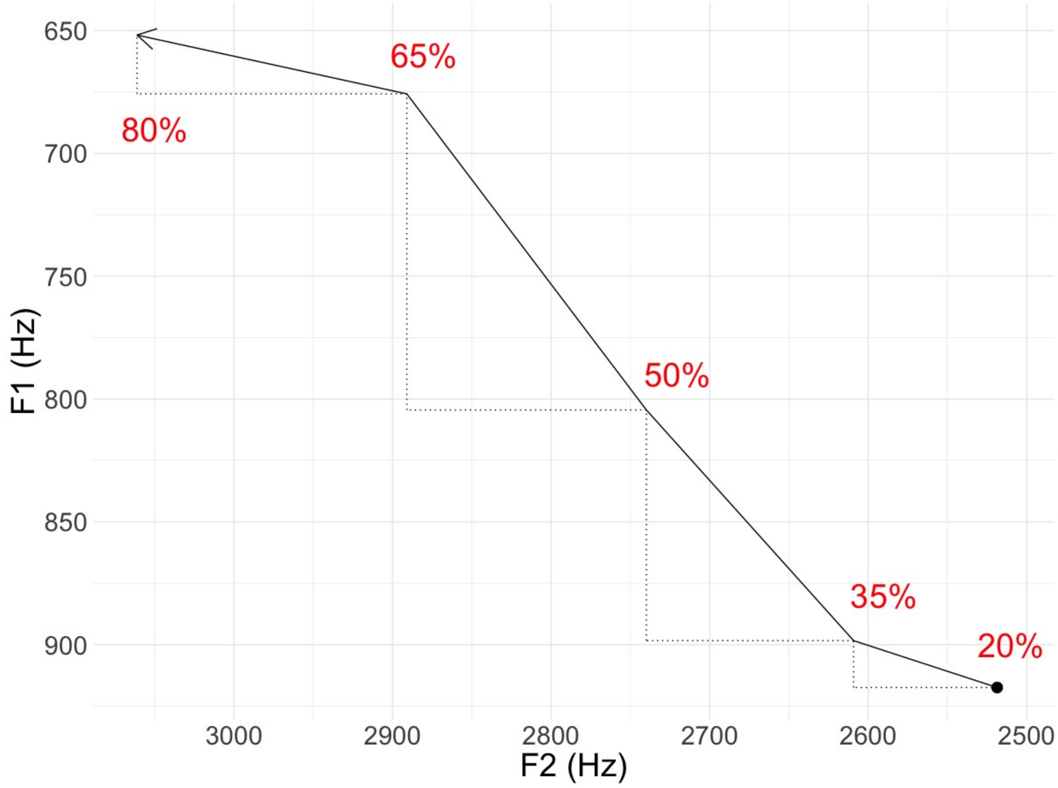


Figure 5.11: A diagram showing the five measurement points used to calculate Trajectory Length

R. Fox and Jacewicz (2009) compare three measures of dynamic variation in F1 and F2: vector length; trajectory length; and spectral rate of change. Vector length is the Euclidean distance in F1xF2 space between 20% and 80% duration points in the vowel segment. The calculation of trajectory length involves calculating the Euclidean distance in F1xF2 space within four sections of the vowel – 20–35%, 35–50%, 50–65% and 65–80% – and then taking the summing of these four vectors. The spectral rate of change is simply the trajectory length divided by 60% of the vowel duration. R. Fox and Jacewicz (2009) argue that vector length can underestimate formant movement, particularly if the vowel involves U-shaped movement in formant space (the example they give is of /ae/ pronounced in the "Southern drawl" of North Carolina). Trajectory length may be a more accurate means of comparing dialectal differences in the formant movements of particular vowels. Meanwhile, dialectal differences were also found in spectral rate of change for the vowels examined.

Other recent studies also pursue a proportional distance approach, with the difference that the number of the measurement point is entered as a fixed effect in the regression model. This approach is taken by e.g. Wormald (2016); Cardoso (2015). In her analysis of FACE and GOAT, Wormald (2015) took measurements at 11 points within the vowel segment, at 10% duration intervals from the onset to the offset. D. Williams and Escudero

(2014) take 30 equally spaced measurement points between 20% and 80% of the vowel (Elvin, Williams, & Escudero, 2016). van Son and Pols (1992) take measurements 16 equidistant points in the vowel segment.

R. Fox & Jacewicz's (2009) Trajectory Length measure was adopted for the current study as a measure of monophthongisation. It was desirable to have a measure of diphthong dynamics that took account of change in both F1 and F2 together, and R. Fox and Jacewicz (2009) found that Trajectory Length outperformed other metrics in measuring (ay)-monophthongisation. As such, measurements of the first and second formant were taken at 20%, 35%, 50%, 65% and 80% time points in each vowel token. As Trajectory Length can only take on positive values, the natural logarithm was taken and used as the dependent variable in the statistical models. Trajectory Length is hereafter abbreviated to TL.

The onset qualities of FACE, PRICE and GOAT are operationalised as the F1 frequency of FACE, and the F2 frequency of PRICE and GOAT, at the 20% duration point. Hereafter, "onset F1" or "onset F2" may be used to mean the F1/F2 frequency at the 20% duration point. The choice of selecting F1 or F2 as the dependent variable for each diphthong is based on the MLE tendencies described above in Table 5.2.

5.3.5 Normalisation

It was described above how vowel formant frequencies vary as a function of the length of the speaker's vocal tract and, consequently, the fundamental frequency of their voice. An acoustic analysis of vowel variation needs to take into account the fact that although the frequencies of different speakers' first and second formants for a particular vowel may vary wildly in Hertz, listeners still perceive that these speakers are uttering the same vowel. Normalisation of the raw formant frequencies is one way to reduce interspeaker differences in formant frequencies and make the data of speakers of different ages and sexes comparable.

According to Disner (1980) and Thomas (2002) (quoted in Thomas (2011, p.161)), there are four goals of a normalisation technique:

1. eliminating variation caused by physiological differences among speakers (i.e. differences in vocal tract length);
2. preserving sociolinguistic/dialectal/cross-linguistic differences in vowel quality;
3. preserving phonological distinctions among vowels;
4. modelling the cognitive processes that allow human listeners to normalise vowels uttered by different speakers

In sociolinguistic studies, the first two of these goals are usually the most important (Thomas, 2011, p.161): sociolinguists aim to reduce physiological differences such that males and females or adults and children can be compared with one another, but preserve sociolinguistic differences that exist between these different groups.

The method selected for normalisation in the current project was the modified Watt-Fabricius method (Fabricius, Watt, & Johnson, 2009). The Watt-Fabricius method (or the S-centroid procedure) set out in Watt and Fabricius (2002) is intended to be a speaker-intrinsic method of normalising vowel formant data. It requires: the F1 and F2 of the FLEECE vowel, on the assumption that this vowel represents the top left hand corner of a British English speaker's vowel space; the F1 and F2 of the TRAP vowel, on the assumption that this is the most open vowel in the speaker's system; and a hypothetical close back vowel *u'*, where F1 = the F1 of FLEECE, and F2 also = the F1 of FLEECE. The difference between the Watt-Fabricius method as laid out in Watt and Fabricius (2002) and the modified Watt-Fabricius method as described in Fabricius et al. (2009) is that the latter does not use real measurements for the F2 of TRAP, but rather calculates the F2 of TRAP as being the midpoint between the F2 of FLEECE and the F2 of *u'*. A centroid point for F_n for each speaker is calculated using these three corner vowels:

$$S(F_n) = \frac{[\text{FLEECE}]F_n + [\text{TRAP}]F_n + [u']F_n}{3}$$

The observed measurements in Hertz of F_n are then divided by $S(F_n)$. The $S(F_n)$ is calculated by speaker, so each speaker's formant measurements are divided by that speaker's centroid measure. Each speaker's centroid value is different, depending on how widely spaced the corners of his/her vowel space are. Dividing the formant measurements by this speaker-specific centroid measure reduces the differences in formant frequencies that arise as a consequence of speaker physiology (goal 1), whilst preserving other variation (goal 2).

The modified Watt-Fabricius has been successfully used in many recent studies in sociophonetics (e.g. Wormald, 2015; Haddican, Foulkes, Hughes, & Richards, 2013; Podesva, D'Onofrio, Van Hofwegen, & Kim, 2015; Sóskuthy, Foulkes, Hughes, Hay, & Haddican, 2015; Wong, 2012; Holmes-Elliott, 2015).

5.3.6 Exclusions

Tokens were included where the duration of the vowel segment was no less than 50ms and where the diphthong was not reduced, i.e. did not have a schwa-like quality. Tokens were also only included where there was not substantial speaker overlap or background noise, which would have obscured the formants in the spectrogram. Tokens from at least 15 minutes into the interview were analysed (Podesva et al., 2015). Tokens with a coda lateral or approximant were excluded because of the tendency for these types of segment to produce allophonic changes in preceding vowels, and/or cause "breaking" of the vowel. The number of tokens of FACE, PRICE and GOAT included in the analyses can be found in Appendices A and B.

5.4 Independent factors

This section will justify the inclusion of the dependent and independent variables in the models.

5.4.1 Language-internal factors

Preceding and following linguistic context were included in the models not just because identifying the linguistic constraints on variation is one of the goals of variationist sociolinguistics (Weinreich, Labov, & Herzog, 1968), but because this is one of the ways in which the current study can be most helpful to future researchers working on MLE.

The language-internal factors included in the models were:

- Duration
- Preceding environment
- Following environment

It was important to control for duration of the vowel segment because we expect this to have an effect on monophthongisation (R. Fox & Jacewicz, 2009). The duration of each vowel token was automatically extracted from the Praat TextGrid. This was log-transformed, as is standard in phonetics (e.g. Gahl & Baayen, 2019; Englund, 2018; Zellou & Tamminga, 2014; Desmeules-Trudel & Brunelle, 2018), and z-scored before being entered into the models.

Preceding phonological and following phonological environment were included in the models for two purposes: to control for the influence of coarticulatory effects of the preceding consonant and following consonants on the vowel formants; and in the interest of exploring what language-internal constraints influence sociophonetic variation in the vowels. Preceding environment was categorised according to a combination of place and manner, and the levels of this variable are described in Table 5.4.

The following phonological environment variable was categorised by syllable coda type in the first instance: i.e. whether the vowel occured word-finally in an open syllable; word-medially in an open syllable; or in a closed syllable. Within the closed syllable category, the different types of coda were further categorised according to whether the consonant in the syllable coda was a voiced obstruent, a voiceless obstruent, or a nasal (tokens with a coda approximant or glide had already been excluded). Taking into account whether the syllable is open or closed is known to be an important factor in monophthongisation in other dialects (e.g. Fridland, 2003), hence this was prioritised above the voicing, place and manner of the following segment.

Both preceding and following environment were sum-coded before being entered into the models – for the contrast matrices, see Appendices A.2 and B.2.

Table 5.4: Language-internal independent factors: summary of categories used in the analysis

Variable	Level	Description
Preceding segment	Approximant/glide	The onset consonant is an approximant /ɹ w/, glide /j/ or lateral /l/
	Coronal	The onset consonant is a coronal obstruent, either voiced /d ð z ʒ/, or voiceless /t θ s ʃ/
	Labial	The onset consonant is a bilabial or labiodental obstruent, /p b f v/
	Nasal	The onset consonant is a nasal, /m n ŋ/
	Other	
	Velar	The onset consonant is a velar obstruent, /k ɡ/
	Vowel	There is no syllable onset and the immediately preceding syllable is open, e.g. in contexts like *so overcoming its fear of being shot*, *my A-levels*
	Zero	There is no syllable onset and there is no preceding word/utterance
log(duration)		
Coda type	Coda nasal	Syllable coda is an underlying nasal e.g. *don't, time, game*
	Final open	The vowel is word-final and the syllable has no coda, e.g. *play, so, tie*
	Medial open	The vowel is not word-final but the syllable is open, e.g. in *tiger* [taɪ.ɡə], *socialise* [so.ʃə.laɪz], *station* [steɪ.ʃən]

Voiced	The syllable coda is a voiced obstruent
Voiceless	The syllable coda is a voiced obstruent
Other	

5.5 Statistical analysis

5.5.1 Bayesian modelling

Statistical analysis for the current project was carried out in the Bayesian paradigm, which is becoming more and more popular as an alternative to the standard Frequentist methodology in Psychology, Linguistics and the social sciences. This section will outline the motivations for using Bayesian as opposed to Frequentist statistics.

The primary reason for using Bayesian statistics is that it allows us to quantify our degree of belief in the hypothesis of interest. Frequentist statistics deals only with the probability of the current data given that the null hypothesis is true – $P(data|H_0)$. In the world of Frequentist statistics, we imagine that we can hold the hypothesis – the null hypothesis – constant, and vary the data, repeating the same experiment infinite times. This means that Frequentist hyopthesis-testing "answers a question that we do not actually want the answer to (can we reject the null?), and which relies on the imagined (and usually unrealistic) properties of data that we did not collect" (Nicenboim, Roettger, & Vasishth, 2018, p.46). Thus, in Frequentist statistics, the researcher sets long-term error rates in advance of collecting and analysing the data, such that one could be confident of only getting a certain proportion of false positives in the long-term. This long-term false positive rate is alpha and is conventionally set to 5%. However, in practice, it is rarely convenient to repeat experiments, particularly in social sciences and linguistics. In analysing our data, we would usually much rather know the probability of our hypothesis given the data – $P(hypothesis|data)$ – and this is something that can only be done with Bayesian statistics, as opposed to Frequentist.

This means that while employing Bayesian methods involves the complexity of learning to apply these methods, it brings the practical benefit of giving results that are easier and more intuitive to interpret (Nicenboim & Vasishth, 2016). "A large part of the confusion [with Frequentist statistics] probably comes from people unwittingly having a Bayesian understanding of statistics and unthinkingly believing Neyman-Pearson statistics give them Bayesian answers (Oakes 1986)" (Dienes, 2008, p.80). Nicenboim and Vasishth (2016) explain how many of the errors in interpreting Frequentist p-values come from researchers instinctively wishing to interpret p-values as a Bayesian measure of evidence.

Bayesian models also have the advantage of coping well with small sample numbers. This is a huge advantage in variationist studies, where the data consist of spontaneous speech and therefore the researcher has no control over the number of tokens of different categories. If you do not have statistical power, it is probably better to use Bayesian statistic with weakly informative priors (Vasishth & Nicenboim, 2016, p.359; Gelman & Carlin, 2014).

Finally, computational methods for running Bayesian regression analyses are able to cope with a more complex random effects structure than can be handled by commonly-used Frequentist alternatives (Vasishth, Nicenboim, Beckman, Li, & Kong, 2018). It is known that running models with e.g. random intercepts for participants affects the generalizability of the results (Barr, Levy, Scheepers, & Tily, 2013); in spite of this, it is often not possible to create models that will converge with the full random effects structure justified by the design (Barr et al., 2013) when using packages such as *lme4* (Bates, Mächler, Bolker, & Walker, 2015). One way to avoid this problem is to use a Bayesian method of running mixed-effects models (Franke & Roettger, 2019; Vasishth et al., 2018; Nicenboim & Vasishth, 2016).

5.5.2 Choice of priors

If there is a parameter of interest, the methods of Bayesian statistics involve eliciting prior beliefs about possible values that this parameter could take, and then collecting data. The prior beliefs take the form of a probability distribution. The evidence collected is known as the likelihood – all Frequentist inference is based solely on the likelihood. The researcher's state of knowledge after the experiment is known as the posterior. The goal of all Bayesian inference is the calculating of the posterior distribution, which, like the prior, is a probability distribution. Because the posterior is based both on prior beliefs and current evidence, it is expressed as follows (Nicenboim & Vasishth, 2016):

$$posterior \propto prior * likelihood \tag{5.1}$$

Priors can be described as uninformative, weakly informative or informative. An uninformative prior means a uniform prior: this distribution looks like a flat line, where all values are treated as equally likely. The brms package by default uses an uninformative prior on fixed-effect coefficients (Buerkner, 2017, p.3). It is called an "improper" prior because it is not actually a valid probability distribution (Vasishth et al., 2018; Lambert, 2018). Whereas Objective Bayesians believe in setting "improper"/uniform priors, many practitioners advocate setting weakly informative priors (Franke & Roettger, 2019; Lambert, 2018; Vasishth et al., 2018; Nicenboim & Vasishth, 2016). Weakly informative priors, also called regularizing priors, constrain the parameter to take on only plausible values – e.g., if the response variable is fundamental frequency in Hertz, we would want the parameter to take on only positive values, and we would expect it to be in the order of hundreds, not thousands or hundreds of thousands. "'Regularizing' here means that extreme values are disallowed or downweighted; for example, a prior on a correlation parameter would be regularizing if it disallows or downweights extreme values such as -1 or +1, which are quite unlikely in data. Weakly informative priors give some minimal amount of information and have the objective of yielding stable inferences" (Vasishth et al., 2018, p.150).

The priors set in the current study were as follows:

1. Normal prior on the intercept: mean 0, standard deviation 3 – a weakly informative prior. The response variables were all scaled to be expressed in z-scores, so the mean of the response for each model is 0. And as the response variables were expressed as z-scores, we would expect 99.7% of observations to be within +3 or -3 standard deviations of the mean, hence this choice of prior.

2. Normal prior on all population-level effects: mean 0, standard deviation 3.

3. Random effects correlation matrix: LKJ(2) as recommended by Vasishth et al. (2018, p.150)

4. On all standard deviations, the brms default: student-t (3, 0, 10)

5. On sigma, the brms default: student-t (3, 0, 10)

5.5.3 Reporting of results

In the analyses in Chapters 6 and 7, the following key measures are reported:

- The estimated regression coefficient ($\hat{\beta}$), i.e. the median of the posterior distribution

- 95% Highest Density Interval (HDI) about this estimate

- Probability of direction (PD): the probability that the effect is positive or negative. Close to 50% = low probability, close to 100% = high probability.

If $\geq 95\%$ of the posterior is $>$ or $<$ 0, and the 95% HDI does not include 0, we will conclude that there is strong evidence for a positive or negative effect (Tanner, Sonderegger, & Stuart-Smith, 2019; Franke & Roettger, 2019; Nicenboim & Vasishth, 2016). If $\geq 95\%$ of the posterior is $>$ or $<$ 0, but the 95% HDI includes 0, we will say that there is marginal evidence for a positive or negative effect.

The model summary tables in Appendices A and B report:

- The estimated regression coefficient ($\hat{\beta}$), as above

- Lower and upper bounds of the 95% Highest Density Interval (HDI), as above

- Probability of direction (PD), as above

- The percentage of the posterior distribution that falls within the Region of Practical Equivalence (ROPE). The ROPE is a set of values considered to be equivalent to the null for practical purposes (Kruschke, 2015, p.336). The logic of this is that it is possible to end up with a posterior distribution on a parameter of which 99% is > 0 – and so we would consider there to be strong evidence that the parameter is positive – but in fact a large proportion of the posterior is gathered close to zero. The ROPE gives us a way to determine what proportion of the posterior we consider to be effectively similar to zero. In the model summaries, the ROPE is set to [-0.1, 0.1] i.e. $\pm$ one tenth of a standard deviation (because the dependent variables were converted to z-scores before being entered into the model). The smaller the proportion of the posterior falling within the ROPE, the stronger the evidence for a non-null effect.

- Effective Sample Size (ESS). The modelling procedure uses a sampling algorithm – Markov chain Monte Carlo (MCMC) – to create a sample of the posterior distribution. This is a form of dependent sampling, where each successive sample is dependent on the previous one. A greater number of dependent samples is needed to characterise the posterior than would be needed if independent sampling were possible. ESS is a measure of how many independent samples the MCMC-generated dependent samples are equivalent too. A higher number is better: Vasishth et al. (2018, p.156) recommend that as a rule of thumb that the effective sample size should be $> 10\%$ of the total number of samples. In the current analysis, 4 MCMC sampling chains were run for 2000 iterations each, 1000 of which were warm-up iterations. This means that for each parameter, inference is made from a total of 4000 iterations. Therefore for each parameter, the ESS should ideally be > 400.

- $\hat{R}$: this is a metric of model convergence and stands for the ratio of between to within chain variance (Vasishth et al., 2018, p.156). $\hat{R}$ of approximately 1 for each parameter indicates that the model has converged, and indicates that the samples from one chain are similar to the samples from another chain; if it is not 1, this means that the different chains are producing very different estimates for the parameter (Lambert, 2018, pp.315–316). 1.1 is often used as a cut-off: $\hat{R} > 1.1$ for a parameter indicates that the model has not converged (Lambert, 2018, p.316).

5.5.4 Worked example

This section demonstrates the differences in interpretation in Frequentist vs. Bayesian statistics by comparing two linear regression models, one created using Frequentist methods implemented via the R packages lme4 and lmerTest (Bates et al., 2015; Kuznetsova, Brockhoff, & Christensen, 2017), and the other using Bayesian methods implemented via the brms package (Buerkner, 2017).

For this example, we use one of the simpler models from the analysis. This is partly because it makes the explanation simpler, and partly because some of the models used in Chapters 6 and 7 would not converge when implemented using lme4, as they had too complex a random effects structure. The model we will be working with predicts the normalised F2 at the 20% time point (as a z-score) in tokens of *like* against the independent variables log(duration), age (adolescent or child; this variable is sum-coded so that the intercept represents the average between the two factor levels), sex (also sum-coded) and the interaction of age and sex, with a random intercept for each participant. The ouput of this model is discussed further in Section 7.4.1.3.

Table 5.5 gives a summary of this model when implemented with Bayesian methods using brms, and Table 5.6 summarises the output when implemented with lme4 and lmerTest. For the sake of comparability with Table 5.5, Table 5.6 includes 95% confidence intervals obtained using the lmerTest package's confint() function.

It can be seen that the estimates and confidence intervals/highest density intervals given by each model are extremely similar (reported in the first three columns of Table 5.5 and 5.6). It also happens that the same parameters that are found to be significant in Table 5.6 have a posterior probability > 0.95 in Table 5.5. However, the model in Table 5.5 allows us to infer something about the probability of our hypothesis, whereas the model in Table 5.6 does not.

The p-values given in Table 5.6 give us the probability that the same results or more extreme would be obtained, given that the null hypothesis is true ($P(data|H_0)$). They are a measure of evidence against the null hypothesis, but do not tell us how far we can be confident that the null hypothesis is true, nor how confident we can be that our alternative hypothesis is true (e.g., how confident we can be that the log(duration) coefficient really is negative, or that the intercept is positive) (Vasishth & Nicenboim, 2016, p.353). Under the assumptions of Frequentist statistics, if we hypothetically repeated our data collection and analysis hundreds of times, 95% of the time, the confidence intervals we obtained would contain the true value of the parameter; but we cannot say how likely it is that the specific confidence intervals that we have obtained this time (e.g., the 95% confidence interval [-0.12, -0.03] about the log(duration) coefficient) contain the true value of the parameter.

The probabilities of direction given in Table 5.5 tell us the posterior probability that each coefficient really is greater than or less than zero, given our data ($P(hypothesis|data)$) – e.g. the probability that the intercept is positive is 97%, while the probability that the coefficient for log(duration) is negative is 100%. For example, with the model reported in Table 5.5, it is correct to say that we can be 100% confident that an increase in log(duration) is associated with a decrease in onset F2, and we can be 95% confident that with an increase of 1 standard deviation in log(duration), the expected decrease in onset F2 is between 0.03 standard deviations and 0.12 standard deviations.

In sum, there are various assumptions that it is tempting to make from Table 5.6, but that are actually incorrect. When the same model is implemented using Bayesian methods, the results are more intuitive, allowing us to make statements about how likely a model coefficient is to be positive or negative, and giving us a credible interval within which we can be 95% confident that the true parameter value lies.

Table 5.5: Model implemented using brms: PRICE in *like* onset F2: adolescents and children

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
Intercept	0.28	0.01	0.59	0.97	0.10	1107	1.00
log(duration)	-0.07	-0.12	-0.03	1	0.86	3761	1.00
age	-0.18	-0.46	0.08	0.91	0.25	1300	1.00
sex	-0.21	-0.49	0.06	0.93	0.20	1289	1.00
age * sex	-0.3	-0.58	-0.02	0.98	0.08	985	1.00

Table 5.6: Model implemented using lme4 and lmerTest: PRICE in *like* onset F2: adolescents and children

Parameter	$\hat{\beta}$	**95% CI lower**	**95% CI upper**	**Std. Error**	**df**	**t-value**	**p-value**
Intercept	0.28	0.02	0.55	0.13	32.99	2.17	0.0374 *
log(duration)	-0.07	-0.12	-0.03	0.02	940.46	-3.35	0.0008 ***
age	-0.17	-0.44	0.08	0.13	33.10	-1.34	0.1896
sex	-0.21	-0.48	0.05	0.13	32.88	-1.65	0.1089
age * sex	-0.29	-0.55	-0.03	0.13	32.89	-2.24	0.0321 *

5.6 Summary

This chapter has set out the methodology for the acoustic analysis of the diphthongs FACE, PRICE and GOAT. It explained how the data was transcribed, how the vowel tokens were segmented and how the first and second formant frequencies were extracted. We also covered the vowel normalisation procedure – the modified Watt-Fabricius method, as is typically used in variationist sociolinguistics (Fabricius et al., 2009) – and the Trajectory Length (TL) measurement that is used in the current study to operationalise how diphthongal or monophthongal a vowel token is. Finally, Section 5.5 explained the motivation and the methods for conducting statistical analysis in the Bayesian paradigm.

Chapter 6

Diphthong variation in the speech of the adolescents

6.1 Introduction

Chapter 4 gave an overview of which MLE innovations were found in the current data. This chapter affords the opportunity to examine a subset of these features in more detail, by looking at shifting and monophthongisation of the diphthongs FACE, PRICE and GOAT (for more detail on these variables and the methodology used to analyse them acoustically, see Chapter 5). This thesis is concerned with whether and how MLE features are spreading in London, and the possibility that different innovations may arise in different parts of London. In order to address RQ1 of the study (see Chapter 1), we will examine vowel variation among the adolescent participants in relation to a number of language-internal and language-external factors.

The analysis of diphthong variation will be guided by the following questions:

1. What language-internal constraints govern vowel variation in this speech community? Does this align with what has previously been reported for MLE (e.g. S. Fox, 2015; Gates, 2018)?

2. Are there differences between the two Communities of Practice (CofP) in the realisations of these diphthongs?

3. What other social factors play a role? For example, speaker sex, and the area of London the individual comes from?

6.2 Predictions

This section will present predictions for the phonetic analysis, taking into account both the findings of previous MLE studies such as Cheshire et al. (2011) and Kerswill et al. (2008), and also the participant observation conducted as part of the current study (see Chapter 3). Table 6.1 (repeated from Chapter 5) summarises what more advanced MLE diphthongs would look like. The predictions regarding who will use more advanced versions of the MLE diphthongs are as follows:

1. It is expected that adolescents will show more diphthongal variants of these diphthongs in reading passage speech, compared to interview speech. It is hard to predict whether the onset qualities of the diphthongs will vary between speech styles, given that the MLE onset qualities of these diphthongs are quite similar to SSBE. However, if differences are found in the diphthong onset qualities, we expect to see a more advanced version of the MLE diphthongs in interview speech as compared to read speech.

2. Boys will show more advanced versions of the MLE diphthongs compared to the girls (Cheshire et al., 2011; Kerswill et al., 2008)

3. Members of the Studio CofP will show more advanced versions of the MLE diphthongs compared to the Youthclub CofP, given the previous studies that have found involvement with rap, hip-hop and grime music to correlate with multiethnolect use (Drummond, 2018a; Adams, 2018)

4. Outcode – i.e. the first part of the UK postcode (where [1] = same postcode as the youth centre; [2] = Northwest London; [3] = London but outside Northwest London) – is included as a factor in case MLE changes are being introduced to the youth centre by adolescents who live further away. Realistically, we do not expect great differences between adolescents from outcodes 1 and 2, seeing as the outcode 2 adolescents mostly attend high school near the youth centre, i.e. in outcode 1. However, the three participants who are categorised as outcode 3 (Kai, Tony and Denzel) have travelled from much further afield and might potentially be brokers of change, if MLE diphthongs are diffusing to this community from outside.

Table 6.1: Summary of MLE tendencies in terms of onset and monophthongisation of the diphthongs FACE, PRICE and GOAT

Diphthong	MLE tendency
FACE	• Closer realisation i.e. lower F1 • Monophthongisation
PRICE	• More front realisation i.e. higher F2 • Monophthongisation
GOAT	• More back realisation i.e. lower F2 • Monophthongisation

6.3 Methods

6.3.1 Participants & language-external/social factors

The participants in this project are described in more detail in the Fieldwork and Methodology chapters. For the phonetic analysis, only those whose recordings were of high enough quality to support an acoustic analysis were included. The participants who were included in the phonetic analysis are listed in Table 6.2. The language-external factors entered into the model were:

- Speaker sex – female or male; the two sexes may alternatively be referred to as girls and boys.
- Community of Practice (CofP). Two CofPs were identified: the Youthclub CofP, i.e. those adolescents whose primary purpose in coming to the youth centre was to socialise with their friends; and the Studio CofP, i.e. those adolescents whose primary purpose in attending the youth centre was to use the recording studio. For further details, see Chapter 3.
- Outcode. In the UK, the term "outcode" refers to the first four characters of a postcode. In the coding system adopted here, 1 = same outcode as the youth centre;

2 = a different outcode from the youth centre, but still Northwest London; 3 = outside Northwest London.

Table 6.2: Summary of background information on adolescent participants. Empty cells = information not given

Pseudonym	**Age**	**Sex**	**CofP**	**Outcode**	**Self-defined ethnicity**
Amanda	17	F	Youthclub	1	
Lola	17	F	Youthclub	1	Black African
Chris	16	F	Youthclub	1	Black African
Jessica	16	F	Youthclub	1	White British, White Irish
Shantel	16	F	Youthclub	1	White Irish
Chantelle	17	F	Youthclub	1	British Black Caribbean
Lucy	16	F	Youthclub	2	White British
Ali	16	M	Youthclub	2	Black British African
CB	16	M	Youthclub	2	
Ibrahim	16	M	Youthclub	2	
Joe	16	M	Youthclub	2	
Omar	16	M	Youthclub	2	
Sami	16	M	Youthclub	2	African, Arab
Tariq	16	M	Youthclub	2	Somali
ZR	16	M	Youthclub	1	
Khadir	17	M	Youthclub	1	Black British, British African
Matisse	19	M	Studio	1	
SD	20	M	Studio	1	
GW	20	M	Studio	1	Black British
Daniel	20	M	Studio	1	African
Moses	24	M	Studio	1	
Kai	17	M	Studio	3	Mixed Other
Denzel	18	M	Studio	3	Black African, British African
Tony	18	M	Studio	3	

6.3.2 Statistical model structure and modelling procedure

The dependent variables were as follows: F1 at the 20% time point for FACE (hereafter "onset F1"); F2 at the 20% time point for GOAT and for PRICE (hereafter "onset F2"); and Trajectory Length (TL) for all three diphthongs. TL was log-transformed before being entered into the models to satisfy the assumption of being able to take on any positive or negative value. All dependent variables were z-scored before being entered into the models, meaning that the units of the regression coefficients are standard deviations from the mean of the response variable.

There were two stages to the modelling procedure: first, a model was run containing the following predictors: log(duration), CofP, preceding environment, following environment, and interactions CofP*preceding environment and CofP*following environment, with by-speaker and by-word random intercepts, and by-speaker random slopes for preceding environment and following environment. The purpose of this model was to establish to what extent the two CofPs are similar or different in their realisations of these diphthongs, and whether they show the same language-internal constraints on this variation.

At the next stage, for each dependent variable two separate models were run including the data and relevant social factors of each CofP. The Youthclub CofP model in each case included log(duration), preceding environment, following environment, task (interview or reading passage), participant sex, outcode, and the following interaction terms: outcode*sex; outcode*preceding environment; outcode*following environment; outcode*task; sex* task. This model also contained by-speaker and by-word random intercepts, and by-speaker random slopes for preceding environment, following environment and task.

The Studio CofP model for each dependent variable included log(duration), preceding environment, following environment, task (interview or reading passage) and outcode, and the following interaction terms: outcode*task; outcode*preceding environment; outcode*following environment. It also included by-speaker and by-word random intercepts, and by-speaker random slopes for preceding environment, following environment and task.

In total, for this chapter, 2027 tokens of FACE, 1315 tokens of PRICE, 1046 tokens of PRICE in *like*, and 1915 tokens of GOAT were analysed. Tabulation of their distribution by social factors and preceding and following phonological environment can be found in Appendix A.1.

Log(duration) was z-scored before being entered into the models. The categorical predictors CofP, outcode, speaker sex, preceding environment, and following environment were sum-coded – the contrast matrices for these can be found in Appendix A.2. Task was dummy-coded, with interview (which accounted for the majority of vowel tokens) being the reference level. This means that for each level, the intercept represents average log(duration), interview speech rather than reading passage speech, and then the mean of the factor levels for each of the other categorical predictors.

6.3.3 The reporting of results

The following are reported in the presentation of results:

- The estimated regression coefficient ($\hat{\beta}$), i.e. the median of the posterior distribution
- 95% Highest Density Interval (HDI) about this estimate
- Probability of direction (PD): the probability that the effect is positive or negative – the proportion of the posterior distribution that is above or below 0. Close to 50% = low probability, close to 100% = high probability.

If $\geq 95\%$ of the posterior is $>$ or $<$ 0, and the 95% HDI does not include 0, we will conclude that there is strong evidence for a positive or negative effect (Tanner et al., 2019; Franke & Roettger, 2019; Nicenboim & Vasishth, 2016). If $\geq 95\%$ of the posterior is $>$ or $<$ 0, but the 95% HDI includes 0, we will say that there is marginal evidence for a positive or negative effect.

6.4 Results

An overview of the results will be presented before the results for the three diphthongs are presented in detail.

6.4.1 Overview

Figure 6.1 compares the vowel systems of the Youthclub and Studio CofP members. Both essentially show an MLE system, like that described by Kerswill et al. (2008): the onset of FACE is close-mid and near-front; the onset of GOAT is similar to FACE, although perhaps slightly more open; the onset of PRICE is central and open, overlapping with TRAP. The frontness of GOAT appears to differ slightly between the two CofPs: in the Studio CofP, GOAT is more back and overlaps with LOT, whereas in the Youthclub CofP, GOAT is similar in frontness to PRICE, does not overlap as much with LOT, and encroaches very slightly on FACE. At the same time, it can be seen that in the Youthclub CofP, there is a great deal of interspeaker variation in GOAT: some boys (shown as triangles) have a backed onset to GOAT, overlapping with LOT, while other participants, especially girls (shown as circles), have a central–front onset to GOAT.

Figure 6.2 excludes girls, and compares just the boys from the Youthclub and Studio CofP. Because only male data is presented, the unnormalised Hz formant frequencies are shown. With the girls' data removed, the differences between the two CofPs are reduced: on both plots, GOAT is backed, overlapping with LOT.

Figure 6.3 shows the onset data just for the Studio CofP. The plot compares the diphthong onsets of the Studio CofP people from the same postcode as the youth centre (1) with those who live in other parts of London (3). Some differences are immediately apparent. The young people from outcode 3 tend to have: a slightly closer FACE onset, compared to outcode 1; and a somewhat closer and backer GOAT onset, compared to outcode 1. Outcode 1 and 3 adolescents in this CofP appear similar to one another in the onset of PRICE.

Figure 6.4 makes the same comparison as Figure 6.3, but within the Youthclub CofP: adolescents from the immediately local area (1), vs. those from elsewhere in Ealing and neighbouring boroughs in Northwest London (2). The main difference between this plot and Figure 6.3 is that within the Youthclub CofP, no differences according to outcode stand out: it seems that adolescent boys from other parts of Ealing and Northwest London do not differ massively from boys who live very locally.

Figure 6.5 compares the female and male members of the Youthclub CofP. The females and males seem to have similar FACE onsets to one another. The girls have a slightly back PRICE onset relative to TRAP, and in comparison to the boys, for whom TRAP and PRICE overlap. There is variation within both sex groups as to the frontness of GOAT, as indicated by the spread of participant means, but on average, the girls have a central onset to GOAT, overlapping slightly with FACE, while the boys have a near-back onset to GOAT, overlapping with LOT.

Figures 6.6a, 6.6c and 6.6e show the smoothed trajectories of the three diphthongs over the 5 proportional distance time points (20%, 35%, 50%, 65% and 80%) for the Youthclub CofP. Because the key social predictor in the Youthclub CofP appears to be speaker sex rather than outcode of residence, these plots are faceted by sex. The F1 and F2 trajectories of FACE are very similar across the sexes, with an F1 starting point of approximately 1.0, and F1 finishing point of around 0.9. The trajectory shape of PRICE appears similar between the two sexes, but for boys, it has a higher F2 at both onset and offset, i.e. an overall more front realisation. For GOAT, the F1 trajectories are similar between the two sexes, but males and females diverge completely in the F2 of GOAT. For boys, the onset of GOAT is back (F2 0.9) and the offset even more back, while for girls, GOAT has a high F2 at onset (around 1.1) and an even higher F2 at offset.

Figures 6.6b, 6.6d and 6.6f show the smoothed trajectories of the diphthongs for the Studio CofP, separated by outcode. Just as was the case in the Youthclub CofP, FACE has an onset of around 1.0 on the normalised scale. For the Studio boys who are not from the area, there is next to no change in F1 over time, as indicated by the flat trajectory, though there is some change in F2. For PRICE, the trajectory is similar to that of the male Youthclub CofP members. For GOAT, the outcode 3 participants show a lower F2 trajectory throughout the vowel duration compared to the outcode 1 participants.

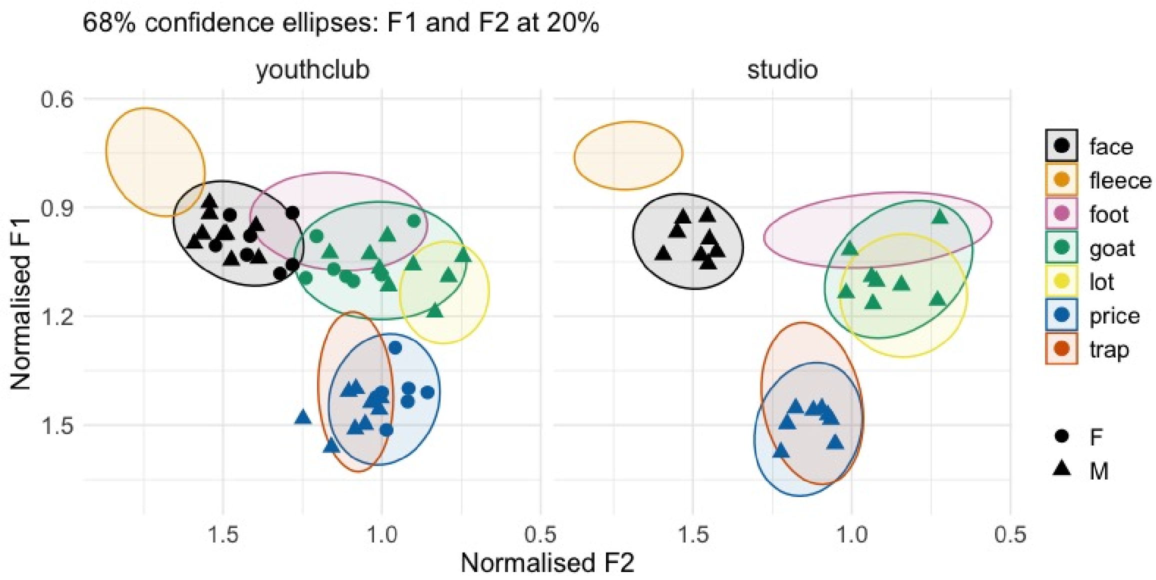


Figure 6.1: Vowel plot showing 1 standard deviation confidence ellipses for the mean F1 and F2 values of the diphthongs FACE, PRICE and GOAT, and the corner vowels FLEECE, TRAP, LOT and FOOT, separated by CofP (Youthclub vs. Studio) and with participant means for the diphthong onsets shown as points

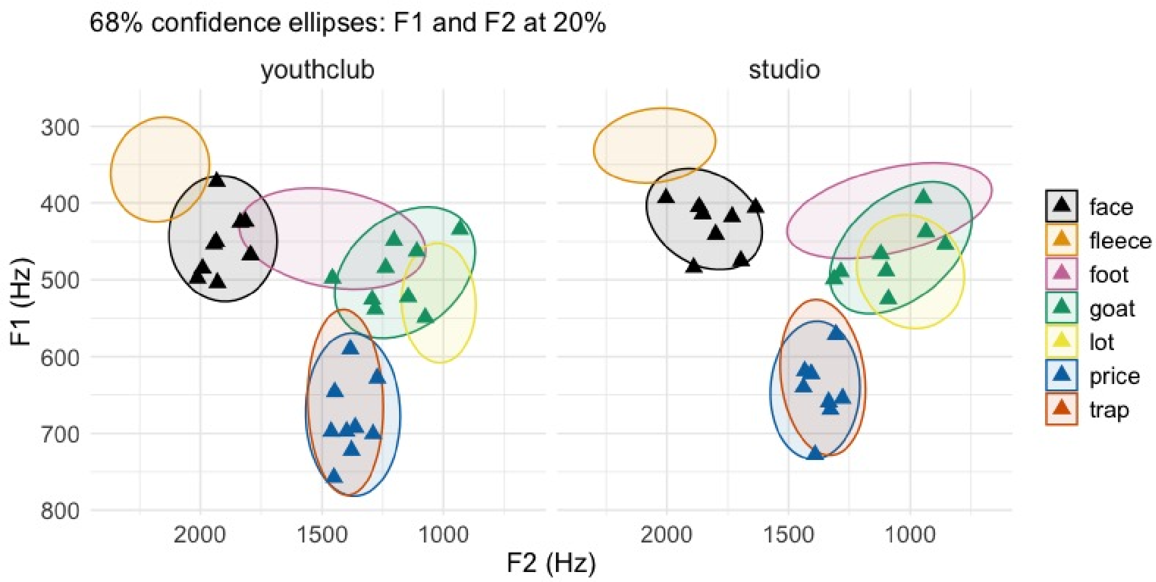


Figure 6.2: Vowel plot of male data, showing 1 standard deviation confidence ellipses for the mean F1 and F2 values of the diphthongs FACE, PRICE and GOAT, and the corner vowels FLEECE, TRAP, LOT and FOOT, separated by CofP (Youthclub vs. Studio) and with participant means for the diphthong onsets shown as points

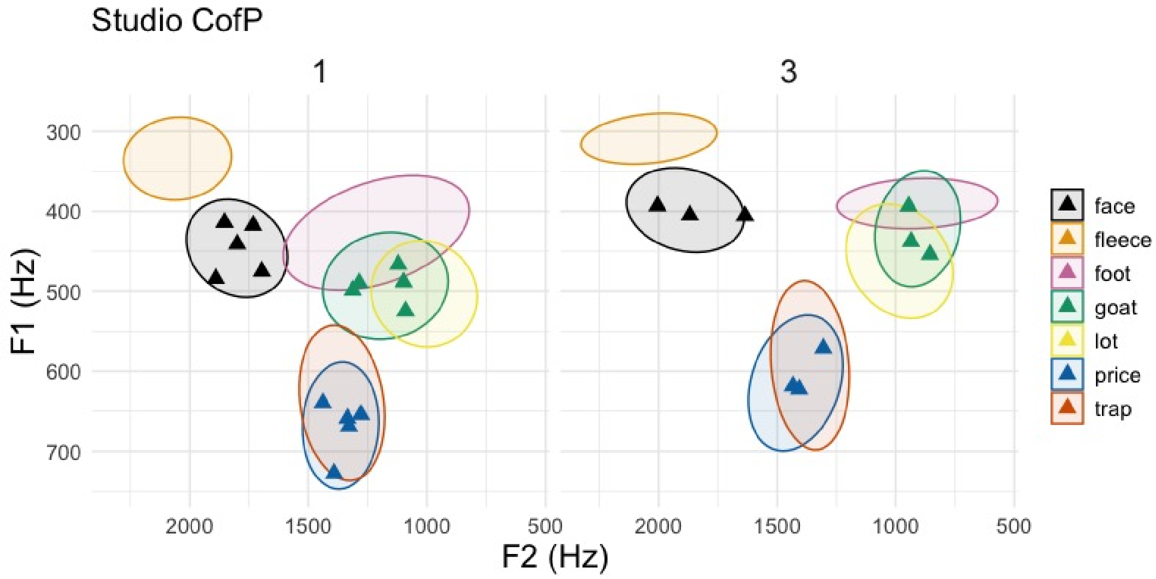


Figure 6.3: Vowel plot showing 1 standard deviation confidence ellipses for the mean F1 and F2 values of the diphthongs FACE, PRICE and GOAT, and the corner vowels FLEECE, TRAP, LOT and FOOT in the Studio CofP, separated by outcode (1 = same postcode as youth centre; 3 = not from Northwest London) and with participant means for the diphthong onsets shown as points

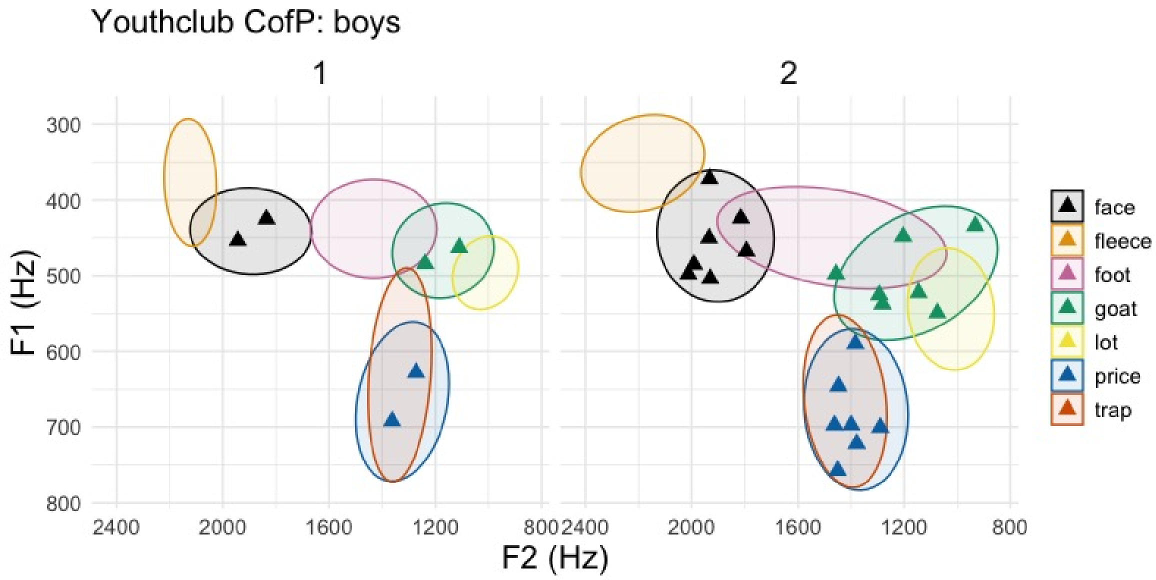


Figure 6.4: Vowel plot showing 1 standard deviation confidence ellipses for the mean F1 and F2 values of the diphthongs FACE, PRICE and GOAT, and the corner vowels FLEECE, TRAP, LOT and FOOT among the boys from the Youthclub CofP, separated by outcode (1 = same postcode as youth centre; 2 = elsewhere in Northwest London) and with participant means for the diphthong onsets shown as points

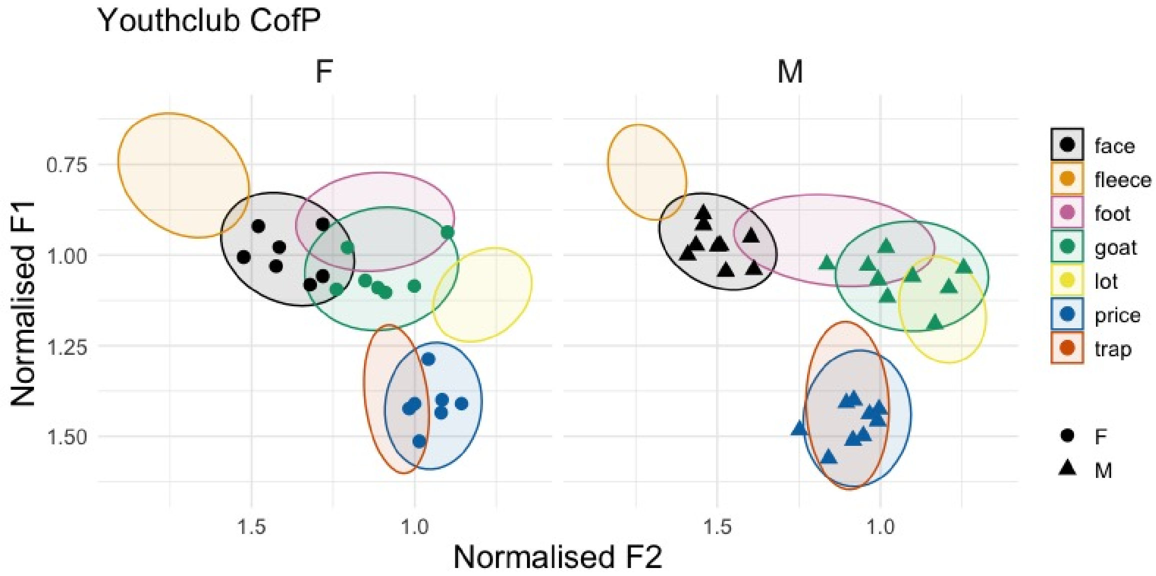
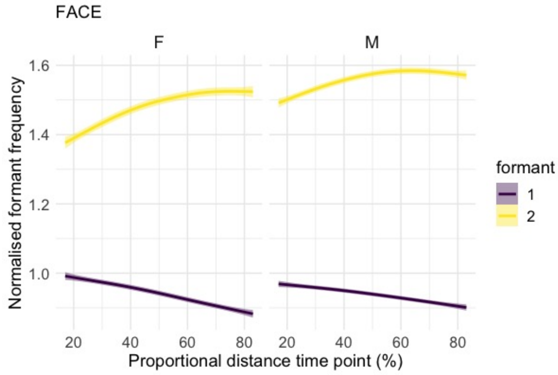
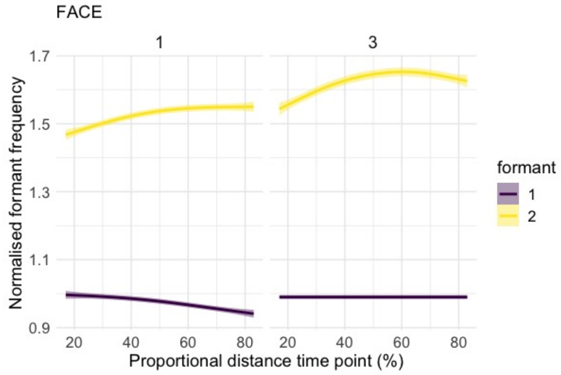


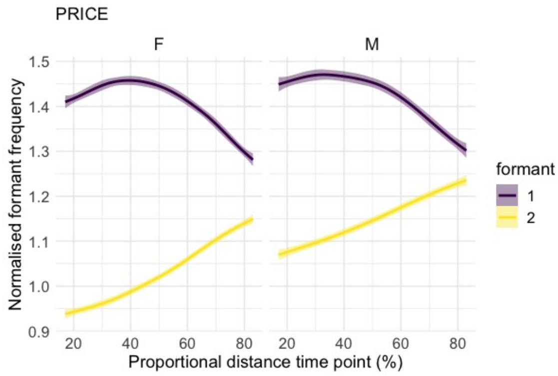
Figure 6.5: Vowel plot showing 1 standard deviation confidence ellipses for the mean F1 and F2 values of the diphthongs FACE, PRICE and GOAT, and the corner vowels FLEECE, TRAP, LOT and FOOT among male and female members of the Youthclub CofP, separated by speaker sex and with participant means for the diphthong onsets shown as points



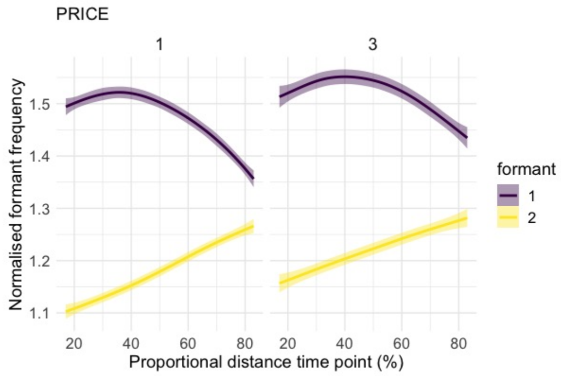
(a) FACE in the Youthclub CofP



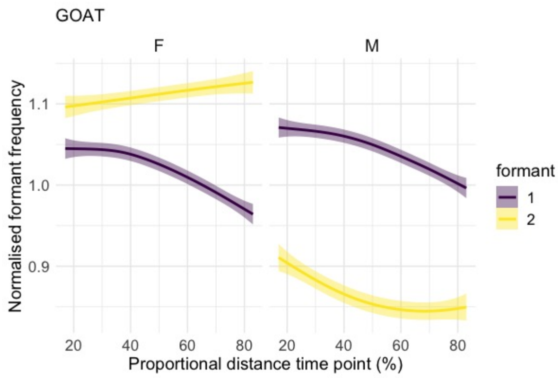
(b) FACE in the Studio CofP



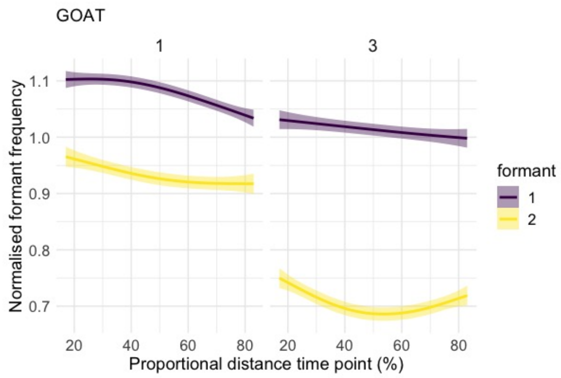
(c) PRICE in the Youthclub CofP



(d) PRICE in the Studio CofP



(e) GOAT in the Youthclub CofP



(f) GOAT in the Youthclub CofP

Figure 6.6: Smoothed normalised F1 (purple) and F2 (yellow) trajectories over the 20%, 35%, 50%, 65% and 80% time points for the Youthclub CofP (plots a, c, e) and the Studio CofP (plots b, d, f)

6.5 Results of statistical analysis

As in the section above, the results for the diphthong onsets are presented first in Section 6.5.1, followed by the analysis of variation in Trajectory Length in Section 6.5.2.

6.5.1 Diphthong onsets

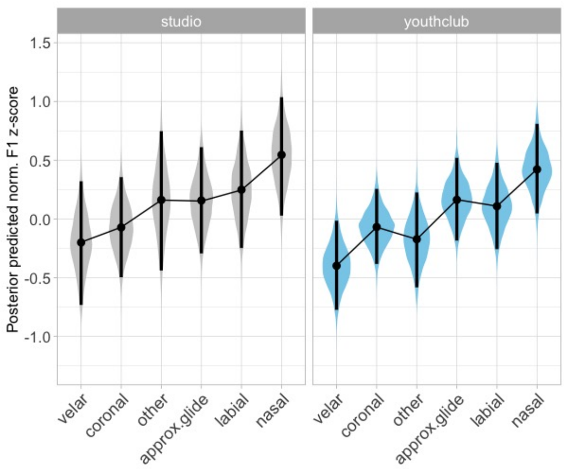
6.5.1.1 FACE: F1 at 20%

Comparing CofPs There was no evidence for a difference in FACE F1 between the two CofPs ($\hat{\beta}$=0.07; 95% HDI [-0.19, 0.32]; 70% PD).

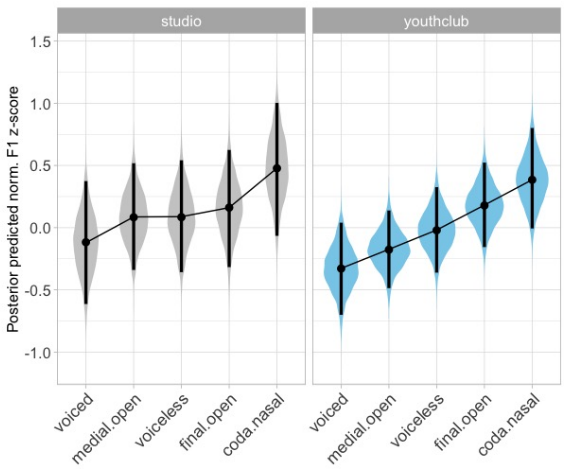
There was strong evidence that a preceding coronal is associated with a lower onset F1 ($\hat{\beta}$=-0.14; 95% HDI [-0.27, -0.02]; 99% PD); strong evidence that a preceding nasal is associated with a higher onset F1 ($\hat{\beta}$=0.41; 95% HDI [0.19, 0.62]; 100% PD) and marginal evidence that a preceding velar is associated with a lower onset F1 ($\hat{\beta}$=-0.3; 95% HDI [-0.61, 0.05]; 96% PD).

There was strong evidence that a coda nasal favours a higher onset F1 ($\hat{\beta}$=0.36; 95% HDI [0.15, 0.54]; 100% PD); marginal evidence that FACE occurring word-medially in an open syllable tends to have a lower onset F1 ($\hat{\beta}$=-0.12; 95% HDI [-0.24, 0.01]; 97% PD); and strong evidence that a coda voiced obstruent favours a lower onset F1 ($\hat{\beta}$=-0.30; 95% HDI [-0.47, -0.12]; 100% PD). This last finding is different from what was found by S. Fox (2015), who found that a coda voiceless obstruent favoured a raised, monophthongal FACE. However, the finding that a coda nasal favours a higher onset F1 is in line with the findings from Gates (2018).

No interactions of CofP and preceding or following environment were found, and indeed Figures 6.7a–6.7b suggest that preceding and following environment effects are relatively consistent across the two CofPs.



(a) FACE onset F1: preceding environment



(b) FACE onset F1: following environment

Figure 6.7: FACE: posterior predicted onset F1 z-score by CofP and by preceding and following environment. Posterior distributions shown as shaded areas, posterior medians as points and 95% HDIs as bars

Youthclub CofP There was no evidence for main effects of speaker sex, outcode, nor task on FACE onset F1, nor interactions of outcode and task, nor sex and outcode.

However, there was strong evidence for an interaction of sex and task ($\hat{\beta}$=0.26; 95% HDI [0.02, 0.49]; 98% PD): females are predicted to have a lower F1 in reading style compared to interview speech.

None of the interactions of outcode and preceding environment, nor outcode and following environment, reached the criteria for strong or marginal evidence.

Studio CofP In Section 6.4.1, it was suggested that members of the Studio CofP from outcode 3 had a more raised FACE onset compared to outcode 1. However, the model that included just the FACE tokens from the Studio CofP gave no evidence of a difference between outcodes ($\hat{\beta}$=-0.02; 95% HDI [-0.55, 0.50]; 54% PD). Similarly, there was no effect of task, nor any interaction between outcode and task.

There was also no evidence for any interactions of outcode and preceding environment, nor of outcode and following environment.

6.5.1.2 PRICE: F2 at 20%

Because of the frequency of the lemma *like* (as discourse marker, quotative etc.), tokens of *like* were modelled separately from the other PRICE words and will be dealt with in Section 6.4 below. There were 1170 tokens of *like* in the adolescent data vs. 2210 tokens of other PRICE words.

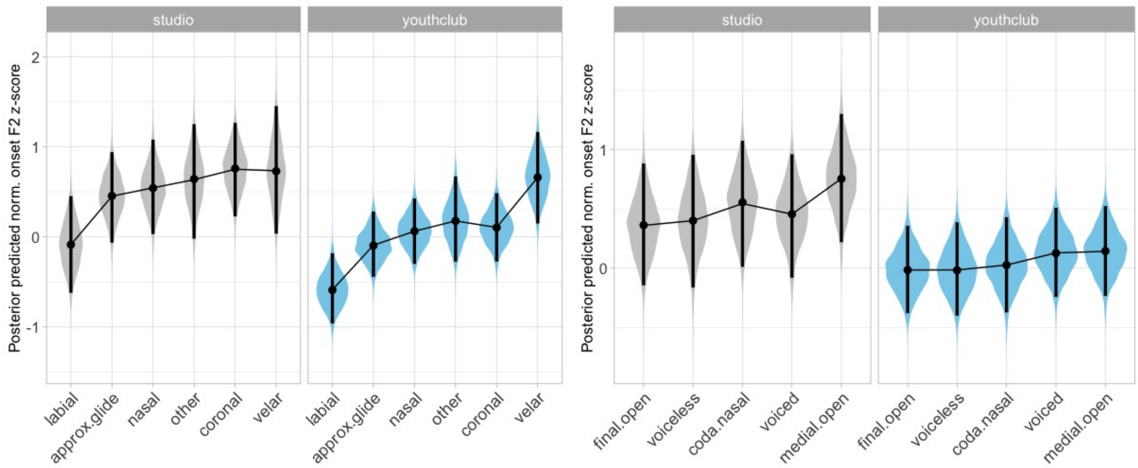
Comparing CofPs The initial model, comparing the two CofPs, gave some evidence – though not strong – that tbe Studio CofP shows a higher onset F2 than the Youthclub CofP ($\hat{\beta}$=0.22; 95% HDI [-0.07, 0.53]; 94% PD).

This model also showed stark simple effects of preceding segment: a preceding coronal is associated with a higher onset F2 ($\hat{\beta}$=0.15; 95% HDI [0.00, 0.30]; 97% PD) a preceding labial is associated with a lower onset F2 ($\hat{\beta}$=-0.61; 95% HDI [-0.79, -0.43], 100% PD); and a preceding velar is associated with a higher onset F2 ($\hat{\beta}$=0.70; 95% HDI [0.22, 1.17]; 100% PD).

There was marginal evidence for an interaction of CofP and a preceding coronal, though the effect size is very small ($\hat{\beta}$=0.10; 95% HDI [-0.02, 0.20]; 95% PD). It can be seen in Figure 6.8a that for the Studio CofP, a preceding coronal favours a slightly higher onset F2, while for the Youthclub CofP, a preceding coronal favours a slightly lower onset F2. Surprisingly, the evidence for an interaction of CofP and preceding velar was not strong (), even though Figure 6.8a suggests a difference between the two CofPs in this environment.

There was strong evidence that PRICE occurring in a word-medial open syllable tends to have a higher onset F2 ($\hat{\beta}$=0.17; 95% HDI [0.04, 0.31]; 99% PD). None of the preceding/following environment effects found by S. Fox (2015) or Gates (2018) were replicated.

None of the interactions of CofP and following environment quite reaches the criteria for strong/marginal evidence. It can be seen from Figure 6.8b that there is a relatively flat distribution of onset F2 across following environment in both CofPs.



(a) PRICE onset F2 by CofP and preceding environment

(b) PRICE onset F2 by CofP and following environment

Figure 6.8: PRICE posterior predicted normalised onset F2 z-score by CofP and by preceding and following environment. Posterior distributions shown as shaded areas, posterior medians as points and 95% HDIs as bars

Youthclub CofP Within the Youthclub CofP, there was no evidence for main effects of outcode or task. The main effect of sex did not quite reach the criteria for strong or marginal evidence ($\hat{\beta}$=-0.33; 95% HDI [-0.81, 0.08]; 94% PD): this means we can be 94% confident that the girls have a lower onset F2 than the boys.

There was not strong or marginal evidence for any interactions of outcode, sex and task.

There was not strong nor marginal evidence for interactions of outcode and preceding or following environment.

Studio CofP No evidence was found for any main effects of task or outcode, nor any interaction between them.

There was not strong nor marginal evidence for interactions of outcode and preceding or following environment.

6.5.1.3 PRICE in *like*: F2 at 20%

Comparing CofPs Section 6.5.1.2 found that the Studio CofP seems to show a higher F2 than the Youthclub CofP in the onset F2 of PRICE, although this did not reach our threshold for marginal evidence. In *like*, there is strong evidence that the Studio CofP shows a higher onset F2 ($\hat{\beta}$=0.48; 95% HDI [0.21, 0.77]; 100% PD).

Youthclub CofP The results for onset F2 in *like* in the Youthclub CofP were much the same as the results for other PRICE words: again, there appears to be a sex difference (girls show a lower onset F2) but this does not quite reach the criteria for marginal evidence ($\hat{\beta}$=-0.34; 95% HDI [-0.80, 0.08]; 94% PD).

There was no evidence for a main effect of outcode, nor for an interaction of sex and outcode.

Studio CofP Within the Studio CofP, the difference between outcodes did not reach the criteria for marginal evidence ($\hat{\beta}$=-0.38; 95% HDI [-1.03, 0.31]; 90% PD).

6.5.1.4 GOAT: F2 at 20%

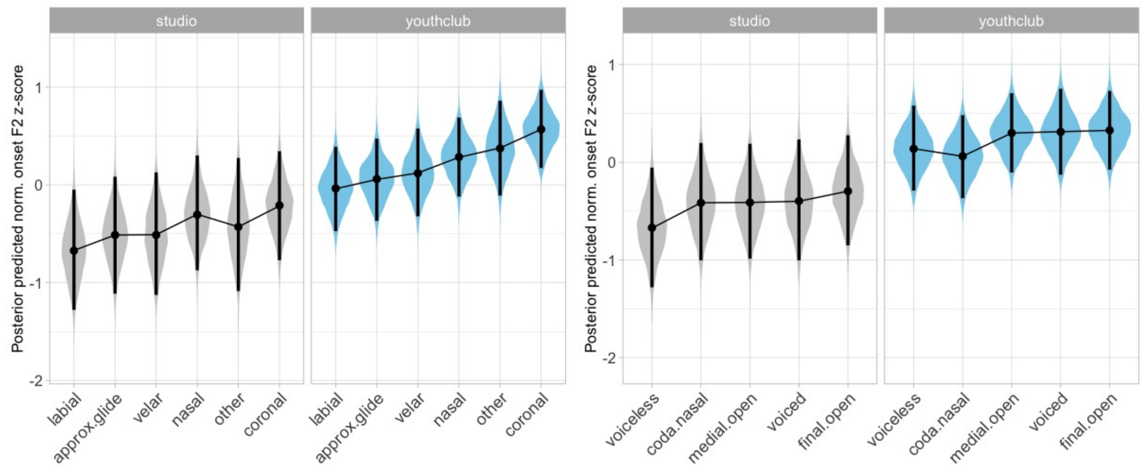
Comparing CofPs There was strong evidence for a difference between the two CofPs in the onset F2 of GOAT ($\hat{\beta}$-0.31; 95% HDI [-0.63, 0.02]; PD 97%): the Studio CofP members are predicted to have a lower onset F2 than the Youthclub CofP members. This confirms what was suggested by the vowel plots presented in Section 6.4.1.

In terms of preceding environment effects, there was strong evidence that a preceding coronal favours a higher onset F2 ($\hat{\beta}$=0.28; 95% HDI [0.13, 0.42]; PD 100%). There was also strong evidence that a preceding labial is associated with a lower onset F2 ($\hat{\beta}$=-0.25; 95% HDI [-0.41, -0.09]; PD 100%) There was some evidence that a preceding approximant or glide favours a lower onset F2, though this did not quite reach the criteria for marginal evidence ($\hat{\beta}$=-0.12; 95% HDI [-0.27, 0.03]; PD 94%).

None of the interactions of CofP and preceding environment reached the criteria for strong or marginal evidence. It can be seen from Figure 6.9a that the preceding segment effects are relatively consistent between the two CofPs, with a preceding coronal favouring a higher onset F2 and a preceding labial favouring a lower onset F2.

In terms of following environment effects, there was marginal evidence that GOAT occurring word-finally tends to have a higher onset F2 (e.g. in words such as *no, so, go*) ($\hat{\beta}$=0.12; 95% HDI [-0.02, 0.27]; PD 95%); and some evidence, though not strong, that a coda voiceless obstruent favours a lower onset F2 ($\hat{\beta}$=-0.29; 95% HDI [-0.66, 0.12]; PD 93%).

There was marginal evidence for a small interaction of CofP and the coda nasal environment ($\hat{\beta}$=0.10; 95% HDI [-0.01, 0.21]; PD 96%).



(a) GOAT onset F2 z-score by CofP and preceding environment

(b) GOAT onset F2 z-score by CofP and following environment

Figure 6.9: GOAT: posterior predicted onset F2 z-score by CofP and by preceding and following environment. Posterior distributions shown as shaded areas, posterior medians as points and 95% HDIs as bars

Youthclub CofP In the Youthclub CofP, there was marginal evidence that females have a higher onset F2 than boys ($\hat{\beta}$=0.43; 95% HDI [-0.08, 0.94]; PD 96%). Recall that this was also suggested above in Section 6.4.1 on the basis of Figure 6.5.

Main effects of outcode and task were not found, nor interactions of outcode, task and speaker sex.

In terms of interactions between preceding environment and outcode: there was strong evidence for a small interaction of outcode and preceding coronal ($\hat{\beta}$=-0.10; 95% HDI [-0.19, -0.01]; PD 99%); and strong evidence was found for an interaction of outcode and the preceding "other" environment ($\hat{\beta}$=0.22; 95% HDI [0.03, 0.41]; PD 99%). This is because within the Youthclub CofP, outcode 1 adolescents are predicted to have the highest onset F2 when preceded by an "other" segment, followed by coronal, whereas outcode 2 adolescents are predicted to have their highest onset F2 when GOAT is preceded by a coronal. It can be seen in Figure 6.10 that outcode 1 adolescents tend to have a higher onset F2 following an "other" segment, whereas outcode 2 adolescents tend to have a lower onset F2 following an "other" segment, and a higher onset F2 after a coronal obstruent.

No interactions of outcode and following environment were found.

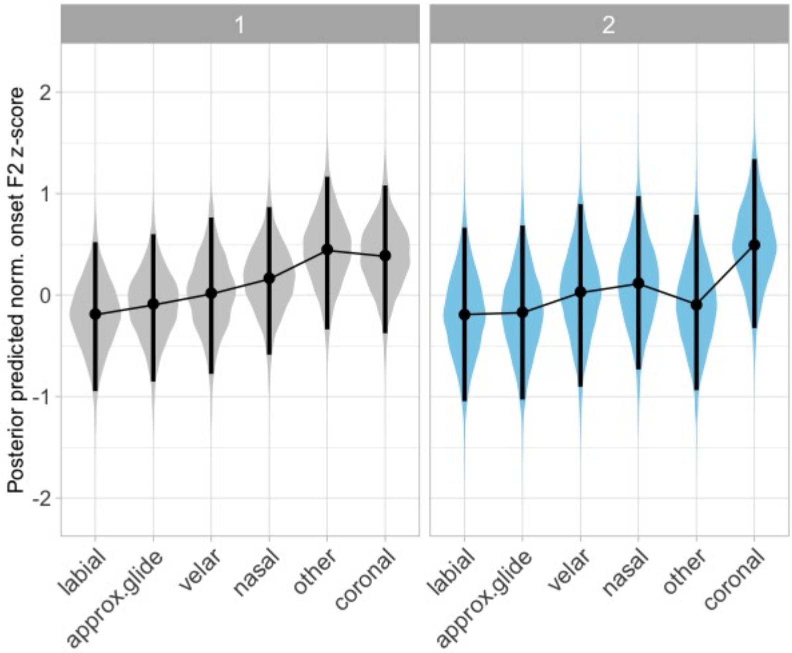


Figure 6.10: GOAT within the Youthclub CofP: posterior predicted onset F2 z-score by outcode (1 vs. 2) and by preceding and following environment. Posterior distributions shown as shaded areas, posterior medians as points and 95% HDIs as bars

Studio CofP Within the Studio CofP, there was strong evidence that adolescents from outcode 1 tend to have a higher onset F2 than adolescents from outcode 3 ($\hat{\beta}$=0.58; 95% HDI [0.26, 0.89]; PD 100%), supporting what was suggested by the vowel plots in Section 6.4.1 above (cf. Figure 6.3).

There was no effect of task, nor any interaction of outcode and task. No interactions of outcode and preceding environment or following environment were found.

6.5.1.5 Summary of results from diphthong onsets

FACE The results suggested no evidence for a difference between the CofPs in F1 – the Youthclub and Studio CofP are similar as far as the height of FACE goes.

Within the Youthclub CofP, there was an interaction of sex and task, such that males are predicted to show a somewhat lower F1 in interview speech compared to reading passage speech, whereas this effect is not present among females. Otherwise, within each CofP, social factors did not appear to affect variation in FACE F1.

Distinct effects of preceding and following environment were found on FACE F1, and these effects were consistent across the two CofPs: a a preceding or following nasal favours a higher onset F1; a preceding coronal favours a lower onset F1; a preceding velar favours a lower onset F1; and a coda voiced obstruent favours a lower onset F1.

PRICE The Studio CofP appears to have a higher onset F2, i.e. more front onset, in PRICE compared to the Youthclub CofP. There was strong evidence for this in *like*, though it did not quite reach the criteria for marginal evidence among the other PRICE words.

Within the Youthclub CofP, there was some evidence for a sex difference, such that the girls are predicted to have a lower onset F2 (more back onset) compared to the boys.

No main effects of outcode nor any interactions of outcode and preceding or following environment were found in either the Youthclub or the Studio CofP.

The preceding environment effects on F2 were consistent within each CofP, though they differed between CofPs. Across the two CofPs, the consistent preceding environment effects were that a preceding labial predicts a lower onset F2, and a preceding velar predicts a higher onset F2. However, there was an interaction between the preceding coronal environment and CofP, such that in the Studio CofP, a preceding coronal predicts a higher onset F2, while in the Youthclub CofP this effect is not found.

GOAT There was strong evidence that the Studio CofP favour a more back GOAT than the Youthclub CofP. Within the Studio CofP, outcode 3 adolescents had a more back GOAT than their peers, so it seems likely that the difference between the CofPs in GOAT F2 is actually due to the outcode 3 adolescents.

Within the Youthclub CofP, a sex difference was found such that girls are likely to have a more front GOAT onset compared to boys.

A preceding coronal was found to favour a higher onset F2, and this effect was consistent across the two CofPs. However, within the Youthclub CofP, the effect of a preceding coronal was carried by adolescents from outcode 2. Thus, there is a complex interaction of social factors with the preceding coronal environment.

6.5.2 Trajectory Length

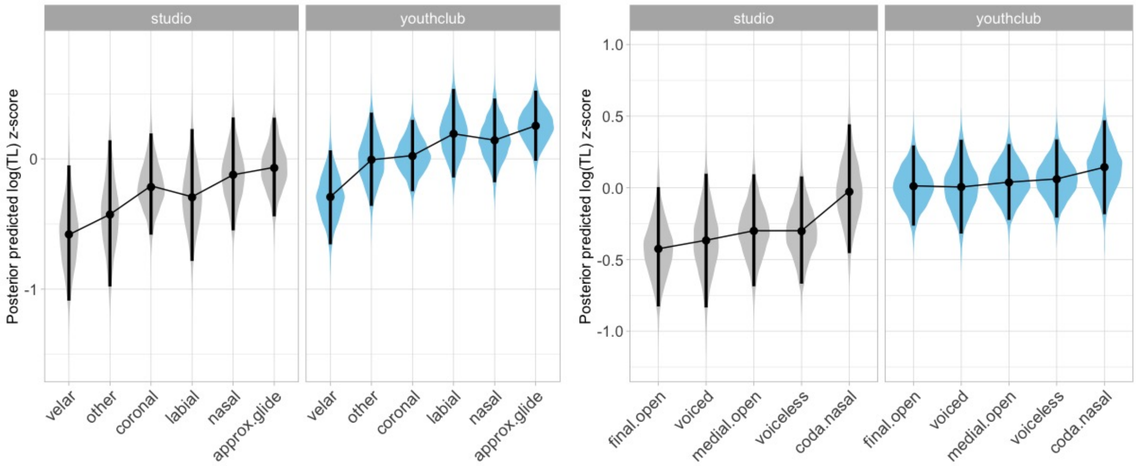
6.5.2.1 FACE: Trajectory Length

Comparing CofPs The initial model found some evidence, though not strong, for a difference between the two CofPs: the Studio CofP are estimated to have a smaller TL than the Youthclub CofP ($\hat{\beta}$=-0.17; 95% HDI [-0.39, 0.06]; 93% PD).

There was strong evidence that a preceding /l, ɹ, w, j/ is associated with a greater TL ($\hat{\beta}$=0.21; 95% HDI [0.09, 0.33]; 100% PD); and strong evidence that a preceding velar is associated with a smaller TL ($\hat{\beta}$=-0.44; 95% HDI [-0.75, -0.12]; 100% PD). The first finding again diverges from the findings of S. Fox (2015), as Fox found that a preceding /l, w/ favoured a more monopthongal realisation of FACE.

There was strong evidence that a coda nasal favours a greater TL ($\hat{\beta}$=0.18; 95% HDI [0.02, 0.33]; 99% PD); and some evidence, though not reaching the criteria for marginal, that FACE occurring word-finally has a smaller TL ($\hat{\beta}$=-0.09; 95% HDI [-0.20, 0.03]; 94% PD). This latter condition is in line with S. Fox (2015), who found that the word-final environment favoured newer variants of FACE.

None of the interactions of CofP and preceding or following environment reached the criteria for strong or marginal evidence.



(a) FACE TL by CofP and preceding environment (b) FACE TL by CofP and following environment

Figure 6.11: FACE posterior predicted normalised onset F1 z-score and log(Trajectory Length) z-score by CofP and by preceding and following environment. Posterior distributions shown as shaded areas, posterior medians as points and 95% HDIs as bars

Youthclub CofP No main effects of sex, outcode or task, were found, nor interactions between these. Similarly, there was no evidence for any interactions of outcode and preceding or following environment.

Studio CofP There were no main effects of outcode and task, nor any interaction of outcode and task. Similarly, no interactions of outcode and preceding or following environment were found.

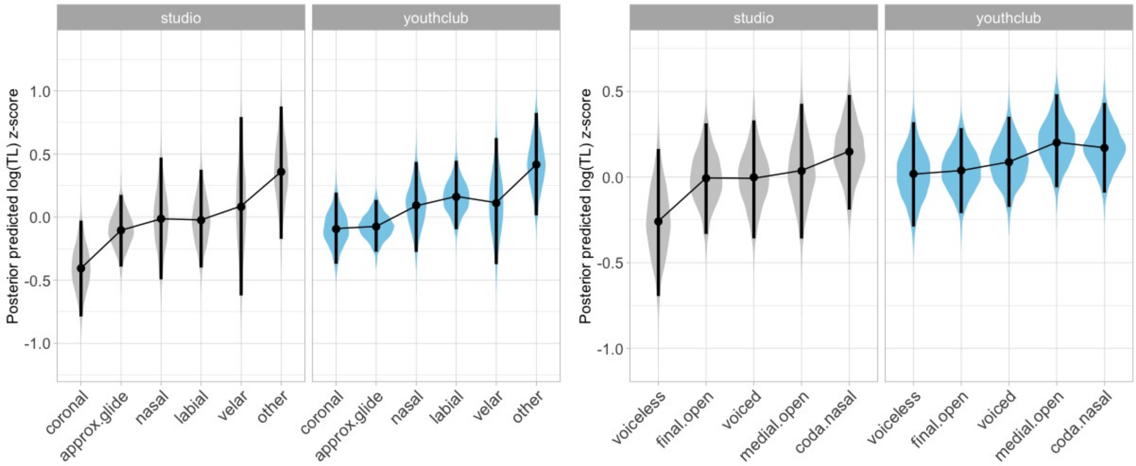
6.5.2.2 PRICE: Trajectory Length

Comparing CofPs The initial model comparing TL between the CofPs suggested that there is little to no difference between the two ($\hat{\beta}$=-0.06; 95% HDI [-0.22, 0.11]; 78% PD).

There was marginal evidence that a preceding /l, ɹ, w, j/ favours a smaller TL ($\hat{\beta}$=-0.13; 95% HDI [-0.27, 0.00]; 97% PD); strong evidence that a preceding coronal favours a smaller TL ($\hat{\beta}$=-0.29; 95% HDI [-0.48, -0.09]; 100% PD); strong evidence that a preceding "other" favours a greater TL ($\hat{\beta}$=0.35; 95% HDI [0.05, 0.66]; 99% PD).

There was not strong nor marginal evidence for any simple following environment effects, although there was some evidence that a coda nasal favours a greater TL ($\hat{\beta}$=0.12; 95% HDI [-0.04, 0.27]; 93% PD). This contrasts with the results of S. Fox (2015), who found that a following nasal favoured newer i.e. more monophthongal variants of PRICE.

There was not strong nor marginal evidence for any interactions of CofP and preceding or following environment, suggesting that these effects are relatively consistent across the two CofPs. This is also suggested by Figures 6.12a–6.12b.



(a) PRICE: posterior predicted log(TL) z-score by CofP and by preceding environment

(b) PRICE: posterior predicted log(TL) z-score by CofP and by following environment

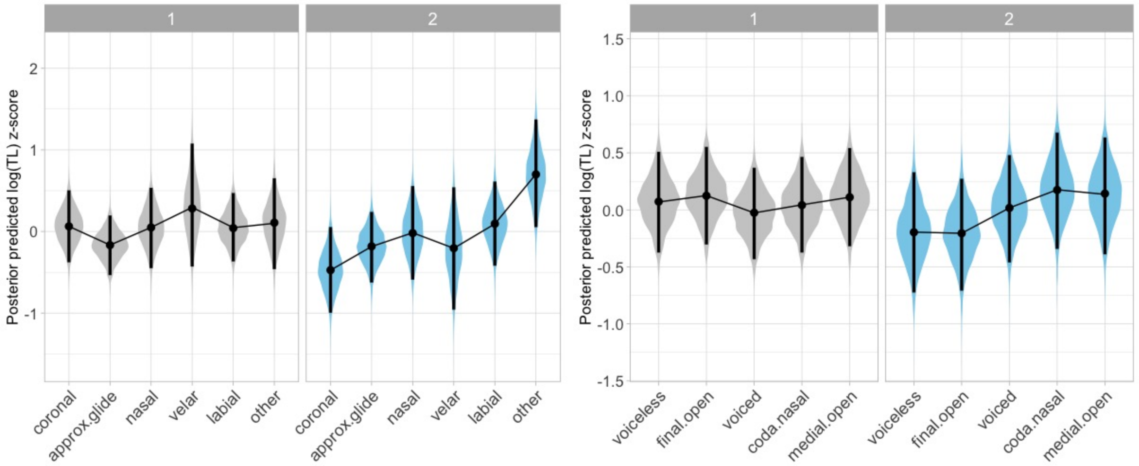
Figure 6.12: PRICE: posterior predicted log(TL) z-score by CofP and by preceding and following environment. Posterior distributions shown as shaded areas, posterior medians as points and 95% HDIs as bars

Youthclub CofP Within the Youthclub CofP, there was marginal evidence that girls are likely to have a greater TL than boys ($\hat{\beta}$=0.21; 95% HDI [-0.01, 0.46]; 96% PD). There was no evidence for main effects of outcode or task, nor interactions between outcode, task and sex.

Some interactions between outcode and preceding environment were found. There was strong evidence for an interaction between outcode and preceding coronal ($\hat{\beta}$=0.23; 95% HDI [0.00, 0.43]; 98% PD); and strong evidence for an interaction between outcode and preceding “other” ($\hat{\beta}$=-0.34; 95% HDI [-0.62, -0.08]; 99% PD)

It can be seen from Figure 6.13a that within the Youthclub CofP, adolescents from outcode 1 and 2 are predicted to show different preceding environment constraints on PRICE TL.

There was also strong evidence for an interaction between outcode and the word-final environment ($\hat{\beta}$=0.12; 95% HDI [0.02, 0.23]; 99% PD), though the effect size is small (cf. Figure 6.13b).



(a) PRICE: posterior predicted log(TL) z-score by preceding environment and outcode

(b) PRICE: posterior predicted log(TL) z-score by following environment and outcode

Figure 6.13: PRICE in the Youthclub CofP: posterior predicted log(TL) z-score by outcode (1 vs. 2), and by preceding and following environment. Posterior distributions shown as shaded areas, posterior medians as points and 95% HDIs as bars

Studio CofP Within the Studio CofP, there was no evidence for main effects of task or outcode.

There was marginal evidence for the following interactions of outcode and preceding environment: outcode and preceding approximant/glide ($\hat{\beta}$=-0.22; 95% HDI [-0.46, 0.02]; 97% PD); outcode and preceding coronal ($\hat{\beta}$=-0.26; 95% HDI [-0.54, 0.04]; 96% PD); and some evidence for an interaction of outcode and preceding velar ($\hat{\beta}$=0.43; 95% HDI [-0.10, 1.02]; 94% PD). It can be seen from Figure 6.14 that adolescents from outcode 1 have different preceding environment constraints compared to those from outcode 3.

None of the interactions of outcode and following environment reached the criteria for strong or marginal evidence.

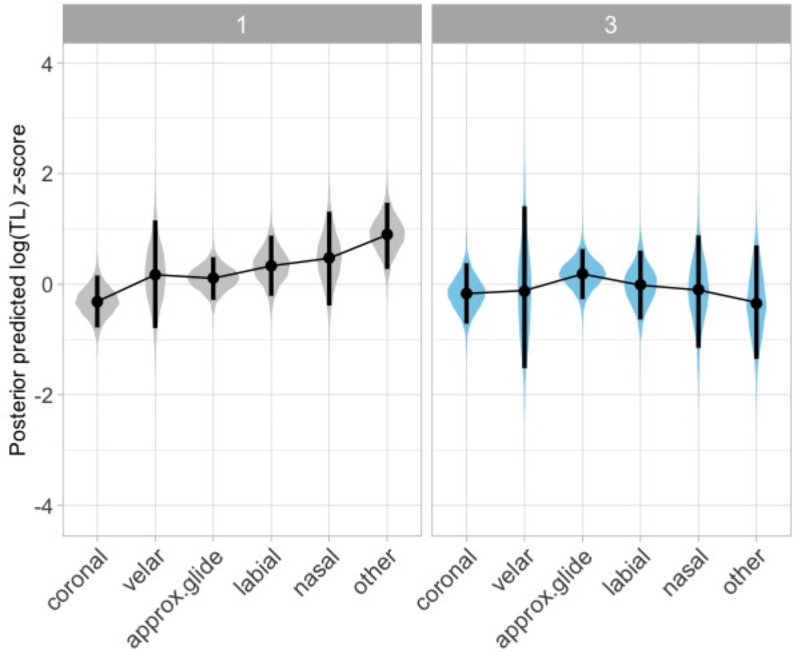


Figure 6.14: PRICE within the Studio CofP: posterior predicted log(TL) z-score by outcode (1 vs. 3) and by preceding environment. Posterior distributions shown as shaded areas, posterior medians as points and 95% HDIs as bars

6.5.2.3 PRICE in *like*: Trajectory Length

Comparing CofPs There was no evidence for a difference between the two CofPs in the log(TL) of *like* ($\hat{\beta}$=0.02; 95% HDI [-0.19, 0.22]; 47% PD).

Youthclub CofP Within the Youthclub CofP, there was no evidence for main effects of sex or outcode, nor for the interaction of sex and outcode.

Studio CofP Within the Studio CofP, there was no effect of outcode ($\hat{\beta}$=0.01; 95% HDI [-0.47, 0.49]; 52% PD).

6.5.2.4 GOAT: Trajectory Length

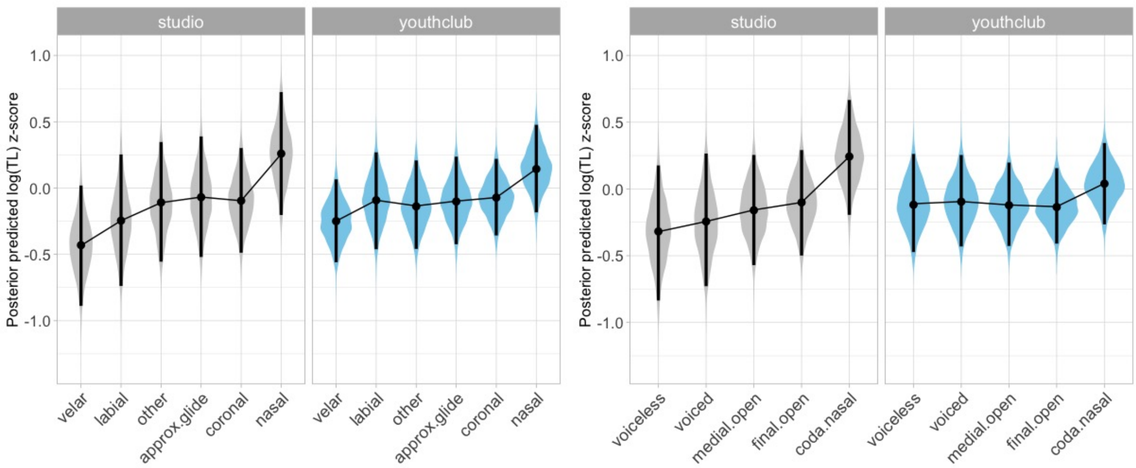
Comparing CofPs The model suggests that the CofPs do not differ greatly in the TL of GOAT ($\hat{\beta}$=-0.01; 95% HDI [-0.24, 0.21]; PD 56%).

As to preceding environment effects: there was strong evidence that a preceding nasal favours a greater TL ($\hat{\beta}$=0.30; 95% HDI [0.14, 0.46]; PD 100%); and strong evidence that a preceding velar (e.g. in words such as *go*) favours a smaller TL ($\hat{\beta}$=-0.34; 95% HDI [-0.61, -0.06]; PD 99%).

No evidence was found for interactions between CofP and preceding environment, suggesting that the preceding environment effects are relatively consistent across the two

CofPs. Figure 6.15a shows that in both CofPs, a preceding velar favours a more monophthongal realisation and a preceding nasal favours a more diphthongal realisation – though these effects appear more pronounced in the Studio CofP.

In terms of following environment effects, there was strong evidence that a coda nasal favours a greater TL ($\hat{\beta}$=0.24; 95% HDI [0.08, 0.42]; PD 100%). Although this effect appears to be stronger in the Studio CofP (Figure 6.15b), there was not quite strong evidence for an interaction with CofP ($\hat{\beta}$=0.12; 95% CI [-0.06, 0.27]; PD 92%).



(a) GOAT log(TL) z-score by CofP and preceding environment

(b) GOAT log(TL) z-score by CofP and following environment

Figure 6.15: GOAT: posterior predicted log(TL) z-score by CofP and by preceding environment and coda type. Posterior distributions shown as shaded areas, posterior medians as points and 95% HDIs as bars

Youthclub CofP In the Youthclub CofP, there was not strong nor marginal evidence for main effects of outcode, sex or task, nor for any interactions between these factors.

One interaction between outcode and preceding environment was found: there was strong evidence for an interaction of outcode and preceding labial ($\hat{\beta}$=0.27; 95% CI [0.00, 0.55]; PD 97%). It can be seen from Figure 6.16 that while the outcode 1 adolescents favour a greater TL when there is a preceding labial, the outcode 2 adolescents are predicted to have a more monophthongal GOAT after a labial consonant. This means that within the Youthclub CofP, adolescents from outcodes 1 and 2 differ from one another in the preceding environment constraints on both the onset F2 and the Trajectory Length of GOAT.

There was not strong nor marginal evidence for any interactions of outcode and following environment.

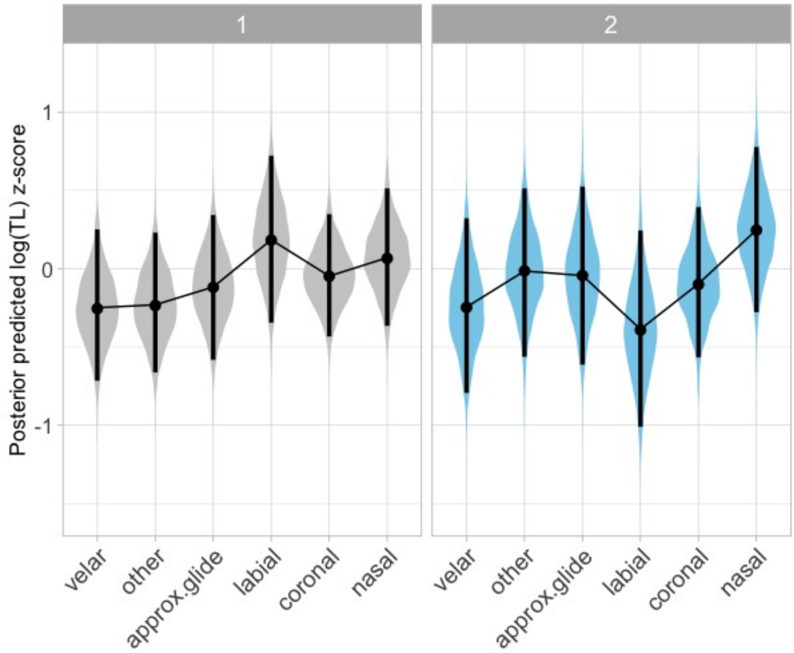


Figure 6.16: GOAT within the Youthclub CofP: posterior predicted log(TL) z-score by outcode and by preceding environment. Posterior distributions shown as shaded areas, posterior medians as points and 95% HDIs as bars

Studio CofP Within the Studio CofP, there was no evidence for a main effect of outcode or task, nor any interaction between the two.

No interactions between outcode and preceding or following environment were found.

6.5.2.5 Summary of results from diphthong trajectories

FACE There was some evidence that the Youthclub CofP have a more diphthongal FACE vowel than the Studio CofP, but this did not reach the criteria for strong evidence. Otherwise, social factors did not appear to affect variation in FACE TL.

There was some evidence for preceding and following environment effects on FACE TL. In terms of preceding environment effects: a preceding velar favours a smaller TL as well as a closer onset (cf. Section 6.5.1.5), i.e. more MLE-like realisation; and a preceding approximant or glide favours a greater TL.

In terms of following environment effects, a coda nasal predicts a more diphthongal realisation as well as a more open onset (cf. Section 6.5.1.5), i.e. a less MLE-like realisation; and FACE occurring word-finally is predicted to be more monophthongal.

PRICE The two CofPs did not differ from one another in the TL of PRICE, either in *like* or in the other PRICE words.

Within the Youthclub CofP, the girls have a more diphthongal realisation of PRICE compared to the boys generally, but this is not the case in *like*: in *like*, boys and girls are similarly monophthongal.

Whule main effects of outcode were not found in either the Youthclub CofP or in the Studio CofP, complex interactions of outcode and preceding and following environment were found in both CofPs.

GOAT The social factors did not appear to affect variation in GOAT TL.

6.6 Qualitative analysis of variation in PRICE and GOAT

Given the complex kinds of variation uncovered in PRICE and GOAT, this section takes a closer look at the uses of variation in these diphthongs in interaction.

Other researchers (e.g. Gates, 2018; Drummond, 2018b) have highlighted the issues associated with applying statistical analysis to "tease apart the effects of social factors on language variation in diverse communities" (Gates, 2018, p.43). Moreover, Third Wave sociolinguistics emphasises that sociolinguistic variation has to look beyond the correlation of linguistic variation with social categories, and towards how variables "connect to the social through their role in enacting and re-enacting personae" (Eckert, 2016). Relatedly, in the study of multiethnolects, Quist (2008) has drawn attention to the importance of combining variety-focused approaches with practice-focused analysis. In this spirit, while a full "stylistic practice" analysis is beyond the scope of this chapter, this section presents two examples of how MLE diphthongs are used interactionally by the Youthclub CofP. The quantitative results showed that for several variables, the Studio CofP was ahead of the Youthclub CofP in their use of MLE features. Nonetheless, these examples indicate that members of the Youthclub CofP may "turn up" their use of MLE diphthongs for interactional purposes.

6.6.1 PRICE

Recall that in the analysis of the onset F2 of PRICE, it was found that a preceding velar favoured a higher onset F2 in the Youthclub CofP, but that this is not so clearly the case for the Studio CofP (see Figure 6.8a). On closer inspection, this seems to be not necessarily a genuine language-internal constraint; this may actually be due to the frequency of the word *guy* in the Youthclub data, and the discourse contexts in which this word tends to appear. In the Studio CofP, the words with an initial velar are roughly evenly distributed between *kind, guy, kinda, guys* and *kindness* – each has fewer than 10 tokens. In the PRICE tokens from the Youthclub CofP, the words with an initial velar are *guy, guy's, kinda, kindness, guys, disguise, mankind* and *skype*, and each of these has fewer than 10 tokens except for *guy*, of which there are 58 tokens. In the data from the Youthclub CofP, the word *guy* tends to appear in narrative contexts, as in the examples below.

(69) a. cos another **guy** got er caught with me

b. and this **guy** he came back he ran as well but then he came back

c. and they don't know that cos they just kicking and punching the **guy**

d. and some **guy** was like to me "you fucking mussi" and he ran off the bus

e. some **guy** he whipped out a machete and started swinging it

A recurring finding in variationist sociolinguistics is that in personal narratives, speakers use a higher proportion of vernacular variants. Because of this, the sociolinguistic interview is designed in such a way as to elicit highly involved narratives, for example the "danger of death" question, in order to tap speakers' vernacular. This offers one explanation as to why members of the Youthclub CofP show a fronted PRICE in the word *guy* – the examples above indicate tales of danger, daring, and barely believable events. However, regardless of criticisms of the attention-to-speech model of style (e.g. Bell, 1984; Coupland, 1980), I would not see these excerpts as instances when the interviewees are necessarily paying minimum attention to speech – if anything, these examples highlight points in the interview when the interviewee is paying extra attention to his/her self-presentation. Bamberg and Georgakopoulou (2008) see narratives as "functional in the creation of characters in space and time, which in turn are instrumental for the creation of positions vis-a-vis co-conversationalists": narratives are fundamental in the narrator's portrayal of a fictional self, and in the narrator's self-positioning in the current interaction. These uses of *guy* can be seen as instances where the interviewee highlights his or her bravery when confronted with an anonymous assailant (or in opposition to a coward, in examples (a) and (c)), and simultaneously positions him- or herself as knowledgeable and streetwise relative to either the peer who is also present for the interview, or the interviewer who is an outsider asking to hear the story. In this respect, these instances of fronted PRICE in *guy* also speak to the concept of personae in variationist sociolinguistics (e.g. D'Onofrio, 2015; Eckert, 2016; Podesva, 2007): PRICE is fronted when the speaker is indexing a persona that is tough, who doesn't snitch on the other guy that got caught with him, and who is unphased by the sight of a machete.

6.6.2 GOAT-backing and -fronting

This section aims to tentatively suggest that the indexicality associated with GOAT-backing makes it useful for taking an emphatic stance, in intensifier *so*.

The process of transcribing and coding the interviews suggested a wide range of variation in GOAT frontness/backness, from a front monophthong similar to [ɵː], to a back monophthong, [oː] or [ɔʊ]. Moreover, other studies indicate that there may be stylistic or interactional factors at play in the use of fronted or backed GOAT. Kerswill (2016) presents a case of style-shifting among 18-year-old Afro-Caribbean girls. One of them, Courtney, uses [əʊ] for GOAT at the beginning of the sociolinguistic interview, but when she is joking around with her friends, the variant she uses is [ɔʊ]. Conversely, Oxbury (2016) found that a group of 11-year-old girls in Hackney favoured a fronted, monophthongal GOAT when in the playground with their friends, but when reading a wordlist in their interviews, they used a back GOAT variant.

The first example is from Amanda, who typically shows quite a fronted GOAT. Figure 6.17 shows that in Amanda's GOAT tokens, there is one outlier. Its F2 at onset is < 1000Hz, overlapping with her LOT tokens, and the direction of its glide is up and backwards. Initially, this looks like an erroneous measurement, but in the recording, this token can be heard as extremely back and monophthongal. It occurs in *so* as an intensifier. In this example, Amanda is trying to explain what the estate where the youth centre is situated used to be like when she was growing up, and how that compared to how the estate is now, at the time of the interview. She pauses and appears to be searching for the right words, with pauses of around 0.2 seconds between her utterances. She upgrades her assessment of how rough the estate used to be from "more rough" to "so rough" to "bare rough". Amanda may be seen as indexing a stance of authority and/or authenticity, which she does by employing backed GOAT, saying "I remember", and then using the MLE intensifier *bare*.

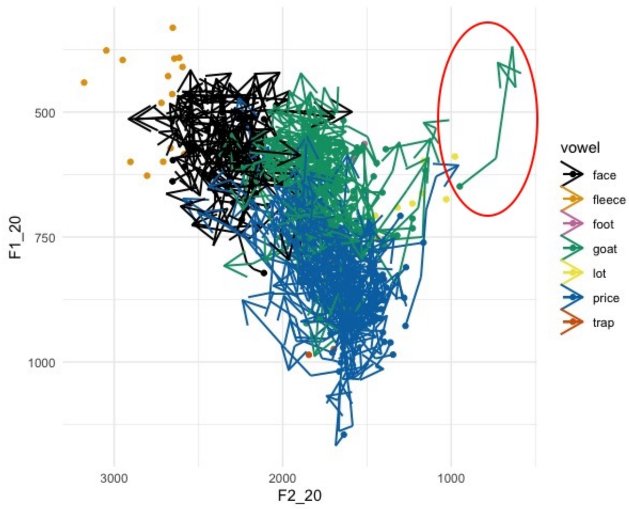


Figure 6.17: Amanda's vowel tokens. A red oval identifies the outlying GOAT token

(70) "used to be *so* rough"

01	Amanda	it used to be
		(0.25)
02	Amanda	more rough than it is n-
03	Amanda	now it's not even that rough
		(0.13)

04	Amanda	but
		(0.26)
05	Amanda	before used to be **so:** rough
06	Amanda	I remember
07	Amanda	like
		(0.23)
08	Amanda	used to be bare rough

A similar example of GOAT-backing appears in Example 71. Chris also uses backed GOAT in intensifier *so*. Chris uses the lemma *so* three times in Example 71, twice as a connective in l.1 and l.5, and then as an intensifier in l.8, and in this short extract she varies between three different variants of GOAT: reduced [ə], SSBE-like [əʊ] and backed, monophthongal [oː] in the instance of intensifying *so*. Phonetic transcription is given in order to show Chris's categorical use of t-glottaling, l-vocalisation in l.2 and variable (dh)-stopping throughout: variation in GOAT is accompanied by variable and categorical use of other MLE variables.

Both of these examples show adolescent girls adopting an extremely backed GOAT apparently for emphasis, in intensifying *so*. This suggests that the variant has accrued indexical meaning that allows it to be recruited in stance-taking. The stance they are taking is similar to what Sharma (2018) calls "the 'real me', signalling stripped-down frankness and personal commitment": Amanda is trying to convey to the interviewer how "rough" the estate she grew up on was, and Chris is emphasising that she and her friend made a truly realistic replica gun. It may be that, as per Sharma (2018), these girls are using "their dominant or default style" to do this, just as the girl in Kerswill (2016) uses backed GOAT to talk to her friend, but not to the interviewer; it is known that interviews only capture a snapshot of a speaker's available speech forms, and may not be representative of that speaker's repertoire (Rickford, 2014). Alternatively, if we assume that, for example, Amanda's GOAT is by default central rather than back (cf. Figure 70), adopting the backed GOAT variant that is mostly associated with boys from the Studio CofP is interactionally useful to her at that moment, to take a stance of authority and authenticity.

These examples demonstrate the need for further analyses of intraspeaker variation in the use of these diphthongs, and further interactional analysis in order to illuminate the meanings that this variation has in interaction.

(71) "*so* much better"

01	Chris	cos that was the last lesson, **so** I was in the class after that
		kz ðaʔ wz ðə lɑs lɛsən sə ɐ wz ɪn ðə klɑs ɑftə ðaʔ
		(0.50)

02	Chris	and they still had the shape out of the gun
		an ðeɪ stɪw had ðə ʃe(ɪ)p aʊʔ əf də gʌn
		(0.29)
03	RO	mm
04	Chris	mhm
		mhm
		(0.35)
05	Chris	**so** then me
		səʊ dɛn mi
		(0.90)
06	Chris	uh
		ʌː
07	Chris	[name]
		(0.78)
08	Chris	we make it look **so** much better
		wi meɪk ɪʔ lʊk soː mʌtʃ bɛʔə

6.7 Discussion

The chapter "List of MLE features" showed that MLE features are used by the adolescents in this study. The descriptive analysis in in Section 6.4.1 confirmed that at the level of group averages, these adolescents also show a vowel system similar to that described for Hackney/inner-London adolescents by Kerswill et al. (2008). The statistical analysis even showed that some language-internal constraints that had been found in studies of MLE in East London also appeared in the speech of Ealing adolescents, namely the constraint that a preceding or following nasal predicts a more open FACE onset (see Section 6.5.1.1; cf. Gates, 2018).

The statistical analysis also showed that different kinds of variation are apparent in the different diphthongs. FACE does not appear to vary with any of the social factors, either in terms of the height of the onset, or in terms of how monophthongal it is. For GOAT and PRICE, social factors are good predictors of variation in the frontness of the onset, but the CofPs do not differ from one another in monophthongisation of these vowels. The analysis of PRICE furthermore suggested that even when the social factor of outcode does not result in a main effect, it may interact with language-internal factors. The analysis of PRICE TL found interactions of outcode and language-internal constraints within both CofPs, meaning that even when adolescents from different parts of London do not differ in their surface realisations of PRICE, they differ in the constraints on that variation, i.e.

in the underlying variable system.

The discussion that follows is organised according to key themes and questions that arise from the analysis: the differences that were found between the two CofPs; the leaders of GOAT-backing in the Studio CofP – Kai, Denzel and Tony; the evidence for both GOAT-backing and GOAT-fronting in the results; the sex differences that emerged in the analyses of PRICE and GOAT; and the lack of an effect of social factors on FACE.

6.7.1 Differences between the CofPs

Differences between the two CofPs were found for the following variables: the onset F2 of PRICE in *like* (but not in other PRICE words); and in the onset F2 of GOAT. At the same time, there was evidence not quite reaching the criteria for marginal evidence for differences between the two CofPs in the TL of FACE and the onset F2 of the other PRICE words. For each of these, it was found that the Studio CofP were ahead of the Youthclub CofP with respect to the MLE changes – i.e. members of the Studio CofP were statistically likely to have a shorter FACE TL, more front PRICE onset, and more back GOAT onset, compared to the Youthclub CofP.

The differences between the two CofPs can be partly explained by the presence of both sexes within the Youthclub CofP – this was already suggested in Section 6.4.1 (see Figures 6.1 and 6.2). In fact, the variables that show evidence for a difference between the CofPs are also those that show a sex difference within the Youthclub CofP: GOAT F2, and PRICE F2. Recall from 6.4.1 that when the girls' data were removed and only the boys' data were plotted, differences between the two CofPs were minimal.

It is also possible that these differences are due to age. The members of the Studio CofP are mostly older than the members of the Youthclub CofP: the oldest members of the Youthclub CofP are aged 17, while this is the age of the youngest member of the Studio CofP. We cannot rule out the possibility that as the boys from the Youthclub CofP grow older, their vowel pronunciations will grow more similar to those of the Studio CofP.

This also relates to the issue of style-shifting, and the repertoires of diphthong variants that each participant commands. As mentioned briefly in Section 3.7.2, the members of the Studio CofP were more mature, and some seemed more confident in their interviews, and more willing to steer the conversation. Accordingly, it may be that they were less likely to accommodate their speech in any way to that of the interviewer, and hence for some variables, the Studio CofP members showed more MLE-like realisations in their interviews than did the Youthclub CofP.

With these caveats in mind, differences between the two CofPs are discussed further in the sections below.

6.7.2 GOAT-backing in the Studio CofP

The evidence from GOAT suggests that GOAT-backing is led particularly by the boys from outcode 3 (i.e. not from Northwest London) within the Studio CofP. It is tempting to assume that these three – Kai, Denzel and Tony – brought this sound change to the Ealing fieldsite, and that it diffused from them to the other members of the Studio CofP, and then to the boys from the Youthclub CofP. However, this explanation misses the fact that Kai, Denzel and Tony each come from a different part of London themselves: if geographical diffusion were the determining force, Kai, Tony and Denzel might each show a slightly different pattern of vowel variation. Therefore, the fact that Kai, Tony and Denzel all show a more back GOAT than other members of the Studio CofP must come from something else they have in common.

A better explanation lies in Kai, Tony and Denzel's commitment to their music. Whereas the other members of the Studio CofP could walk to the youth centre from their homes, Kai, Tony and Denzel lived far enough away that they had to spend considerable time and money on taking the tube or the overground to reach the youth centre (see Chapter 3 for more detail). Many of the members of Studio CofP who were from outcode 1 reported that they first started attending the youth centre for social reasons, and then later began using the Studio. By contrast, Kai Denzel and Tony reported that they came to the youth centre to use the studio, having heard about it from connections at college.

A body of research has posited links between the use of contemporary urban vernaculars and involvement in rap, hip-hop or grime. For example, Drummond (2018a) found that among a group of Manchester young people, rapping – i.e. whether or not they would rap during their everyday interactions – was a statistically significant predictor of rates of TH-stopping. P. Pichler and Williams (2016) state that it can be difficult to separate "the kind of multi-ethnolect spoken by young Londoners" from Hip-Hop Speech Style (HHSS; Cutler, 2003), a more global register. In their study, phonetic features and vocabulary from UK youth language are part of a "tool kit" used to index membership in hip-hop culture (P. Pichler & Williams, 2016, p.571), and some features, such as DH-stopping, are part of both MLE and HHSS. Brunstad, Røyneland, and Opsahl (2010) suggest that affiliation with hip-hop culture is partly responsible for the emergence of multiethnolects. In this respect, it makes sense that affiliation with hip-hop would predict use of MLE features.

This raises the question, why do Kai, Tony and Denzel differ from the rest of the Studio CofP in GOAT-backing, but are not in the lead in FACE-raising or PRICE-fronting? It may be the case that GOAT-backing, more than these other two MLE features, is indexically associated with HHSS. Cardoso et al. (2019) found that a speaker who used higher rates of GOAT-backing, DH-stopping and /k/-backing was rated as less educated, and was more likely to be identified as Black, than a speaker who also used MLE features but showed lower rates for the three phonetic features just mentioned. This suggests that GOAT-backing, DH-stopping and /k/-backing are perceived differently from other features such as GOOSE-fronting, TH-fronting and /l/-vocalisation. Cardoso et al. (2019) note that features such as GOOSE-fronting, TH-fronting and /l/-vocalisation are found in several vernacular varieties of British English, while GOAT-backing, DH-stopping and /k/-backing are more specific to MLE. It may be that backed GOAT is indexically linked with the persona of a typical MLE speaker – someone male, Black, and from an underprivileged background – to a greater extent than fronted PRICE or raised FACE.

We can tentatively speculate on the sound symbolism of GOAT-backing. Ohala's (1994) frequency code posits that high front vowels are associated with smallness, and secondarily with submissiveness and cooperativeness, while back vowels are associated with largeness, and secondarily with dominance (see also Eckert, 2011). Podesva et al. (2015) found links between GOAT-fronting and perceived positive affect. Relatedly, Pratt (2018) suggests that speech sounds that involve retraction of the tongue dorsum can be drawn on to index toughness. Hence it may be the case that GOAT-backing is able to index toughness sound symbolically, while PRICE-fronting does not have this affordance.

6.7.3 Two directions of change? GOAT-backing and -fronting

There seem to be contrary pulls within the GOAT diphthong. On the one hand, Kerswill et al. (2008) and Cheshire et al. (2011) identify backing and raising of GOAT as being particularly associated with: (a) inner-London speech rather than outer-London and the southeast; (b) Non-Anglo ethnicity; (c) male adolescents and young adults rather than females. GOAT-backing is also one of the features identified by Cardoso et al. (2019) as a potentially salient indicator of MLE that leads to negative perceptions by outsiders, as we saw in the section above. MLE is therefore supposed to be characterised by an avoidance of and resistance to GOAT-fronting.

At the same time, fronting of at least the offset of GOAT is a widespread change supposed to be taking place in the southeast of England (A. Williams & Kerswill, 1999; Kerswill & Williams, 2000a). GOAT-fronting is actually an expected feature of MLE, given that MLE involves fronting of GOOSE: it is known that the fronting of GOOSE and GOAT are structurally linked and that the latter tends to follow the former (Labov, 1994, p.108).

As mentioned in Section 6.6.2, a wide range of GOAT pronunciations can be heard in the data, including fronted monophthongal realisations. Importantly, the social factors that predict GOAT-backing do not correlate with Trajectory Length – for example, the girls are found to have a more front onset of GOAT compared to the boys, but do not appear to differ from the boys in TL. This means that there is not a binary choice between a backed, monophthongal GOAT and central, diphthongal variants, but that there are potentially also central–front monophthongal variants in the mix. While there is inter- and intraspeaker variation in participation in GOAT-backing, at the same time, participants also seem to be variably participating in GOAT-fronting. This finding is worth noting because it potentially indicates that the kind of GOAT-fronting found in the current data is different from that attested elsewhere in England: studies of GOAT-fronting in two different regions of the UK (Milton Keynes, and Yorkshire) have found that the offset tends to be fronted while the nucleus remains back or central (Kerswill & Williams, 2000a; Haddican et al., 2013).

It is also known that coronals play a role in the fronting of GOAT and GOOSE (Hall-Lew, 2009, p.163; Cheshire et al., 2011, p.171), such that the most fronted realisations are found after a coronal. This was also found in the current data: in both CofPs, it was found that a preceding coronal favoured a more front onset to GOAT. Unfortunately, the current study did not include an analysis of GOOSE-fronting. Further analysis is needed of the fronting of GOOSE and GOAT, and whether they are both favoured by a preceding coronal in Hackney, Ealing and other parts of London.

6.7.4 Differences between *like* and other PRICE words

The findings for PRICE in *like* did not always match up with the findings from the other PRICE words.

Similarities in the results of PRICE and *like* were as follows. In *like*, there was strong evidence that the Studio CofP had a higher F2 (more front onset) compared to the Youth-club CofP. There was also evidence, not quite reaching the threshold for marginal evidence, that the Studio CofP had a higher F2 in other PRICE words. In both *like* and in other PRICE words, there was evidence – again not quite reaching the criteria for marginal evidence – that within the Youthclub CofP, boys have a more front onset compared to the girls. There was no indication of a difference between the two CofPs in *like* or in other PRICE words.

However, there was strong evidence within the Youthclub CofP that the girls had a more diphthongal pronunciation of PRICE compared to the boys; but this effect was not found in *like*. This suggests that the boys are not pronouncing *like* as noticeably more monophthongal than other PRICE words, but that the girls are pronouncing *like* as more monophthongal than other PRICE words. Why should this be the case?

At one level, we would expect *like* to be more monophthongal than other words, simply because it is so frequent. It is known that more frequent words are more prone to vowel reduction (Aylett & Turk, 2006), which in the case of a diphthong would lead to monophthongisation. Therefore, it is to be expected that the girls have a more monophthongal pronunciation of *like* compared to other PRICE words; what this result highlights is that the boys have relatively monophthongal PRICE across all lexemes. Girls have diphthongal PRICE in most words, and a relatively monophthongal pronunciation in *like*; the boys have a realisation of PRICE that is similarly monophthongal to the girls' *like* tokens across the board.

6.7.5 Sex differences

Main effects of sex were found for the following variables: the Trajectory Length of PRICE; and the onset F2 of GOAT. In addition, there was evidence not quite reaching the criteria for marginal evidence for main effects of sex on PRICE F2, both in *like* and in other PRICE words. In each of these cases, females appeared more conservative than males with respect to the MLE tendencies, i.e. GOAT-backing, and PRICE-fronting and -monophthongisation (summarised above in Table 6.1). This finding aligns with previous studies that have found differences between the sexes in the use of MLE.

Various studies have suggested that there is a gendered aspect of MLE/MUBE use. For example, Drummond (2018b, p.196) posits a scale of "urban/street-style" orientation that characterises the adolescent participants in his study of youth language in Manchester. Drummond (2018b) presents two fictional characters who represent opposite ends of the continuum: at one end is a boy (Malachi) who is into grime and uses MLE/MUBE features as well as regional accent features; at the other end is a girl (Chantelle) who listens to house music and RnB and speaks with a Mancunian accent. He states that the Chantelle end of the continuum could just as easily have been male, but that the Malachi end is "definitely more male than female". Relatedly, we have already seen that P. Pichler and Williams (2016) see HHSS and MLE as linked, and both as linked to masculinity.

Studies of MLE have also found interactions of gender and ethnicity. Gates (2018) reports "stark gender differentiation" in FACE and PRICE in a cohort of secondary school students in Newham, East London: the boys in her study showed a more front PRICE and raised FACE compared to the girls. Gates (2018) suggests that gender, ethnic, and peer group identities intersect, such that the White girls (who name themselves the White Squad) use distinctly more conservative diphthong variants in comparison to the ethnic minority girl peer groups. Similarly, S. Fox (2015) found that, in an East London youth centre, Bangladeshi boys were leading in the innovation of MLE-like FACE and PRICE vowels, and that the innovative variants were diffusing through friendship networks to White British boys; but the White British girls did not seem to be adopting the innovative variants to the same extent, and favoured more conservative variants.

Cheshire et al. (2011), too, found complex ethnicity- and gender-based differences in the adoption of MLE vowels among Hackney 16–19: Anglo girls led in the fronting of FOOT; Anglo girls were more conservative than either Anglo boys or Non-Anglo girls and boys in the raising of FACE; and Non-Anglo boys had the most back GOAT, while

Anglo boys and Anglo and Non-Anglo girls were all similar to one another with respect to GOAT. Gender differences were not apparent among the 8–9 year olds, but in the 12–13 year old age group, the Anglo girls began to show more conservative realisations of FACE and PRICE. On this basis, they state “We conclude from this that there is some gender differentiation in the MLE vowel system, which increases as children pass into adolescence” (Cheshire et al., 2011, p.170).

In cases of stable sociolinguistic variation and changes nearing completion, men use higher rates of the non-standard variant, while in new changes in progress, women lead in the use of the incoming variant (Cheshire, 2002). We could see the findings here as (a) indicating that the changes of GOAT-backing and PRICE-fronting are nearing completion, and (b) confirming the trend that women are more conservative than men. However, this would be a premature conclusion to draw – not least because the data presented here include only one age group, so we do not actually have any evidence of change in progress.

We also need to take into account the possibility of an additional change: GOAT-fronting in Ealing – either diffusing from elsewhere in outer London and the southeast, or motivated by GOOSE-fronting (see Section 6.7.3 above). In the current data, the girls seem to be leading in GOAT-fronting – although this is a generalisation, and boys, too, use fronted variants, and indeed some boys favour fronted variants. If the girls are leading in GOAT-fronting, and if GOAT-fronting is a supralocal change, this supports the suggestion of J. Milroy, Milroy, Hartley, and Walshaw (1994) that females are more likely than males to lead in the adoption of supra-local changes.

6.7.6 Variation in FACE

The quantitative analysis revealed remarkably little influence of social factors on FACE variation, and the descriptive analysis in Section 6.4.1 suggested a low degree of inter-speaker variation in FACE production. This contrasts with the findings of Gates (2018), who found that boys tended to have a more raised FACE onset than girls, and that in particular, the “White Squad” had a lower FACE onset compared to the other groups of girls.

At the same time, FACE showed strong evidence for language-internal constraints, and these were consistent across and within the two CofPs. One of these – the finding that a preceding or following nasal favours a higher F1 i.e. a more open onset to FACE – is also consistent with Gates’ findings in East London (Gates, 2018). This provides some indication that the same language-internal constraints operate on FACE-raising in both East London and the Ealing fieldsite. Alternatively, this may simply be a coarticulation effect: it is known that nasalised vowels may show a higher F1 compared to their non-nasalised counterparts (Carignan, Shosted, Shih, & Rong, 2011; Arai, 2005).

6.8 Chapter summary

This chapter analysed variation in the onset qualities and diphthong dynamics in FACE, PRICE and GOAT among the adolescent participants. The descriptive overview suggested that overall, the adolescents show a vowel system that is extremely similar to the emerging MLE one described by Kerswill et al. (2008). It was found that the Studio CofP tend to have a more front PRICE onset (especially in *like*) and more back GOAT onset compared to the Youthclub CofP. Within the Youthclub CofP, however, the boys have more front and monophthongal PRICE and more back GOAT compared to the girls. In sum, boys seem to have more MLE-like realisations of PRICE and GOAT than the girls.

The Studio CofP has more MLE-like realisations of the diphthongs GOAT and PRICE than the Youthclub CofP – though it was suggested that this may not be the case if it were not for the contributions of the girls within the Youthclub CofP – and within the Studio CofP, the boys who travel from further away to use the studio show the most backed realisation of GOAT. This chapter has suggested that GOAT, more than PRICE, is indexically linked to HHSS and that possibly the sound symbolic associations of GOAT-backing make it more useful than PRICE in indexing a stance of toughness and masculinity.

At the same time, GOAT-fronting also seems to be in progress in this community. As has been found with studies of GOOSE- and GOAT-fronting, a preceding coronal favours a fronted realisation of the vowel. Also in line with other studies, girls lead in GOAT-fronting.

Variation in FACE did not correlate with any of the social factors, but did show some of the same language-internal conditioning found by Gates (2018): a preceding or following nasal favours a more open onset.

Chapter 7

Diphthong variation in the speech of the children

7.1 Background

7.1.1 Children's acquisition of MLE diphthongs

The starting point of this chapter is the claim of Cheshire et al. (2011) that even the youngest age group in the MLE project, the 4–5 year olds, had acquired the same vowel system as the adolescents. This vowel system, as we have seen, is purported to involve (a) reversal of Diphthong Shift and (b) monophthongisation of the diphthongs (see review in Chapter 5). The 4–5 year olds' vowels from the MLE project are shown in Figure 7.1. The diphthongs of interest for the current project are shown in Figure 7.1 (b): FACE is close, similar in height to DRESS and also GOAT; GOAT itself is backed; PRICE is similar in frontness and openness to TRAP and MOUTH.

Cheshire et al. (2011) distinguish different apparent-time patterns for endogenous and exogenous variables. Reversal of Diphthong Shift and monophthongisation of the diphthongs are considered an endogenous change, due to the high-contact scenario, and these show a flat age distribution, with the system already in place by age 4–5 (Cheshire et al., 2011, pp.171–172; Kerswill et al., 2013, p.272). The fronting of GOOSE, meanwhile, is a globally diffusing feature, and this appears to show incrementation, with an adolescent peak in the 16–19 year old group.

To examine the relationship between caregiver input and the children's production of these variables, Cheshire et al. (2011) examined the correlation between children's vowel productions (the children were aged 4–5, 8–9 and 12–13) and those of their principal caregivers for two variables: the F1 of FACE and the F2 of GOOSE. They found no significant correlation between the children's and caregivers' F1 in FACE, also noting that the 4–5 year olds tended to have a lower F1 compared to the other age groups. For GOOSE F2, the caregivers showed a bimodal distribution: some had a fronted GOOSE vowel, while others – especially those who were not born in the UK – had a back GOOSE. All of the children bar two, however, had a fronted GOOSE vowel. Cheshire et al. (2011) contrast this finding with the findings regarding GOAT-fronting in Milton Keynes, where 4-year-olds tended to copy their principal caregiver's GOAT pronunciation, but 8- and 12-year-olds were participating in the regional change of GOAT-fronting regardless of whether their caregiver had fronted GOAT or not. Cheshire et al. (2011, p.169) suggest "the children's rejection of the back GOOSE vowel may well represent a lessening of the linguistic tie between parent and child in comparison to majority-community language speaking families".

One final finding that is relevant to the current chapter is Cheshire et al.'s apparent-time findings suggested "increasing gender differentiation with age" in the MLE vowels (Cheshire et al., 2011, p.172). In the 4–5-year-old age group, they found no significant effect of gender on the vowels, while the 8-year-old group showed no gender effects for FACE or GOAT, but did show gender and ethnicity differences in FOOT. In the 12–13-year-old group, they found that Anglo girls were more conservative than their peers with respect to FACE and PRICE; while GOAT and FOOT appeared to show an interaction of ethnicity and gender, as Non-Anglo boys showed backer realisations of these two vowels compared to girls and Anglo boys (Cheshire et al., 2011, p.170). These gender and ethnicity differences were also found among the 16–19-year olds: in this age group, Anglo girls led in fronting of FOOT; Anglo boys and girls tended to have a more open FACE than Non-Anglos; and Non-Anglo boys had a markedly backer GOAT onset than the other groups. These findings from the MLE project will become relevant to the discussion later on in the chapter.

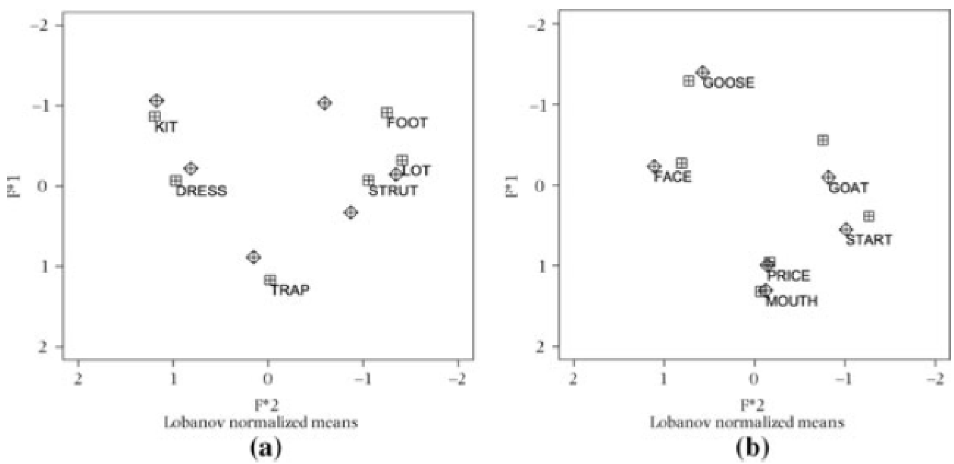


Figure 7.1: From Cheshire et al. (2011, p.165): MLE project 4–5 year olds' short monophthongs (a) and diphthongs plus GOOSE and START (b)

Although Cheshire et al. (2011) and Kerswill et al. (2013) claim that the children have acquired the same vowel system as the adolescents, this claim is based on comparison of diphthong onsets, and not on diphthong dynamics. This chapter compares the diphthong productions of both adolescents and children in Ealing both in terms of the onset qualities, and in terms of diphthong dynamics, i.e. how diphthongal or monophthongal FACE, PRICE and GOAT are. It turns out that the 5–7 year old Ealing children have acquired the same diphthong system as the adolescents in terms of onset qualities, but that they behave quite differently in terms of monophthongisation of the diphthongs. This chapter also replicates Cheshire et al.'s finding that there is no evidence of gender differentiation in the diphthongs among young children.

7.1.2 Influence of ambient languages and accents

We are unlikely to see direct evidence in the children's output of any single input – for example, although one child speaks Tamil at home (see Section 7.2.1 below), this is not particularly numerous in the Feature Pool, and unlikely to be in evidence in his speech to his peers and even less likely to be evident in their speech. Cheshire et al. (2011) for this reason describe the situation as one of group second language acquisition (Winford, 2003). The children have teachers who speak with a Cockney accent, and while this means that Cockney is a significant part of their English input, the children are as likely to converge to one another as they are to their teacher.

There are two languages that are numerous in the Feature Pool both in that several of the children in this sample receive input in this language, and in that large proportions of the local Ealing population speak these languages – and therefore these languages are worth considering as having an influence on the output of the Feature Pool, i.e. the children's speech production. These languages are Somali and Arabic. We assume that the children with caregivers who were born in Arabic-speaking or Somali-speaking countries receive input both in that language, and in L2-accented English. Therefore in this section, we review what those kinds of input might look like, and particularly how they might influence the children's production of the diphthongs FACE, PRICE and GOAT.

There is relatively little literature on Somali-L1 speakers' vowel productions in English. This is most likely because Somali has a relatively rich vowel system, and so acquisition of the English vowel contrasts is unlikely to present a problem for Somali learners of English (Conway, 2008, p.27). Standard Somali has a system of ten vowels, arranged as five back-front contrastive pairs: /ɪ– i/, /ɛ– e/, /ɑ– æ/, /ɔ– o/ and /u – ʉ/ (Saeed, 1999). These also have long and short alternate versions, and the length contrast is phonemic rather than allophonic. Finally, these vowels can combine into diphthongs, which can also be seen as forming five pairs, again contrasting on the front–back dimension (Saeed, 1999). Koffi (2012) investigates Somali L1 speakers' production of the

English lax vowels. Although the monophthong inventories of Somali and English are similar, Koffi (2012) finds that the back vowels are fronted by the Somali speakers, and hence that the English vowel space of the Somali bilinguals is smaller compared to that of the American English monolinguals. All of the diphthongs in the present study have close equivalents in Somali: FACE has potential equivalents in /ei/ and /ɛɪ/; PRICE corresponds to /æi/ and /ɑɪ/ – although the former is somewhat more front and the latter somewhat more back than SSBE PRICE, and neither corresponds to Cockney PRICE; GOAT has potential equivalents in /oː/ and /öi/ (Saeed, 1999). Börjesson (2014) lists the contrasts /i – e/, /ae – aː/ and the FOOT vowel as areas of difficulty for Somali learners of English, but the diphthongs are not mentioned as liable to phonetic transfer. In sum, we would not expect any innovations in the diphthongs to come from Somali.

Various varieties of Arabic are also present in the Feature Pool: while a number of the children in this sample had caregivers who are speakers of Kuwaiti or Yemeni Arabic, we should also take into account the presence locally of Syrian, Lebanese and Jordanian Arabic. Khattab (2013) states that L1-accented patterns for Lebanese Arabic-English bilinguals would produce [eɪ] or [eː] for FACE, and [oː] or [əʊ] for GOAT – in other words, the onset would be similar to the SSBE pronunciation of these vowels, but with the possibility of monophthongisation. Evans and Alshangiti (2018) investigated the intelligibility of L2 English by high and low proficiency speakers whose L1 was Saudi Arabic. SSBE-L1 listeners correctly identified the high proficiency speakers' productions of PRICE 97% of the time. They correctly identified the high proficiency speakers' productions of FACE 61% of the time, and incorrectly identified it as PRICE 33% of the time. GOAT was correctly identified only 34% of the time: the high proficiency speakers' productions of GOAT were incorrectly identified as GOOSE, FOOT and CHOICE. Munro (1993) compares the vowel productions of L1 speakers of a variety of Arabic dialects with L1 English speakers. PRICE was unfortunately not part of that study, but FACE and GOAT were: the Arabic-L1 speakers did not differ significantly from the L1-English speakers in the F1 or F2 at 30% of FACE, but they did differ significantly in the F2 at 30% of GOAT. In Munro (1993), the Arabic group was also found to differ significantly from the English group in the amount of formant movement in these two diphthongs – the Arabic-L1 speakers produced more monophthongal tokens of FACE and GOAT than did the English L1 speakers. Taken together, these studies suggest that the presence of Arabic in the local Feature Pool supports backed, monophthongal tokens of GOAT, and potentially more open and/or monophthongal variants of FACE.

7.2 Methods

7.2.1 Participants

19 children were originally recorded and of these, 14 were selected for the phonetic analysis. The other 5 children were excluded because they were later discovered to have a known speech/hearing impairment, or changed school before the second Diapix recording took place, or the recording was not of sufficient quality. All 14 remaining children successfully completed a phonology assessment screen (Dodd et al., 2002) and were reported by their teachers to have no known speech/hearing issues.

These children comprise 3 girls from school year 1 (ages 5–6), 3 boys from year 1, 4 girls and 4 boys each from year 2 (ages 6–7). All were born in London or the UK except for F3, who was born in Brazil, and M4, for whom this information was not available. There are two sets of siblings in the data: F1 is the younger sister of M5, and M1 is the younger brother of M6. Table 7.1 is a condensed version of the information given earlier in Tables 3.5 and 3.6. Of the children whose caregivers provided information about their language background, F3's home language environment is reported to be monolingual Portuguese, M3's is monolingual Somali and F9's is monolingual Arabic. F9's caregivers adopted an explicit "Arabic in the home" policy. F1 and M5 are reported to receive input in both Somali and English from one caregiver, and English from the other; both respond in English. F4 receives input in Somali and English from one caregiver, and just Somali from caregiver 2; she replies in English to caregiver 1, and in both Somali and English to caregiver 2. Of F10, it is reported that both caregivers address her in both Swahili and English, and likewise she communicates to them in both Swahili and English. M8 has a similar pattern of bilingualism but in his home, the languages used are Tamil and English. Finally, M7's caregivers, like F9's had an explicit language policy, and they would address their children only in Arabic, but the children replied in either Arabic or English. M7 attended Arabic Saturday school.

For the comparison between children and adolescents, the adolescents who were from outcode 3 i.e. not from Northwest London were excluded. The adolescents from outcode 1 (same postcode as youth centre and primary school) lived locally and the adolescents from outcode 2 (Ealing/neighbouring boroughs) attended the local secondary school with the outcode 1 adolescents.

For a more detailed description of these children and the adolescents, see Chapter 3, and see Chapter 2 for details on how speech data were recorded from the two age groups – sociolinguistic interviews for the adolescents, vs. the Diapix task for the children (Baker & Hazan, 2011). Chapter 5 gives full details on the methodology for acoustic analysis.

Table 7.1: Summary of background information on the child participants. Empty cells indicate that this information was not provided by the caregivers

Child	Place of birth	School year	Sibling order	Caregiver place(s) of birth	Caregiver 1 speaks to child	Child speaks to caregiver 1	Caregiver 2 speaks to child	Child speaks to caregiver 1
F1	London	1	middle	Mogadishu, Somalia	Somali, English	English	English	English
F3	Brazil	1		Brazil	Portuguese	Portuguese	Portuguese	Portuguese
F4	London	1	youngest	Mogadishu, Somalia; Yemen	English, Somali	English	Somali	English, Somali
F7	London	2		Yemen	Arabic, English	Arabic, English		
F8	London	2						
F9	UK	2	middle	Kuwait	Arabic	Arabic	Arabic	Arabic
F10	London	2	oldest	Kenya	English, Swahili	English, Swahili	English, Swahili	English, Swahili
M1	London	1	middle	Mogadishu, Somalia				
M3	London	1		Somalia	Somali	Somali	Somali	Somali
M4		1						
M5	London	2	oldest	Mogadishu, Somalia	Somali, English	English	English	English

M6	London	2	oldest	Mogadishu, Somalia				
M7	London	2	oldest	Kuwait; Iraq	Arabic	Arabic, English	Arabic	Arabic, English
M8	UK	2	oldest	Sri Lanka	Tamil, English	Tamil, English	Tamil, English	Tamil, English

7.2.2 Checking for effects of addressee and keyword in the child data

A preliminary analysis of the children's data was conducted to check for effects of addressee and keyword.

Controlling for addressee seemed necessary, as children are known to be sensitive to the communicative expectations of their interlocutor (e.g. E. S. Andersen, 1990; Lanza, 1992; Khattab, 2007; Matras, 2009). The data was coded for the addressee factor by turn at talk, according to whether that utterance seemed to be directed at the other child in the recording session, or at the adult interviewer. If utterances were ambiguous or seemed addressed to both, they were categorised as "interviewer"; if it seemed that the child could be talking to him/herself (e.g. while counting how many differences had been found in the Diapix task), this was categorised as "child". In practice there were very few self-addressed tokens and so these were excluded from the statistical analysis presented in Section 7.4.

Keyword was included as a factor because early exploration of the data suggested that the children showed an enlarged vowel space in the Diapix keywords. This is to be expected, as keywords tend to be prosodically prominent. This variable was coded according to whether the word in which the vowel occurs was integral to the Diapix task or not. The Diapix keywords were *cake, gate, table, baby; kite, tiger, spider, five; sheep, cheese; sock, dog; cat, hat; football, book* (see Chapter 2). However, also classed as "keyword" in this binary variable are any words that identify something potentially distinctive about the Diapix scene. For example, under this definition, in the sentences "Is it on the **bush**?", "Is it **black**?", the words in bold text are keywords because they are distinguishing features of the scene.

The preliminary analysis found that in the child data, addressee and keyword had no effect on diphthong onset quality nor on diphthong dynamics for any of the three diphthongs.

7.2.3 Independent predictors, model structure, exclusions

The independent predictors included in the models presented in Section 7.4 were: log(duration); age (adolescent or child); speaker sex (F or M); preceding environment; and syllable coda type.

Duration was log-transformed and included in the models in order to control for this factor – we would expect that tokens with a longer duration would also be more diphthongal.

As in the previous chapter, preceding environment was categorised according to a combination of place and manner. The following phonological environment variable was categorised by syllable coda type in the first instance: i.e. whether the vowel occured word-finally in an open syllable; word-medially in an open syllable; or in a closed syllable. Within the closed syllable category, the different types of coda were further categorised according to whether the consonant in the syllable coda was a voiced obstruent, a voiceless obstruent, or a nasal (tokens with a coda approximant or glide had already been excluded). For convenience, these factor levels are summarised in Table 7.2.

For this analysis, 2431 tokens of FACE (adolescents = 1590, children = 841), 1918 tokens of PRICE (adolescents = 1076, children = 842), 983 tokens of PRICE in *like* (adolescents = 866, children = 117) and 2488 tokens of GOAT (adolescents = 1546, children = 942) were included. A full breakdown of the no. of tokens of each variable by age, sex and preceding and following phonological environment is given in Appendix B. As before, tokens were included where the duration of the vowel segment was no less than 50ms and where the diphthong was not reduced. Tokens were also only included where there was not substantial speaker overlap or background noise.

Table 7.2: Language-internal independent factors: summary of categories used in the analysis

Preceding segment	Approximant / glide	The onset consonant is an approximant /ɹ w/, glide /j/ or lateral /l/
	Coronal	The onset consonant is a coronal obstruent, either voiced /d ð z ʒ/, or voiceless /t θ s ʃ/
	Labial	The onset consonant is a bilabial or labiodental obstruent, /p b f v/
	Nasal	The onset consonant is a nasal, /m n ŋ/
	Other	
	Velar	The onset consonant is a velar obstruent, /k g/
log(duration)		
Coda type	Coda nasal	Syllable coda is an underlying nasal e.g. *don't, time, game*
	Final open	The vowel is word-final and the syllable has no coda, e.g. *play, so, tie*
	Medial open	The vowel is not word-final but the syllable is open, e.g. in *tiger* [taɪ.gə], *socialise* [so.ʃə.laɪz], *station* [steɪ.ʃən]

Voiced	The syllable coda is a voiced obstruent
Voiceless	The syllable coda is a voiceless obstruent

7.2.4 Statistical analysis and reporting of results

The statistical analysis largely followed the methods presented in the previous chapter, i.e. Bayesian generalized mixed effects models. These were fit using the brms package in R (Buerkner, 2017; R Core Team, 2020). The priors used were the same as in the previous chapter, i.e. weakly informative priors.

For the relevant models, onset formant frequency was scaled and centred before being entered into the model, and Trajectory Length was log-transformed and then scaled and centred before being entered into the models. Duration, too, was log-transformed and scaled and centred. All categorical variables were sum-coded before being entered into the models. This means that for each predictor variable, the intercept represents the mean of the factor levels, and the coefficient for each factor level represents the difference between that factor level and the mean of all factor levels. The contrast matrices for the sum-coded predictor variables can be found in Appendix B.2.

As before, in the reporting of results, the following are presented:

- The estimated regression coefficient ($\hat{\beta}$), i.e. the median of the posterior distribution
- 95% Highest Density Intervals (HDI) about these estimates
- Proportion of the posterior $>$ or $<$ 0. This is abbreviated to PD i.e. probability of direction.

In line with other studies, if the posterior probability of the effect direction is 95% or higher and the 95% HDI does not include 0, we will say that there is strong evidence for a non-zero effect (Tanner et al., 2019; Franke & Roettger, 2019). If one but not both of these conditions is true, we will say there is marginal evidence for the effect.

Model summary tables for this chapter can be found in Appendix B.3, and token numbers by vowel, age and preceding/following phonological environment are given in Appendix B.1.

7.3 Descriptive analysis of vowel plots

Figure 7.2 shows confidence ellipses for the point vowels FLEECE, TRAP, LOT and FOOT, and the diphthongs FACE, PRICE and GOAT at 20% of the vowel segment's duration – the onsets of these vowels. We can see that overall, the children's vowel system (top-right and bottom-right) appears similar to that of the adolescents (top-left and bottom-left), with some small differences. The children's ellipses are wider, as is to be expected – we know that children's speech production in general shows greater variability than adolescents' and adults' (Vorperian & Kent, 2007).

Both age groups and both sexes have a FACE onset that is located at approximately 1.0 on the F1 scale, and 1.4–1.5 on the F2 scale (cf. Table 7.3). For all groups, the onset of PRICE overlaps with TRAP, except for the adolescent females (top-left), who have the onset of PRICE slightly back compared to TRAP. The adolescent males (bottom-left) have a back GOAT, overlapping with LOT, while for the adolescent females, GOAT overlaps slightly with FACE but hardly overlaps with LOT. Among the children, neither group has GOAT as far back as do the adolescent males, and for both male and female children, GOAT overlaps with LOT – although for the males, it also overlaps slightly with FACE.

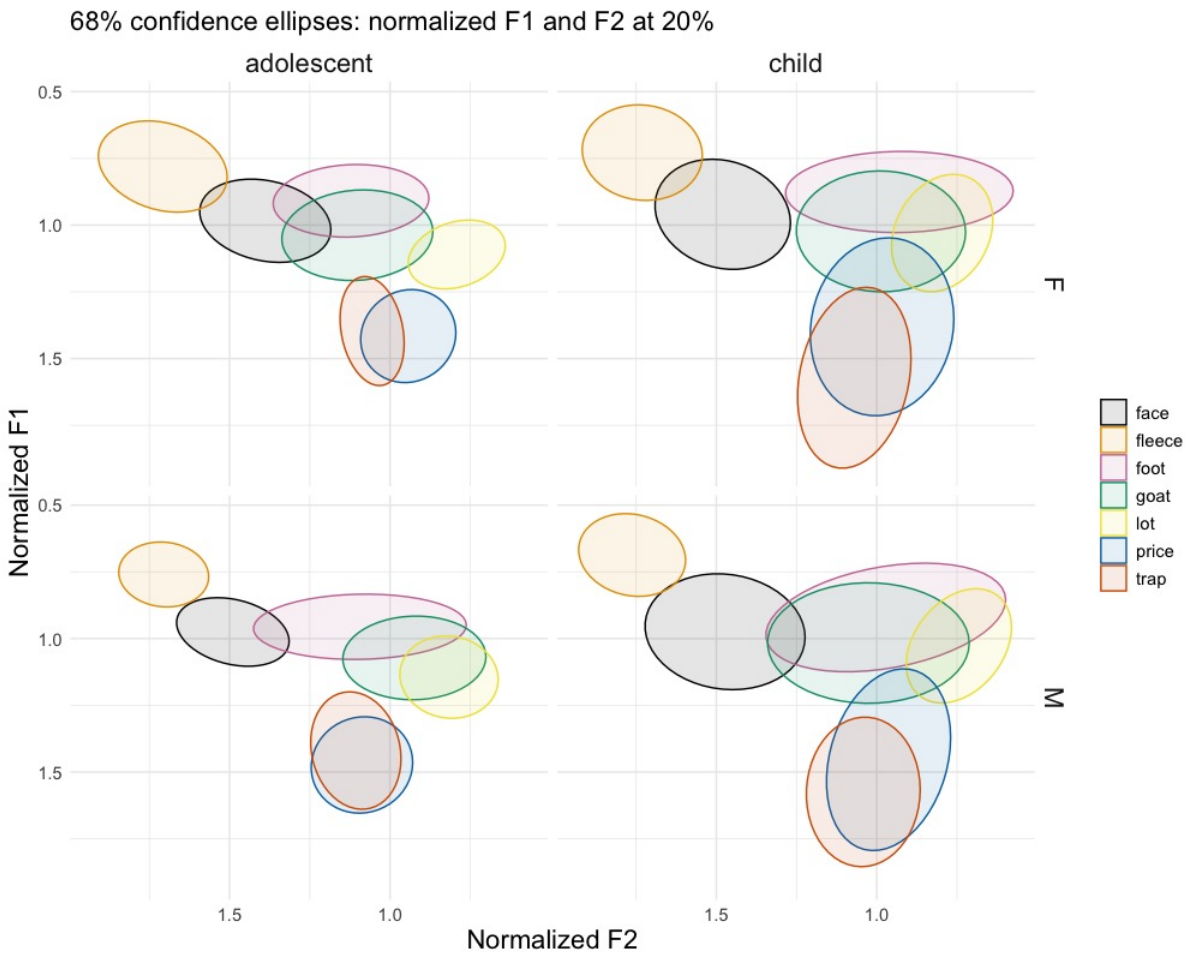


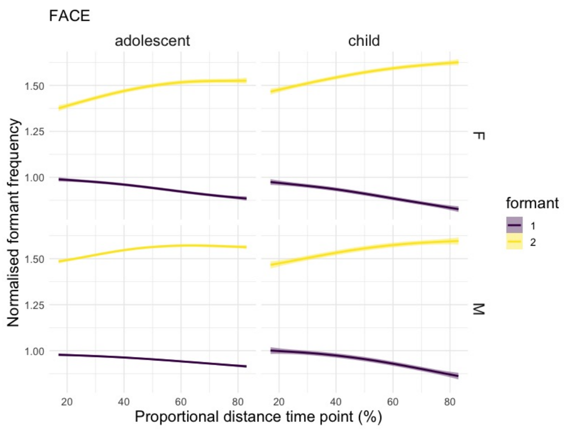
Figure 7.2: Vowel plots by age and sex showing 68% confidence ellipses for the vowels FACE, FLEECE, FOOT, GOAT, LOT, PRICE and TRAP from F1 and F2 measurements taken at the 20% time point (diphthong onset)

Figure 7.2 tells us only about diphthong onsets, and not about diphthong dynamics. To shed more light on the diphthong trajectories, Figure 7.3 shows smoothed F1 and F2 trajectories for the three diphthongs. These plots should be taken with a pinch of salt, as they treat the time scale as continuous, when in fact only five time points were sampled (20%, 35%, 50%, 65% and 80%). Nevertheless, a few useful observations can be drawn from these plots.

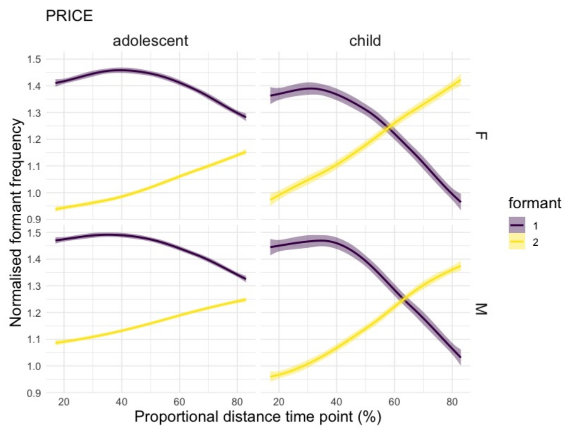
Firstly, the trajectory of FACE is very consistent between the two age and sex groups, and within these groups – as indicated by the narrow confidence intervals.

The plot for PRICE is especially revealing. The trajectory shapes for adolescents and children are similar, i.e. F1 increases initially and then rapidly decreases, while F2 shows a constant increase over time. But for the children, this pattern is exaggerated, leading to a crossover pattern: between the 50% and65% time points, the values of the two formants crossover and F2 finishes at a high value (i.e. front target), while F1 finishes low (i.e. close target). For the adolescents, the overall amount of change in each formant is not so great.

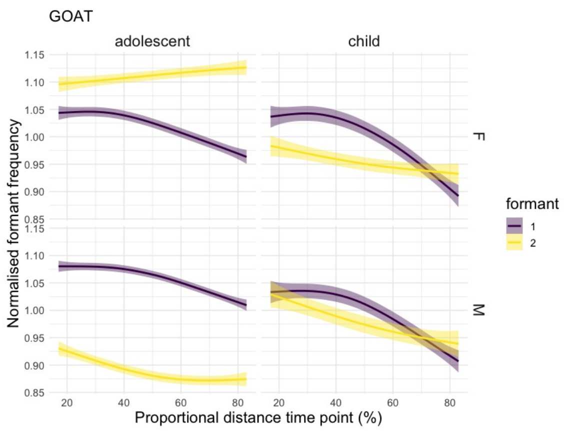
As to GOAT, for all groups, the F1 starting point is similar, between 1.00 and 1.10 (Table 7.3). However, while for children, the F2 starting point is around 1.00, for adolescent females it is 1.09, and for adolescent boys, it is 0.93. For adolescent males, and for children of both sexes, F2 tends on average to decrease over the duration of the vowel. However, for adolescent females, the F2 tends to increase slightly over time. For both ages and both sexes, F1 decreases over time, indicating that the end point of the vowel is closer than the onset.



(a) Smoothed trajectories for FACE



(b) Smoothed trajectories for PRICE



(c) Smoothed trajectories for GOAT

Figure 7.3: Smoothed F1 (dark purple) and F2 (yellow) trajectories for the diphthongs FACE, PRICE and GOAT across the five time points

To allow us to inspect the diphthong offglides more closely, Figure 7.4 is constructed in the same way as Figure 7.2, except that it shows the 80% time points, i.e. the diphthong glides. The most interesting difference between Figures 7.2 and 7.4 is in PRICE. Figure 7.4 shows that for adolescents, the offset of PRICE overlaps with TRAP, indicating that it is relatively monophthongal. By contrast, among the male children, the glide of PRICE just barely overlaps with TRAP, and for the female children, it does not overlap with TRAP at all. Instead, it overlaps with the offset of FACE, suggesting that for the children, PRICE and FACE have similar glide qualities – something like [ɪ].

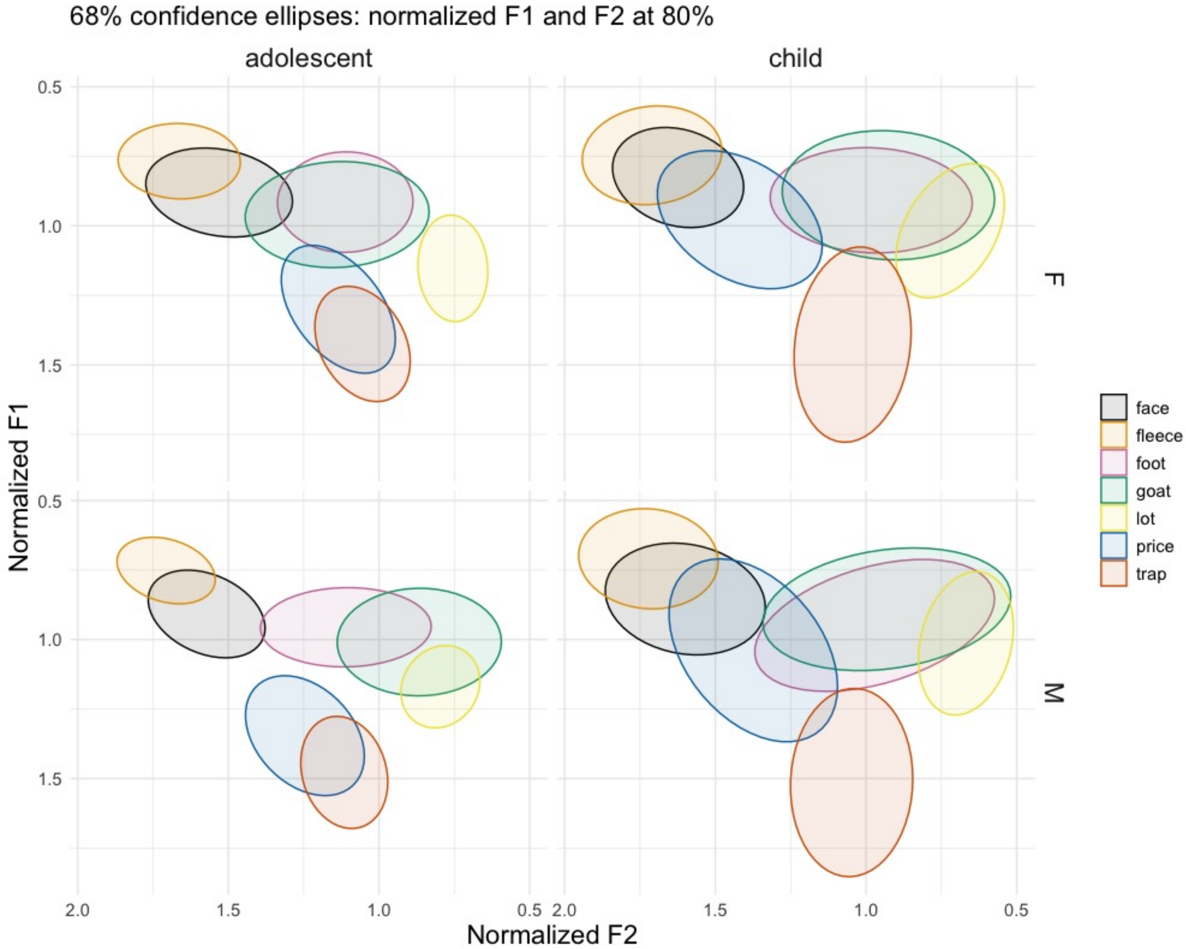


Figure 7.4: Vowel plots by age and sex showing 68% confidence ellipses for the vowels FACE, FLEECE, FOOT, GOAT, LOT, PRICE and TRAP from F1 and F2 measurements taken at the 80% time point (diphthong offglide)

Table 7.3: Means and standard deviations in normalized formant frequencies for FACE, FLEECE, FOOT, GOAT, LOT, PRICE and TRAP by age and sex

			F1 at 20%		F1 at 80%		F2 at 20%		F2 at 80%	
vowel	age	sex	mean	SD	mean	SD	mean	SD	mean	SD
FACE	adolescent	F	0.99	0.12	0.89	0.14	1.39	0.16	1.52	0.19
	adolescent	M	0.98	0.1	0.92	0.13	1.49	0.14	1.56	0.15
	child	F	0.97	0.16	0.83	0.14	1.47	0.16	1.62	0.17
	child	M	0.99	0.18	0.87	0.16	1.47	0.19	1.59	0.21
FLEECE	adolescent	F	0.79	0.13	0.78	0.11	1.71	0.15	1.66	0.16
	adolescent	M	0.77	0.1	0.76	0.1	1.71	0.13	1.71	0.15
	child	F	0.74	0.13	0.75	0.13	1.73	0.15	1.71	0.18
	child	M	0.7	0.12	0.72	0.14	1.76	0.13	1.71	0.2
FOOT	adolescent	F	0.91	0.11	0.92	0.13	1.12	0.19	1.12	0.17
	adolescent	M	0.95	0.09	0.95	0.12	1.09	0.23	1.09	0.22
	child	F	0.9	0.14	0.91	0.15	0.94	0.26	0.98	0.24
	child	M	0.92	0.16	0.95	0.17	0.97	0.27	0.97	0.28
GOAT	adolescent	F	1.04	0.13	0.97	0.15	1.09	0.18	1.13	0.23
	adolescent	M	1.08	0.13	1.01	0.15	0.93	0.17	0.88	0.21
	child	F	1.04	0.18	0.9	0.19	0.99	0.19	0.93	0.26
	child	M	1.03	0.18	0.91	0.2	1.03	0.24	0.94	0.3
LOT	adolescent	F	1.11	0.1	1.16	0.15	0.79	0.12	0.76	0.09
	adolescent	M	1.15	0.12	1.17	0.11	0.82	0.11	0.8	0.1
	child	F	1.04	0.17	1.04	0.2	0.8	0.13	0.74	0.15
	child	M	1.05	0.21	1.05	0.24	0.75	0.13	0.69	0.13
PRICE	adolescent	F	1.42	0.14	1.3	0.18	0.94	0.11	1.14	0.15
	adolescent	M	1.47	0.14	1.34	0.17	1.09	0.12	1.24	0.16
	child	F	1.36	0.25	0.99	0.19	0.99	0.17	1.4	0.21
	child	M	1.45	0.26	1.06	0.26	0.97	0.15	1.36	0.21
TRAP	adolescent	F	1.4	0.15	1.42	0.15	1.05	0.08	1.06	0.12
	adolescent	M	1.42	0.17	1.47	0.15	1.11	0.11	1.12	0.11
	child	F	1.54	0.27	1.43	0.27	1.06	0.14	1.05	0.15

	child	M	1.57	0.21	1.5	0.25	1.04	0.14	1.05	0.15

7.4 Comparing children and adolescents: statistical analysis

This section presents the results of the quantitative comparison of the adolescents' and childrens' diphthong onsets and diphthong trajectories. In this analysis, the key questions are:

1. Is there an effect of age, i.e. do adolescents and children realise this diphthong differently from one another?

2. Do the children and adolescents have the same constraints on variation in the diphthongs, in terms of preceding and following environment?

3. Is there an effect of sex, and/or an age-sex interaction? This relates to the findings of Cheshire et al. (2011) described above (Section 7.1.1)

As in the previous chapter, the results for the diphthong onsets are presented first (Section 7.4.1), followed by the results for the diphthong trajectories (Section 7.4.2).

7.4.1 Diphthong onsets

7.4.1.1 FACE onset F1

Age and sex There was marginal evidence for a main effect of age on FACE onset F1, with adolescents predicted to have a higher onset F1 than children ($\hat{\beta}$=0.16; 95% HDI [-0.03, 0.36]; 95% PD) – i.e. a more open onset to FACE. But there did not appear to be a main effect of sex on onset F1 ($\hat{\beta}$=0.02; 95% HDI [-0.15, 0.20]; 58% PD), nor was there strong evidence for an interaction of age and sex ($\hat{\beta}$=0.04; 95% HDI [-0.14, 0.20]; 66% PD). This can be seen in Figure 7.5, where a sex difference is not evident for either age group, but the children are predicted to have a lower onset F1 than the adolescents.

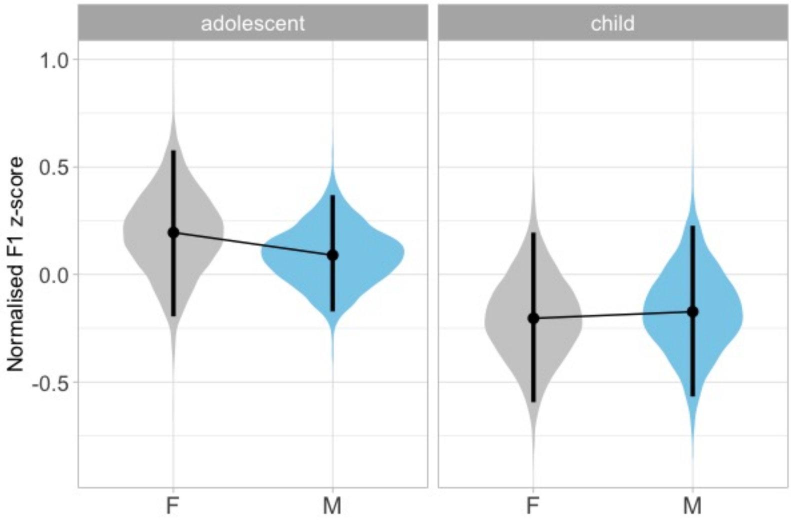


Figure 7.5: FACE onset F1

Preceding environment effects on onset F1 There was strong evidence for a main effect of a preceding approximant or glide: this is associated with a higher onset F1 ($\hat{\beta}$=0.25; 95% HDI [0.11, 0.39]; 100% PD). This category includes words such as *wait, later, racing*. But there was also strong evidence for an interaction of age and preceding approximant or glide ($\hat{\beta}$=-0.15; 95% HDI [-0.28, -0.01]; 99% PD). This is because the effect of a preceding approximant/glide is more pronounced among children (see Figure 7.6).

There was strong evidence for main effect of preceding coronal ($\hat{\beta}$=-0.18; 95% HDI [-0.33, -0.06]; 100% PD), but also marginal evidence for an interaction of age and a preceding coronal ($\hat{\beta}$=0.1; 95% HDI [-0.02, 0.22]; 95% PD). The preceding coronal category includes words such as *they, same, day*. It can be seen from Figure 7.6 that among the adolescents, a preceding coronal is associated with a somewhat higher onset F1, while among the children, a preceding coronal is associated with a lower onset F1.

There was marginal evidence for main effect of preceding nasal ($\hat{\beta}$=0.15; 95% HDI [-0.02, 0.33]; 95% PD), and strong evidence for an interaction of age and preceding nasal ($\hat{\beta}$=0.21; 95% HDI [0.05, 0.39]; 99% PD). This category includes words such as *make, name*. Figure 7.6 shows that for the adolescents, a preceding nasal is associated with a higher onset F1 (this was also found in the previous chapter), whereas for the children, a preceding nasal is associated with a lower onset F1. This will be returned to in the discussion.

There was strong evidence for an interaction of age and preceding "other" ($\hat{\beta}$=-0.24; 95% HDI [-0.41, -0.06]; 99% PD). Words in the "other" category were those where FACE occurs word-initially (e.g. *ache, april*), or where there is an /h/ in syllable onset (e.g. *hay, hey*). It can be seen from Figure 7.6 that adolescents are predicted to have a lower onset

F1 in FACE in this environment, while the opposite is true for the children, such that in this category of FACE words, children are actually predicted to show a higher onset F1 than adolescents.

Finally, there was strong evidence for a main effect of preceding velar, ($\hat{\beta}$=-0.29; 95% HDI [-0.54, -0.02]; 98% PD), and no evidence for an interaction of age and a preceding velar ($\hat{\beta}$=-0.06; 95% CI [-0.34, 0.21]; 68% PD). The most frequently occurring word in this category was *okay* (132 out of 385 tokens with a preceding velar), but this category also includes words such as *cake* (45 tokens), *came* (57 tokens). For both adolescents and children, a preceding velar is associated with a lower onset F1 (see Figure 7.6).

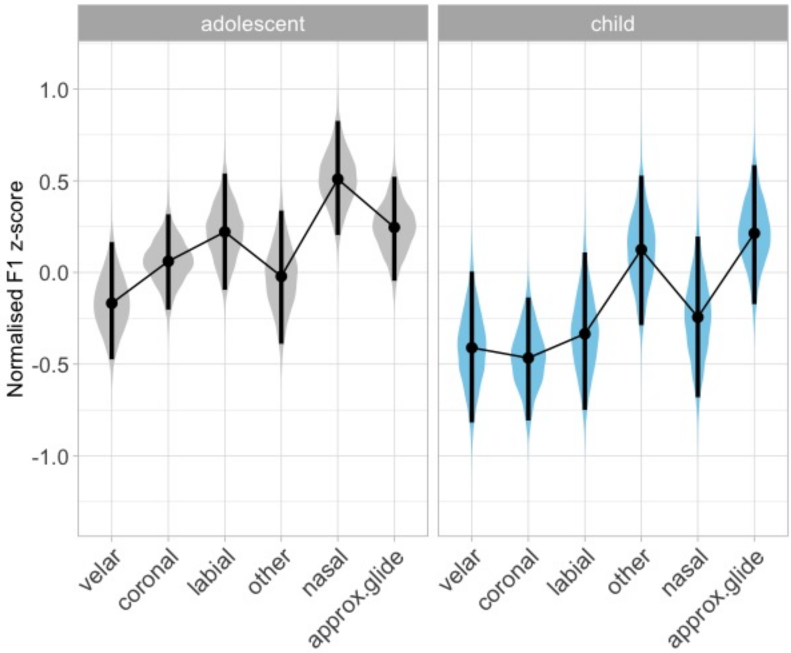


Figure 7.6: FACE onset F1: preceding context

Following environment effects on onset F1 In terms of following environment effects on onset F1, recall that among the adolescents, a coda nasal was associated with a higher F1, while a coda voiced obstruent was associated with a lower F1.

There was marginal evidence for a main effect of FACE occurring word-finally; this is associated with higher onset F1 ($\hat{\beta}$=0.13; 95% HDI [-0.01, 0.28]; 96% PD). Although there was not strong evidence for an interaction with age, Figure 7.7 suggests that the effect of the word-final environment is more pronounced among adolescents.

There was strong evidence for a main effect of a coda voiced obstruent. This is associated with a lower onset F1 ($\hat{\beta}$=-0.2; 95% HDI [-0.4, -0.02]; 98% PD). Again, Figure 7.7 suggests that this effect is more pronounced among adolescents.

There was strong evidence for an interaction of age and a coda nasal ($\hat{\beta}$=0.26; 95% HDI [0.10, 0.43]; 99.90% PD). It can be seen from Figure 7.7 that while a coda nasal predicts a higher onset F1 in adolescent speech, the opposite is true among the children: for them, it predicts a lower onset F1. A similar interaction effect was found for a preceding nasal, and these facts will be returned to in the discussion.

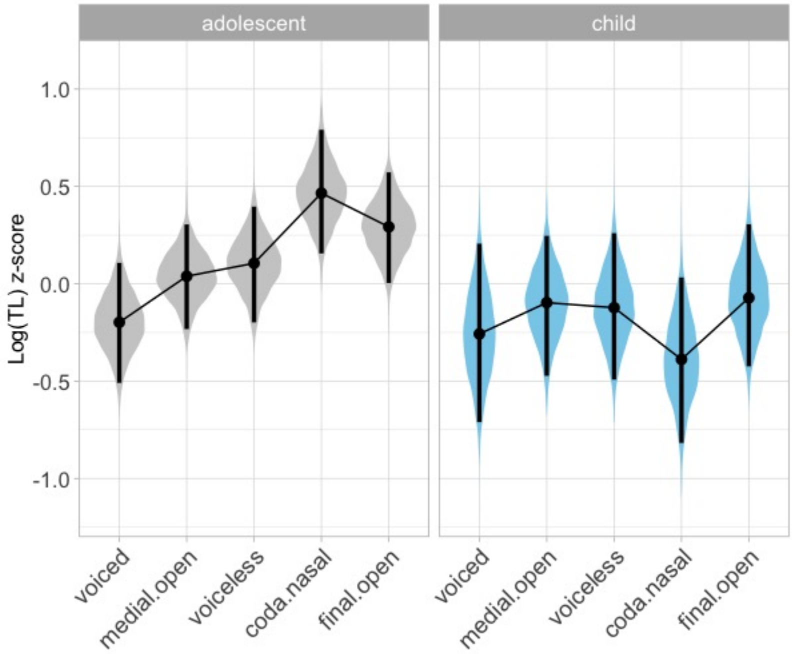


Figure 7.7: FACE onset F1: following context

7.4.1.2 PRICE onset F2

Age and sex There was no evidence for an age difference with regard to PRICE F2, as the posterior was centred at zero ($\hat{\beta}$=0.00; 95% HDI [-0.20, 0.19]; 51% PD). However, there was strong evidence for a main effect of sex ($\hat{\beta}$=-0.27; 95% HDI [-0.45, -0.10]; 99.85% PD), and also strong evidence for an interaction of age and sex ($\hat{\beta}$=-0.25; 95% HDI [-0.40, -0.08]; 99.80% PD). This is because there is a stark sex difference in the F2 of PRICE among adolescents that is not reflected in the children, who instead tend to have a PRICE F2 that is intermediate between that of the adolescent boys and girls. This is shown in Figure 7.8, which shows the posterior predicted median onset F2 for the age and sex groups.

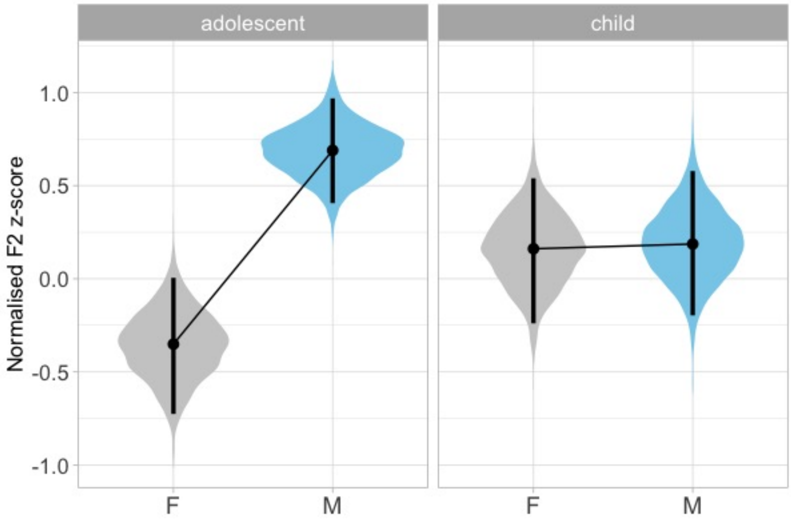


Figure 7.8: PRICE onset F2

Preceding environment effects on onset F2 There was strong evidence that a preceding coronal is associated with a higher onset F2 ($\hat{\beta}$=0.16; 95% HDI [0.01, 0.32]; 98% PD) – this category includes words such as *time, inside, tiger*. However, there was also marginal evidence for an interaction of age and preceding coronal ($\hat{\beta}$=-0.11; 95% HDI [-0.24, 0.02]; 95% PD). This is because while a preceding coronal favours a higher F2 among both age groups, this effect is greater among the children (see Figure 7.9).

There was strong evidence that a preceding labial is associated with a lower onset F2 ($\hat{\beta}$=-0.71; 95% HDI [-0.89, -0.55]; 100% PD). This category includes words with a preceding bilabial or labiodental consonant, e.g. *five* or *bike*. Meanwhile there seemed to be no interaction of age and a preceding labial ($\hat{\beta}$=0.03; 95% HDI [-0.10, 0.16]; 70% PD). This is reflected in Figure 7.9, which shows that both adolescents and children are expected to show a much lower onset F2 following a labial consonant. This is an expected coarticulatory effect, as a preceding bilabial or labiodental consonant is expected to lower F2 (Thomas, 2011, p.101).

There was strong evidence for a main effect of a preceding "other" ($\hat{\beta}$=0.33; 95% HDI [0.08, 0.56]; 99.72% PD), i.e. a preceding "other" predicts a higher onset F2. This category includes words such as *high, ice*. While not strong, there was some evidence for an interaction of age and a preceding "other" ($\hat{\beta}$=-0.15; 95% HDI [-0.35, 0.07]; 91% PD): again, as with a preceding coronal, the effect is greater among the children than among the adolescents, though among both age groups, a preceding "other" is associated with a higher onset F2 (see Figure 7.9).

There was strong evidence that a preceding velar is associated with a higher onset F2 ($\hat{\beta}$=0.46, 95% HDI [0.11, 0.85]; 99% PD), and strong evidence for an interaction of age and a preceding velar ($\hat{\beta}$=0.46; 95% CI [0.14, 0.8]; 100% PD). Words with a preceding velar include *guy, kit, kind*. For both age groups, a preceding velar is associated with a higher onset F2, but this effect is much more pronounced among the adolescents than among the children (see Figure 7.9). The previous chapter linked this effect to the narratives in which the word *guy* is used in adolescent speech.

There was also strong evidence that a preceding approximant or glide is associated with a lower onset F2 ($\hat{\beta}$=-0.12; 95% HDI [-0.25, 0.01]; 96% PD), and there was no evidence for an interaction of age and this environment. This category includes words such as *why, life, right*. Figure 7.9 shows that this effect is consistent across adolescents and children.

In sum, the preceding environment effects are largely consistent across the two age groups (see Figure 7.9), with the exception of the preceding velar environment, which was discussed in the previous chapter.

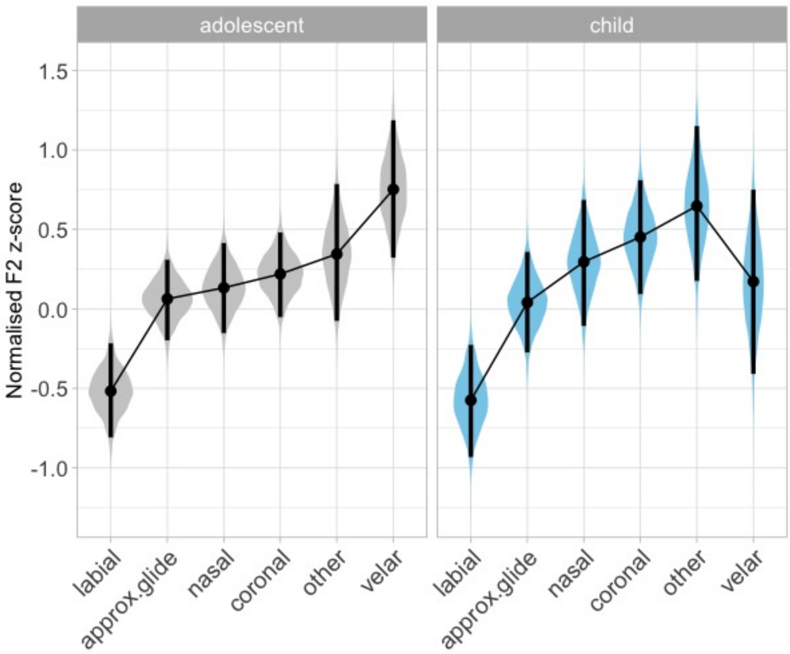


Figure 7.9: PRICE onset F2: preceding context

Following environment effects on onset F2 Recall that the analysis of PRICE variation among the adolescents found no following environment effects on onset F2. When the data from the adolescents and children are analysed together, there was strong evidence that PRICE occurring in a word-medial open syllable tends to have a higher F2 compared to other syllable types, though the effect size is small ($\hat{\beta}$=0.14; 95% HDI [0.01, 0.27]; 97.40% PD). It can be seen from Figure 7.10 that for both adolescents and children, the word-medial open syllable is the environment that favours the highest onset F2. Words in this category include *spider, tiger*.

There was also strong evidence for an interaction of age and the environment of a coda voiced obstruent ($\hat{\beta}$=0.12; 95% HDI [0.01, 0.24]; 97.78% PD): among adolescents, a coda voiced obstruent is associated with a higher onset F2, but among the children, a coda voiced obstruent is associated with a lower onset F2. This can be seen in Figure 7.10. Words in this category include *five, tried*.

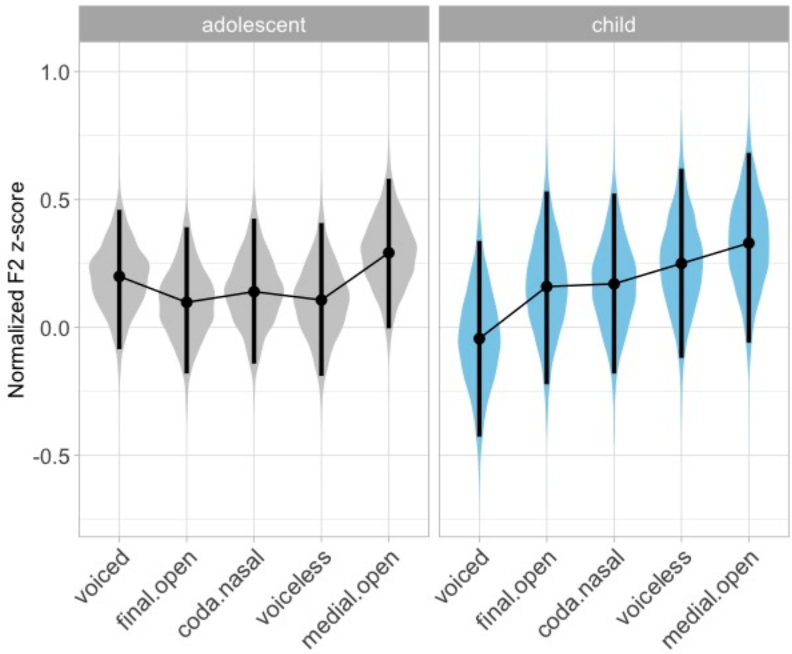


Figure 7.10: PRICE onset F2: following context

7.4.1.3 PRICE in *like* onset F2

Recall that in the previous chapter, we analysed PRICE in *like* separately from other PRICE words, because the lemma *like* accounted for almost half of the PRICE tokens among the adolescents (see token counts in Appendix B). Once adolescents from outcode 3 had been removed from the data, this left 866 *like* tokens from the adolescents, compared to only 117 from the children.

The model did not find strong evidence for an age difference in PRICE onset F2 in *like*; however, it does allow us to be 90% confident that the children actually have a higher onset F2 than the adolescents in the PRICE vowel in *like* ($\hat{\beta}$=-0.18; 95% HDI [-0.46, 0.08]; 90.75% PD). The evidence for a main effect of sex also did not reach the criteria for strong evidence ($\hat{\beta}$=-0.21; 95% HDI [-0.49, 0.06]; 93% PD), though this means that we can be 93% confident that boys have a higher onset F2 than do the girls. There was also strong evidence for an interaction of age and sex, similar to that found for the other PRICE words ($\hat{\beta}$=-0.30; 95% HDI [-0.58, -0.02]; 98% PD): among the children, girls are predicted to have a slightly higher onset F2 than the boys, but among the adolescents, the girls are predicted to have a much lower onset F2 than the boys. These findings are represented graphically in Figure 7.11: the adolescent girls are predicted to have a lower onset F2 than the adolescent boys; the children are predicted to show an onset F2 that is more similar to the adolescent boys than to the adolescent girls.

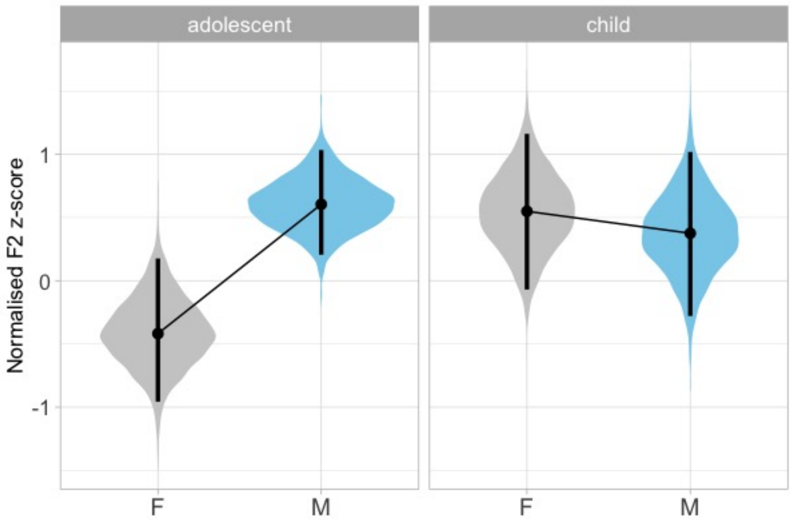


Figure 7.11: PRICE in *like* onset F2

7.4.1.4 GOAT onset F2

Age and sex Adolescents are estimated to have a slightly lower onset F2 compared to the children, but the evidence for this effect was not strong ($\hat{\beta}$=-0.08; 95% HDI [-0.34, 0.17]; 75% posterior < 0).

There was some evidence for a main effect of sex on GOAT F2, though this did not quite reach the criteria for strong evidence ($\hat{\beta}$=0.16, 95% HDI [-0.09, 0.39]; 91% PD). There was also marginal evidence for an interaction of age and sex ($\hat{\beta}$=0.19; 95% HDI [-0.03, 0.39]; 96% PD). This is because the evidence for a sex difference is stronger among the adolescents than among the children: among adolescents, girls are estimated to have an onset F2 that is 0.70 standard deviations greater than the boys (95% HDI [0.09, 1.32]); while among the children, girls are estimated to have an onset F2 that is -0.05 standard deviations lower than that of the boys (95% HDI [-0.72, 0.60]). It can be seen in Figure 7.12 that adolescent girls are predicted to have a much higher onset F2 than the boys, while the children are predicted to have an onset F2 that is intermediate between the adolescent girls and adolescent boys.

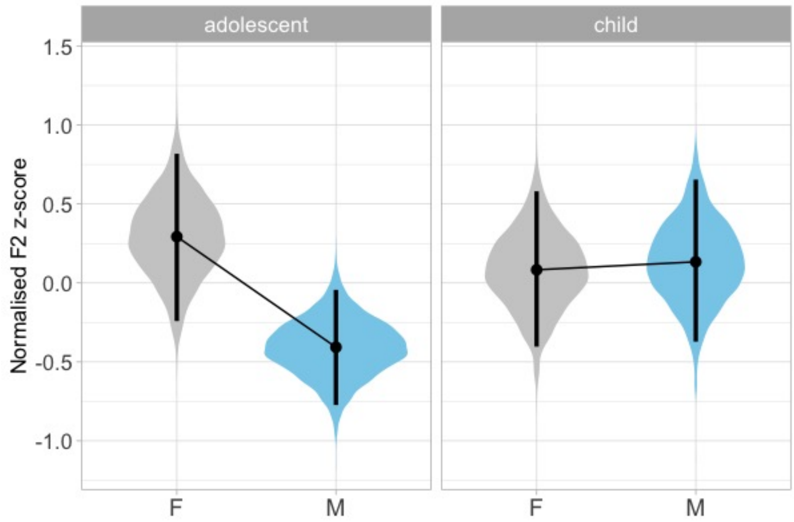


Figure 7.12: GOAT onset F2

Preceding environment effects on onset F2 There was marginal evidence that a preceding approximant or glide is associated with a lower onset F2 ($\hat{\beta}$=-0.14; 95% HDI [-0.30, 0.03]; 95% PD), and no evidence for an interaction effect here. Words in this category include *road, hello, won't*.

There was strong evidence that a preceding coronal is associated with a higher onset F2 ($\hat{\beta}$=0.20; 95% HDI [0.04, 0.35]; 99% PD). But there is also marginal evidence for an interaction of age with a preceding coronal ($\hat{\beta}$=0.14; 95% HDI [-0.01, 0.30]; 96% PD): while among both adolescents and children, a preceding coronal predicts a higher onset F2, this effect is more pronounced among the adolescents (Figure 7.13). Words in this category include *so, don't, though*.

There was marginal evidence that a preceding labial is associated with a lower onset F2 ($\hat{\beta}$=-0.19; 95% HDI [-0.39, 0.01]; 97% PD), as would be expected (Thomas, 2011, p.101). There was no evidence for an interaction of age and this preceding environment. Words with a preceding bilabial or labiodental include *both, phone, supposed*.

There was strong evidence that a preceding nasal is associated with a higher onset F2 ($\hat{\beta}$=0.17; 95% HDI [0, 0.36]; 97% PD). There is some indication of an interaction of age and the preceding nasal environment ($\hat{\beta}$=-0.12; 95% HDI [-0.26, 0.04]; 94% PD): for children, a preceding nasal predicts a higher onset F2, but this is not so clearly the case for adolescents (see Figure 7.13). Words in this category include *no, know, most*.

Looking at Figure 7.13, it seems fair to say that the preceding environment effects are not entirely consistent across the two age groups, but at the same time, the preceding environment effects are not especially pronounced in either age group (with the exception of a preceding nasal) – none of the effect sizes discussed in this section is greater than ± 0.20.

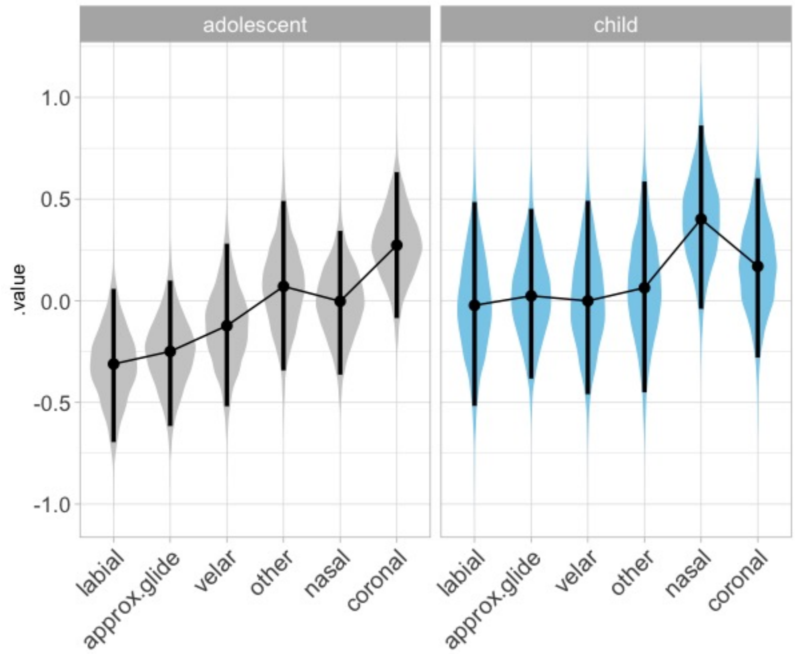


Figure 7.13: GOAT onset F2: preceding context

Following environment effects on onset F2 There was strong evidence that a coda voiced obstruent is associated with a higher onset F2 ($\hat{\beta}$=0.23; 95% HDI [0.06, 0.43]; 99% PD), and at the same time, there was marginal evidence for an interaction of age and a coda voiced obstruent ($\hat{\beta}$=-0.15; 95% HDI [-0.32, 0.02]; 96% PD). While among the children, a coda voiced obstruent favours a higher onset F2, this is not the case for the adolescents (Figure 7.14). Words in this category include *road, goes, those*.

There was marginal evidence for an interaction of age and a word-medial open syllable ($\hat{\beta}$=0.14; 95% HDI [-0.01, 0.29]; 97% PD). Words in this category include *okay, over, joking*. This is because for adolescents, this environment favours a somewhat higher onset F2, while for children, this environment favours a lower onset F2 (Figure 7.14).

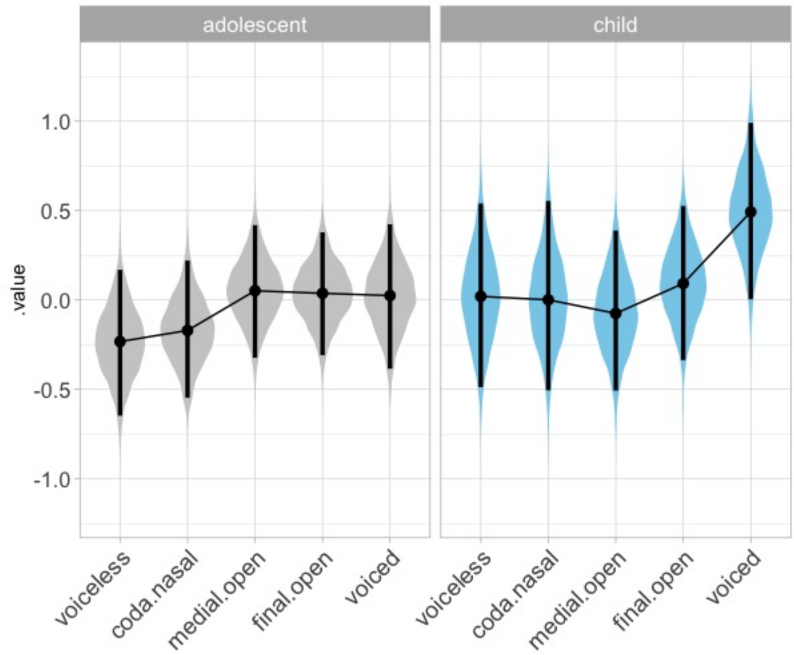


Figure 7.14: GOAT onset F2: following context

7.4.1.5 Summary of results from diphthong onsets

Only one of the diphthong onsets showed even marginal evidence for a main effect of age, and that was FACE F1. The children were found to have a lower onset F1 than the adolescents; this means that the children show a more MLE-like realisation of the FACE onset than the adolescents.

At the same time, age-sex interactions emerged for the remaining onset variables:

- PRICE onset F2: the adolescent girls were predicted to have a lower onset F2 than the adolescent boys; the children were predicted to have an onset F2 in PRICE that is intermediate between the adolescent boys and girls. A sex difference was not found among the children.

- In PRICE in *like*, adolescent boys were again found to have a higher onset F2 than girls, but the children are estimated to show an onset F2 that is more similar to the adolescent boys than to the adolescent girls.

- GOAT onset F2 showed a similar pattern. The adolescent girls were predicted to have a much higher onset F2 than the adolescent boys; not only was a sex difference absent among the children, but the children were predicted to have an onset F2 that was inbetween that of the adolescent boys and girls.

The children showed the following similar preceding/following environment constraints to the adolescents in the PRICE and GOAT onsets:

- Preceding environment constraints on PRICE F2. For both adolescents and children: a preceding coronal favoured a higher onset; a preceding "other" favoured a higher onset F2; a preceding labial favoured a lower onset F2; and a preceding approximant or glide favoured a lower F2.

- Preceding environment constraints on GOAT F2: a preceding coronal favoured a higher onset F2; a preceding labial favoured a lower F2; and a preceding approximant or glide favoured a lower F2.

Meanwhile, interactions of age and preceding/following environment were found for the onsets of all three diphthongs:

- FACE: the preceding and following environment effects on onset F1 differed greatly between the children and adolescents. Among the adolescents, a preceding or following nasal is associated with a higher onset F1 i.e. more open nucleus. In the previous chapter, it was observed that this aligns with findings in Newham (Gates, 2018). However, the children do not show this constraint. Among the adolescents, a preceding velar favoured a lower F1; among the children, a preceding approximant or a preceding "other" favoured a higher F1, while a preceding labial favoured a lower F1. Among the children, no following environment effects on onset F1 were found; while among the adolescents, as already stated, a coda nasal was associated with a higher F1, and a coda voiced obstruent favoured a lower F1.

- PRICE: the adolescents and children showed largely consistent preceding environment constraints on onset F2, but there were several interactions of age and following environment. Among the adolescents, a coda voiced obstruent favoured a higher onset F2, but among the children, a coda voiced obstruent favoured a lower onset F2.

- GOAT: there were also differences in the following environment constraints on GOAT variation between children and adolescents. For example, following environment effects on onset F2 seemed to be absent among the adolescents, but for the children, GOAT occurring word-medially in an open syllable favoured a lower F2, while a coda voiced obstruent favoured a higher onset F2.

7.4.2 Diphthong trajectories

We turn now to the results for the diphthong trajectories.

7.4.2.1 FACE Trajectory Length

Age and sex With Trajectory Length, meanwhile, there was some evidence that the children show a greater TL than the adolescents, though this did not reach the criteria for strong evidence ($\hat{\beta}$=-0.06; 95% HDI [-0.20, 0.07]; 83% PD). But at the same time, there was marginal evidence for a main effect of sex ($\hat{\beta}$=0.11; 95% HDI [-0.01, 0.22]; 96% PD), and strong evidence for an interaction of age and sex ($\hat{\beta}$=0.15; 95% HDI [0.04, 0.26]; 99.48% PD). This is because among the adolescents, there is a clear sex difference, with girls showing a greater TL than boys, but this difference is not found among the children. This can be seen in Figure 7.15, which shows that adolescent boys are predicted to have a smaller TL i.e. monophthongal realisation of FACE, while adolescent girls are predicted to have a more diphthongal realisation, and children are predicted to favour a FACE variant that is intermediately diphthongal. This strongly resembles the pattern found for GOAT onset F2 (Section 7.4.1.4) and PRICE onset F2 (Section 7.4.1.2).

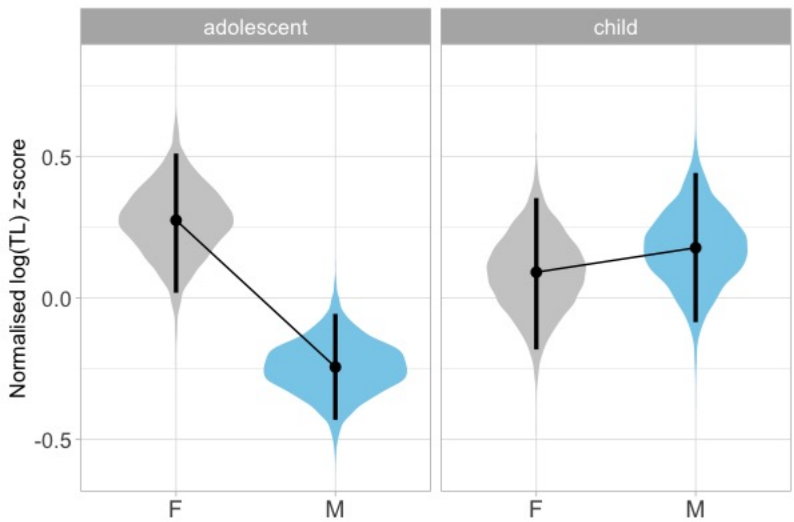


Figure 7.15: FACE TL

Preceding environment effects on TL There was strong evidence that a preceding approximant or glide is associated with a greater TL ($\hat{\beta}$=0.28; 95% HDI [0.18, 0.39]; 100% PD), and strong evidence that a preceding velar is associated with a smaller TL ($\hat{\beta}$=-0.25; 95% HDI [-0.46, -0.07]; 100% PD); and neither strong nor marginal evidence for an interaction of age with either of these types of preceding environment.

Taking this together with the results from FACE onset F1 (Section 7.4.1.1), this means that a preceding approximant or glide is associated with a more conservative realisation of FACE, i.e. one that has a more open onset and is more diphthongal; while a preceding velar is associated with a more MLE-like realisation of FACE, i.e. one with a closer onset and that is more monophthongal; and these effects are consistent across both children and adolescents (cf. Figures 7.6 and 7.16).

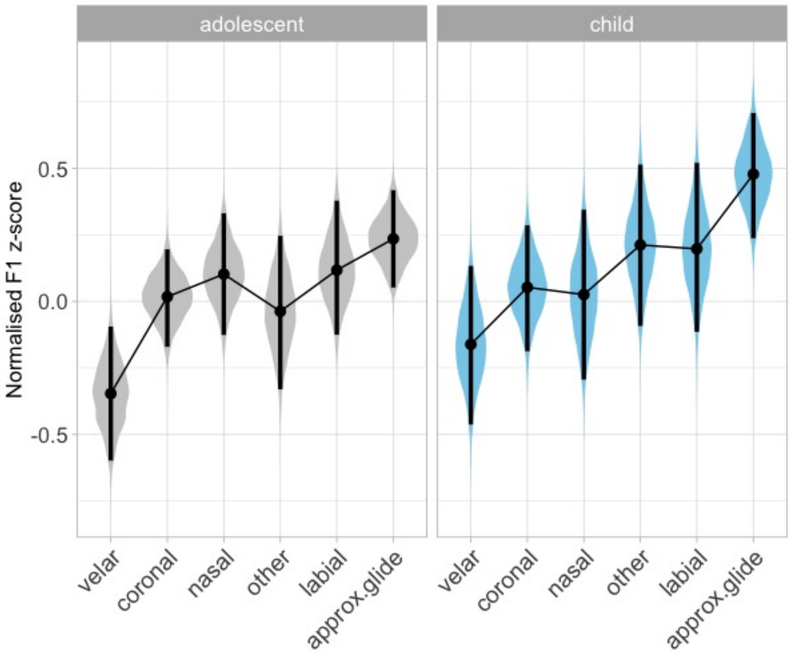


Figure 7.16: FACE TL: preceding context

Following environment effects on TL In the previous chapter, the analysis of FACE TL variation among the adolescents suggested that a coda voiced obstruent and a word-final open syllable both favoured a smaller TL, i.e. a more monophthongal pronunciation.

When the adolescents' and children's data were modelled together, there was marginal evidence for a main effect of coda voiced obstruent: this is associated with a smaller TL ($\hat{\beta}$=-0.14; 95% HDI [-0.3, 0.01]; 96% PD). There was also strong evidence for a main effect of a coda voiceless obstruent, which is associated with a greater TL i.e. more diphthongal realisation ($\hat{\beta}$=0.21; 95% HDI [0.06, 0.36]; 100% PD). Although there was neither strong nor marginal evidence for interactions of age and either coda voiced or coda voiceless obstruent, it can be seen from Figure 7.17 that the effect of a voiced vs. voiceless coda obstruent is more pronounced among the children than among the adolescents.

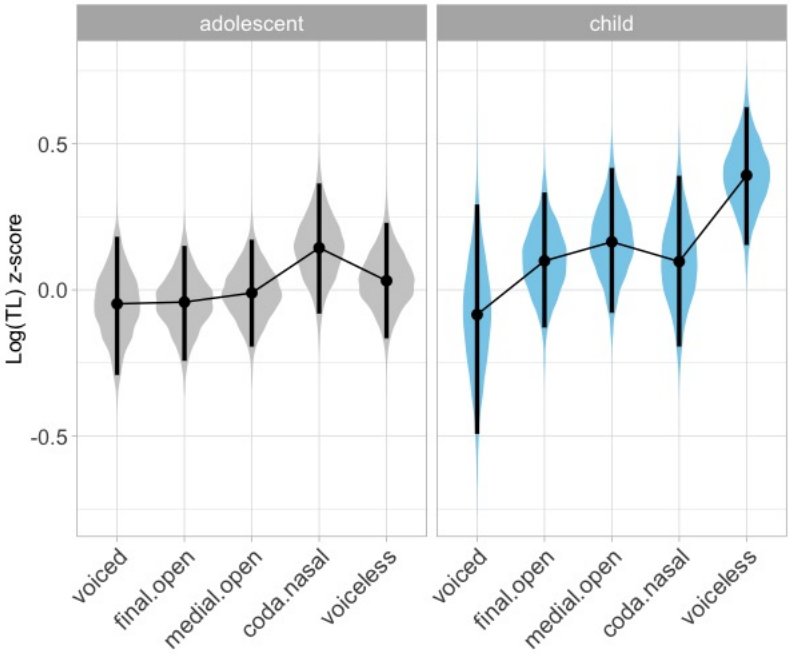


Figure 7.17: FACE TL: following context

7.4.2.2 PRICE Trajectory Length

Age and sex There was strong evidence for an effect of age ($\hat{\beta}$=-0.28; 95% HDI [-0.38, -0.19]; 100% PD), meaning that the adolescents tend to have a smaller TL i.e. more monophthongal realisation of PRICE than the children. Although it did not reach the criteria for strong evidence, there was some evidence for a main effect of sex ($\hat{\beta}$=0.06; 95% HDI [-0.03, 0.14]; 91% PD), and also strong evidence for an interaction of age and sex ($\hat{\beta}$=0.12; 95% HDI [0.04, 0.21]; 99.67% PD). This means that while the children generally have a more diphthongal realisation of PRICE than do the adolescents, within the adolescents, girls are estimated to have a greater TL than the boys; while among the children, the evidence for a sex difference is not so strong. The model's predictions for PRICE TL are represented in Figure 7.18.

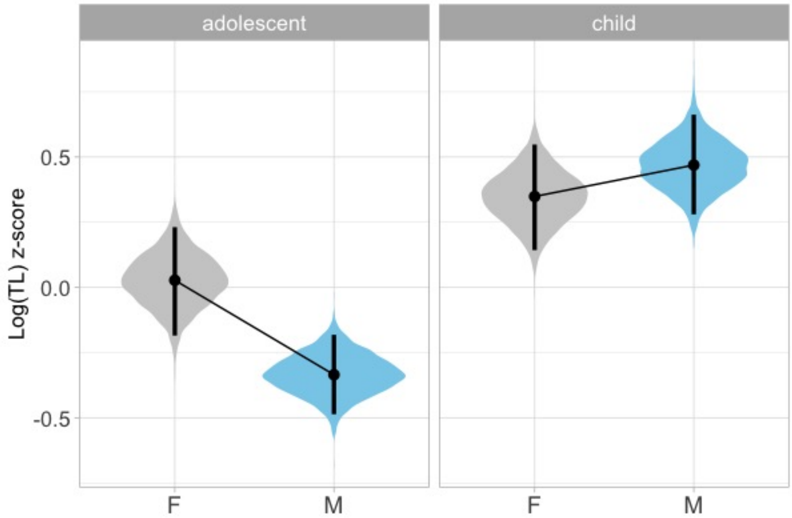


Figure 7.18: PRICE TL

Preceding environment effects on TL There was strong evidence that a preceding coronal is associated with a smaller TL ($\hat{\beta}$=-0.18; 95% HDI [-0.33, -0.04]; 99.20% PD). Taken together with the evidence from onset F2, this means that a more innovative/more MLE-like realisation of PRICE is likely in words with a preceding coronal (e.g. *time, inside, tiger*).

There was strong evidence that a preceding "other" is associated with a greater TL ($\hat{\beta}$=0.35; 95% HDI [0.18, 0.54]; 100% PD); and strong evidence that a preceding approximant or glide is associated with a smaller TL ($\hat{\beta}$=-0.14; 95% HDI [-0.24, -0.04]; 99.5% PD);

There did not appear to be an interaction of any of these preceding environments with age. Indeed, it can be seen from Figure 7.19 that the preceding environment constraints on PRICE TL are relatively consistent across adolescents and children.

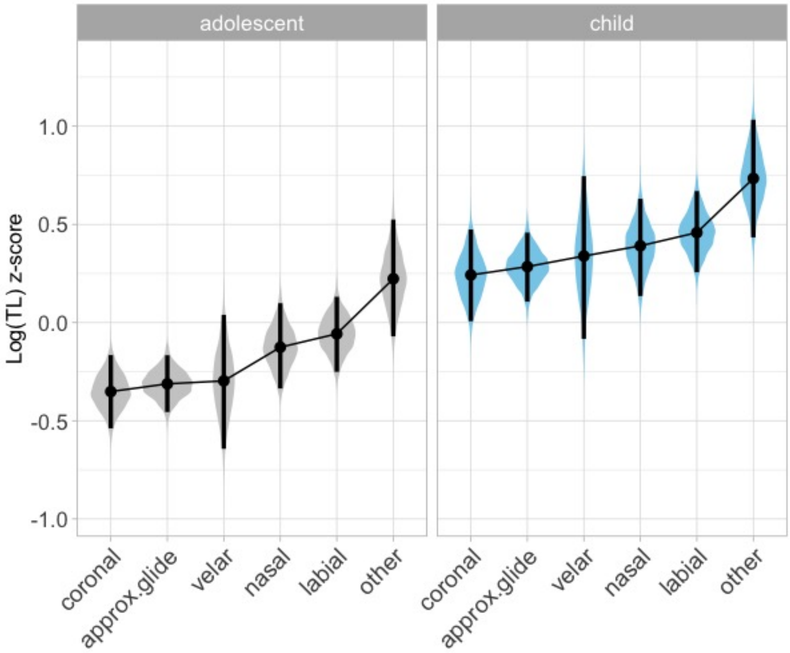


Figure 7.19: PRICE TL: preceding context

Following environment effects on TL There was strong evidence for several interactions of age and following environment, suggesting that the following environment constraints are not entirely consistent between the children and the adolescents.

In terms of main effects, there was strong evidence that PRICE occurring word-finally (e.g. *why, guy*) is associated with a smaller TL ($\hat{\beta}$=0.13; 95% HDI [0.04, 0.21]; 99.85% PD). There was also strong evidence that PRICE occurring word-medially in an open syllable is associated with a greater TL ($\hat{\beta}$=0.13; 95% HDI [0.03, 0.23]; 100% PD); and strong evidence that a coda voiceless obstruent (e.g. *white, kite*) favours a greater TL ($\hat{\beta}$=0.16; 95% HDI [0.01, 0.30]; 98% PD).

Turning now to the interaction effects, there was strong evidence for an interaction of age and a coda nasal ($\hat{\beta}$=0.13; 95% HDI [0.04, 0.21]; 99.85% PD): among adolescents, a coda nasal favours a greater TL, while among children, a coda nasal favours a more monophthongal realisation (see Figure 7.20). Words in this category include *time, nine*.

There was marginal evidence for interaction of the word-final open syllable environment and age ($\hat{\beta}$=0.07; 95% HDI [-0.01, 0.15]; 95% PD): among children, this environment is associated with a more monophthongal pronunciation of PRICE, but this effect is much weaker among adolescents.

Finally, there was strong evidence for an interaction of age and the word-medial open syllable environment ($\hat{\beta}$=-0.10; 95% HDI [-0.20, -0.01]; 97% PD). Among children, this environment is associated with a more diphthongal pronunciation, but this effect is not found among adolescents (Figure 7.20).

In sum, it seems that the adolescents do not have major following environment constraints on PRICE monophthongisation, but that the children do have some constraints. This fact will be returned to in the discussion.

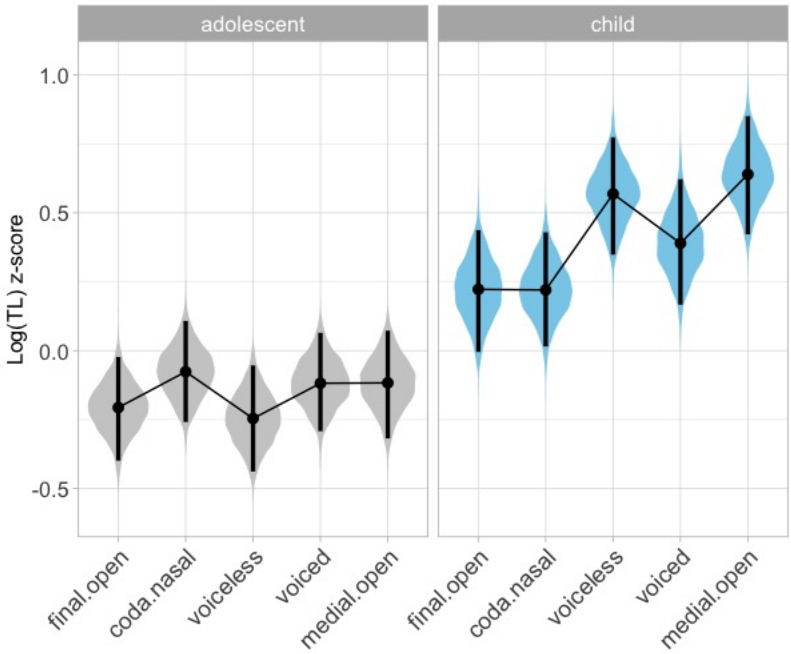


Figure 7.20: PRICE TL: following context

7.4.2.3 PRICE in *like* Trajectory Length

There was strong evidence for an age difference ($\hat{\beta}$=-0.55; 95% HDI [-0.72, -0.39]; 100% PD). The children have a greater TL i.e. more diphthongal realisation of PRICE in *like* than do the adolescents, by around 1.1 standard deviations (95% HDI [-1.44, -0.79]; cf. Figure 7.21). Meanwhile, neither the main effect of sex, nor the interaction of age and sex, reached the criteria for strong evidence.

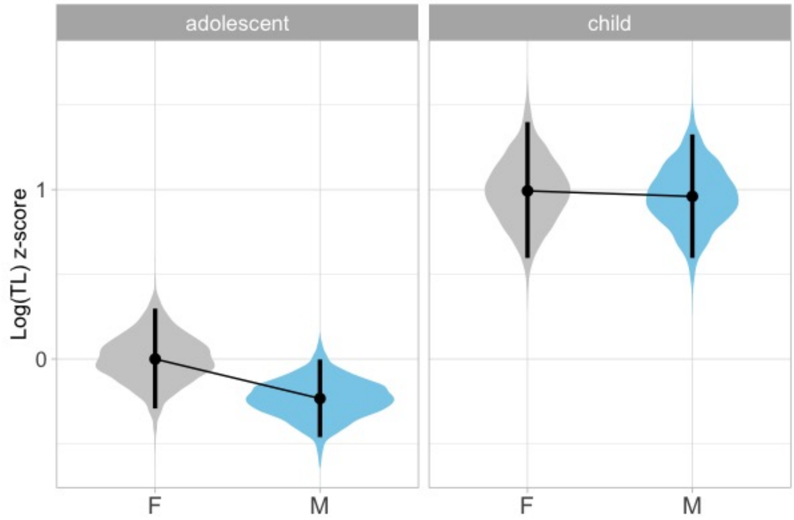


Figure 7.21: PRICE in *like* TL

7.4.2.4 GOAT Trajectory Length

Age and sex There was strong evidence that the adolescents tend to have a smaller TL than the children, i.e. a more monophthongal realisation of GOAT ($\hat{\beta}$=-0.17; 95% HDI [-0.29, -0.04]; 99.5% posterior $<$ 0). This can be seen in Figure 7.22. There was no evidence for a main effect of sex on GOAT TL ($\hat{\beta}$=0.03, 95% HDI [-0.09, 0.15]; 69% PD), and similarly, no evidence for an interaction effect of age and sex on GOAT TL ($\hat{\beta}$=0.02, 95% HDI [-0.10, 0.13]; 62% PD).

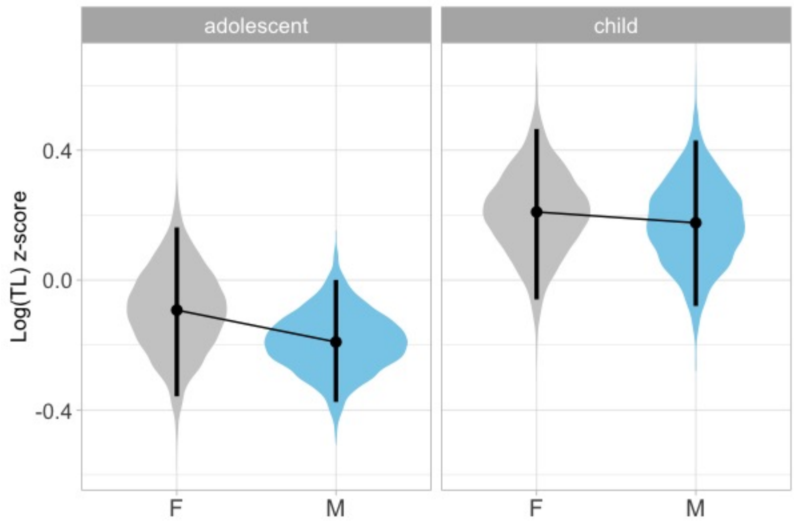


Figure 7.22: GOAT TL

Preceding environment effects on TL Recall that in the previous chapter, it was found that among the adolescents, a preceding velar favoured a smaller TL.

There was strong evidence that a preceding nasal is associated with a greater TL ($\hat{\beta}$=0.18; 95% HDI [0.06, 0.29]; 99.60% PD), and no evidence for an interaction effect of age and preceding nasal.

There was some evidence that a preceding coronal is associated with a smaller TL, and similarly, evidence that a preceding velar is associated with a smaller TL, though this did not quite reach the criteria for strong evidence in either case (preceding coronal: $\hat{\beta}$=-0.07, 95% HDI [-0.17, 0.03], 93% PD; preceding velar $\hat{\beta}$=-0.14, 95% HDI [-0.33, 0.04], 93% PD).

There was not strong not marginal evidence for any interaction effects of age and preceding environment.

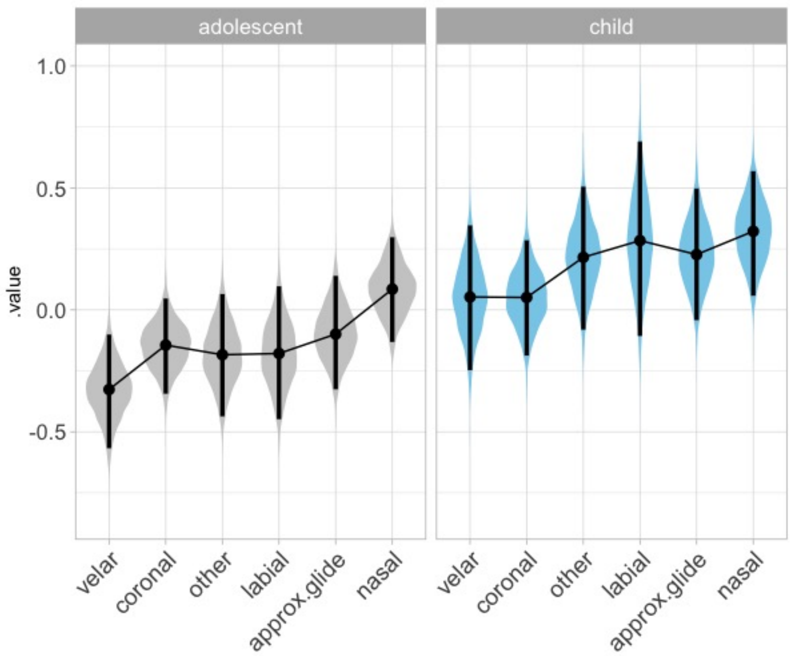


Figure 7.23: GOAT TL: preceding context

Following environment effects on TL There was strong evidence that a coda nasal favours a greater TL ($\hat{\beta}$=0.15; 95% HDI [0.00, 0.29]; 98% PD) – e.g. in words such as *don't, only, home* – and no evidence for an interaction of age with this environment. It can be seen from Figure 7.24 that for both adolescents and children, a coda nasal favours a greater TL.

There was marginal evidence that GOAT occurring word-finally in an open syllable tends to have a smaller TL ($\hat{\beta}$=-0.09; 95% HDI [-0.18, 0.01]; 96% PD), and again, no interaction effect. Words in this category include *no, so, go*.

None of the interactions between age and following environment reached the criteria for strong or marginal evidence; however, Figure 7.24 suggests that the effects of following context on GOAT TL are not entirely consistent between the two age groups.

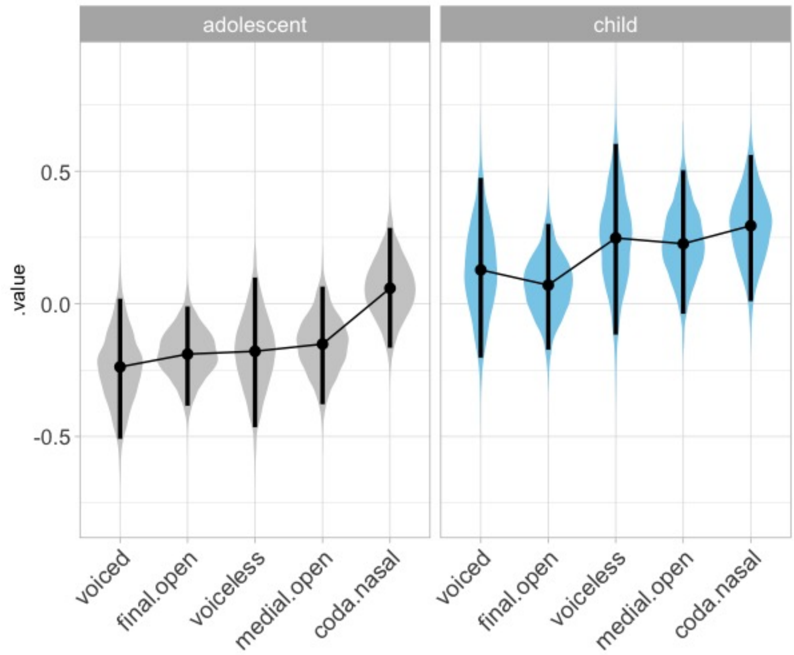


Figure 7.24: GOAT TL: following context

7.4.2.5 Summary of results from diphthong trajectories

Main effects of age were found for both the PRICE (including *like*) and GOAT trajectories, with adolescents predicted to show a more monophthongal realisation. The adolescents were also predicted to show a more monophthongal realisation of FACE than the children, though this did not reach the criteria for strong or marginal evidence.

At the same time, an age-sex interaction emerged for FACE TL, similar to those found for PRICE and GOAT onset F2: the adolescent girls were predicted to have a more diphthongal realisation of FACE than the boys; while the children were predicted to have a FACE TL that was intermediate between that of the adolescent girls and boys. A sex difference was not found among the children.

The children showed similar preceding/following environment constraints to the adolescents for these variables:

- Preceding environment constraints on PRICE TL. For both adolescents and children: a preceding coronal favoured a smaller TL, as well as a higher onset F2, indicating an overall more MLE-like pronunciation; a preceding "other" favoured a greater TL, as well as a higher onset F2; and a preceding approximant or glide favoured a smaller TL as well as a lower F2.

- Preceding environment constraints on FACE TL. For both children and adolescents, a preceding approximant or glide favoured a greater TL i.e. more diphthongal pronunciation; while a preceding velar favoured a smaller TL.

- Preceding environment constraints on GOAT and TL: a preceding coronal favoured a smaller TL, as well as a higher onset F2; a preceding nasal favoured a greater TL; a preceding velar favoured a smaller TL.

- Following environment constraints on GOAT TL. For both children and adolescents, a preceding nasal favoured a greater TL, while a preceding velar favoured a smaller TL.

The following complex interactions of age and preceding/following environment on the following variables were found:

- PRICE: the adolescents and children showed largely consistent preceding environment constraints on TL, but there were several interactions of age and following environment. For the adolescents, a coda nasal favoured a greater TL, while for the children, a coda nasal favoured a smaller TL. Among the children, but not among the adolescents, PRICE occurring word-medially in an open syllable tended to have a greater TL; and among the children, but not so much among the adolescents, PRICE occurring in a word-final open syllable is estimated to be more monophthongal.

7.5 Discussion

The research questions presented at the beginning of Section 7.4 were: Do children realise the diphthongs FACE, PRICE and GOAT in the same way as adolescents in their community, in terms of (a) FACE-raising, PRICE-fronting and GOAT-backing, and (b) monophthongisation of these diphthongs? If there are language-internal constraints on sociophonetic variation in these diphthongs, have the children acquired the same constraints as the adolescents? And finally, is there an effect of sex, and/or an age-sex interaction, similar to that reported by Cheshire et al. (2011)?

The major finding of this analysis is that the children do indeed show the same vowel system as the adolescents in terms of diphthong onsets, replicating the findings of Cheshire et al. (2011); Kerswill et al. (2013). This was suggested by the vowel plots presented in Section 7.3 and confirmed by the statistical analysis presented in Section 7.4: for PRICE and GOAT, there was not compelling evidence of a main effect of age, and for FACE, there was strong evidence that the children have a closer FACE onset than the adolescents – i.e. if there is a change in progress of FACE raising, the children are actually ahead of the adolescents.

However, the findings for diphthong dynamics are quite different. For PRICE and GOAT, we found strong evidence that the adolescents have a more monophthongal pronunciation of these vowels than the children. For FACE, the findings were complicated by an age-sex interaction: the adolescent females were more diphthongal than the adolescent males, and the children were intermediately diphthongal.

In fact, similar age-sex interactions were found for PRICE F2 and GOAT F2. For FACE TL, and PRICE and GOAT onset F2, a particular age-sex interaction pattern was found: the adolescent boys showed the most MLE-like realisation, i.e. the most monophthongal FACE, the most front PRICE onset and the most back GOAT onset; the adolescent girls showed a relatively diphthongal FACE, back PRICE onset and central/front GOAT onset; and the children showed no sex difference among themselves, and also showed phonetic realisations of these variables that were intermediate between the adolescent boys and girls. These findings align remarkably closely with those of Cheshire et al. (2011), who found that gender differences were not in evidence before age 8. They also align with a body of studies that have found no gender differences in sociolinguistic usage in young children; while other studies have found that gender differences are in evidence even in children this young.

The results from the comparison of language-internal constraints presented a mixed picture. If we look particularly at PRICE, the children showed almost identical probabilities of preceding environment effects on both the onset F2 and the TL of PRICE. The finding that the children show similar preceding environment constraints on PRICE monophthongisation is important because it shows that even though the children do not share the same phonetic implementation of PRICE as the adolescents, they have the same constraints on variation in this diphthong – thus, even though the children's pronunciations differ from those of the adolescents, they seem to be participating in the same system of variation. However, we cannot conclude that the children have acquired the same system of variation as the adolescents, because the two age groups show different following environment effects on PRICE monophthongisation: the previous chapter found that the adolescents show a greater TL before a coda nasal, and the current chapter found an age-sex interaction for this environment.

This brings us onto the complicated matter of why the children show such different realisations of PRICE and, to a lesser extent, GOAT, than do the adolescents.

7.5.1 Why are the children more diphthongal than the adolescents in PRICE and GOAT?

The first possibility is that in these children's ongoing acquisition of sociolinguistic norms in their community, the target is not MLE, but actually SSBE or the levelled variety of English typical of the southeast of England. This would explain the central and open onset to PRICE, the close-mid onset to FACE and the close-mid central onset of GOAT. To some extent, these two possible targets – levelling changes in the southeast, and MLE – overlap (cf. Kerswill et al., 2008), but this possibility would certainly explain why the children have more diphthongal diphthongs than the adolescents.

Impressionistically, this seems to be true of some children, but not others. Chapter 4 presented a small-scale qualitative examination of phonetic variation in one child's speech, and noted his use of some MLE variables, and the absence of others from his speech. There is certainly a huge degree of interspeaker variation within the children, with some children variably using SSBE accents, and features linked to the southeast rather than to London youth language – for example, among the female children, there were occasional tokens of GOAT fronted in the way described by Kerswill and Williams (2000a), i.e. [əʏ].

At the same time, we might expect the children's speech to be closer to the standard than that of the adolescents, simply because adolescents are known to be especially innovative in their language use (Kirkham & Moore, 2013; Tagliamonte, 2016b). Even though transmission and incrementation presumably do not occur in the same way in this community as they do in monolingual communities (Labov, 2001a, 2007; see Cheshire et al., 2011), this does not necessarily mean that we should not expect an adolescent peak: the adolescent peak could occur simply because of age-grading. However, if this is the correct explanation – i.e. the adolescents are simply more innovative in their use of sociophonetic variation than the children – this does not explain why this age difference has emerged for monophthongisation of the diphthongs, and not for the diphthong onset qualities.

It is worth considering the influence of (a) the nature of the Diapix task and the recording setting, and (b) the fieldworker's accent. We know that children are enormously sensitive to communicative context and to interlocutor (e.g. Smith & Durham, 2019; Lanza, 1992; Matras, 2009; Khattab, 2013). Although the preliminary analysis did not find any effects of addressee or Diapix keyword on the children's pronunciation of these diphthongs, that analysis only compared diphthong tokens from within the same kind of recording session – even if the child was addressing their peer rather than the fieldworker, the fieldworker was still present, and it is possible that the children accommodated to her SSBE/RP accent. Moreover the recording took place during school hours, and even though it was stressed to the children that the Diapix task was a game, their interpretation of the communicative context would have been influenced by the fact that the activity was carried out in a room at school, and overseen by an adult. By contrast, the adolescent data analysed here came from sociolinguistic interviews conducted in their youth centre.

Self-recordings were made by some of the children in the playground and at lunchtime, but these were not included in the current analysis because of the difficulty of firstly, conducting acoustic analysis on self-recordings, and secondly, the problem of comparing data recorded in a quiet setting with data recorded in a noisy setting. Future work will endeavour to analyse the diphthong pronunciations of these children in their playground speech, to see if the communicative context is indeed the reason why the children's diphthong productions were more diphthongal than the adolescents'.

It may also be relevant to consider the local Feature Pool. It may be the case that children in this area of Ealing are more likely to be exposed to diphthongs in the English input they receive from caregivers than children in other parts of London, such as Hackney. Section 7.1.2 reviewed the languages and varieties of L2-accented English that the children are likely to be exposed to. There is no reason to expect a monophthongal substitution for PRICE in Somali-accented English; similarly, the evidence from the study by Evans and Alshangiti (2018) suggests that Gulf Arabic-accented English would have a diphthongal PRICE variant – unfortunately there is not a similar study with speakers of other Arabic dialects. Cross-linguistically, diphthongs are not common (Maddieson, 1984, p.133), and so we would expect that in a community like the one described in Hackney, where no one linguistic group is in a majority, a likely outcome of the Feature Pool would be monophthongisation of the English diphthongs. But in Ealing, the most 5 populous languages are English, Polish, Punjabi, Somali and Arabic (Mangara, 2017). In the specific ward where the research was conducted, the prevalence of Somali and Arabic is likely to be even higher. This means the local Feature Pool is likely to support diphthongal variants of the English diphthongs, given that a majority of the input languages and input accents have diphthongal variants of these diphthongs. If the situation of indirect language contact among children supports a diphthongal variant of PRICE, this would suggest that monophthongisation of PRICE is an age-graded feature, adopted in adolescence.

Chapter 9 will expand on this idea further, by considering whether the MLE diphthong onsets and monophthongisation of the diphthongs represent two distinct kinds of sociolinguistic variation. Building on Cheshire et al.'s (2011) consideration of "under the counter" and "off the shelf" features, we will link this distinction to the evidence of age-grading in the current results. It seems possible that in this community, the diphthong onsets, being supported by the local Feature Pool, are endogenous, "under the counter" features, and hence are acquired by children from peers at school. Meanwhile, monophthongal variants of the diphthongs, not having the same kind of support from the local Feature Pool, but being enregistered features of MLE, are adopted by adolescents in an "off the shelf" way. This discussion is postponed to Chapter 9 in order to discuss the findings from the diphthongs together with the observations from Chapters 4 and 8.

Finally, it is worth noting that the children showed the same preceding environment constraints as the adolescents on Trajectory Length variation in all three diphthongs. This means that although the children on average have more diphthongal realisations of at least two out of three of these diphthongs (PRICE and GOAT) than do the adolescents, they are nonetheless participating in the same sociolinguistic system to the extent that they show the same constraints on variation.

7.5.2 Why is this not the case for FACE?

It was noted in the previous chapter that the social factors do not seem to influence FACE variation among the adolescents. This finding is mirrored in the current chapter, in that the children seem to have acquired the same onset quality of FACE as the adolescents and do not show a more diphthongal realisation of FACE than do the adolescents.

One explanation as to why the children do not show a more diphthongal realisation of FACE than the adolescents could be the position of FACE in the vowel space. FACE has an onset that is near-front and close-mid. Given that its glide direction is forwards and upwards, there is limited space for this glide, and so limited scope for variation in the diphthong dynamics of FACE. As the children have a closer FACE onset than do the adolescents, it may be that it is not possible for them to show a more diphthongal realisation of FACE than the adolescents. The adolescent girls are able to have a more diphthongal realisation of FACE than the adolescent boys and the children because the adolescent girls' ONSET is more open – in other words, the fact that the adolescent girls have a relatively diphthongal FACE could just be a by-product of the height of the onset.

By contrast, both GOAT and PRICE have more room for movement in the diphthong. The onset of GOAT for the children varies between near-back, central, and near-front. The plots presented in Section 7.3 above suggested that (a) the height of GOAT is similar to FACE, and (b) that the direction of the glide varies – whereas FACE invariably has a forward glide. This means that there is space for GOAT to be more diphthongal than FACE. The onset of PRICE, meanwhile, is maximally open, and so there is room for it to be diphthongal in the children's speech.

More studies are needed of children's acquisition of monophthongisation changes in progress, and especially studies that control for the specific language backgrounds of children acquiring those changes. Without more data on children's acquisition of this kind of change, it is impossible to know whether in the current case, it is the local "Feature Pool" that leads to the children having more diphthongal variants of PRICE and GOAT than the adolescents, or whether it is to do with how diphthongs and monophthongisation of diphthongs are acquired.

7.5.3 Summary

This chapter has replicated the finding of Cheshire et al. (2011) that children aged 5–7 show the same diphthong system as adolescents in their community in terms of diphthong onsets. It also replicates Cheshire et al.'s finding that while there are gender differences in production among the adolescents, these are not found in the child data. However, the children and adolescents showed different diphthong dynamics compared to one another for GOAT and especially for PRICE.

Chapter 8

Epistemic phrases

8.1 Introduction

This chapter takes a different approach from the previous two. Chapters 6 and 7 presented a quantitative sociophonetic analysis of the diphthongs FACE, PRICE and GOAT in order to see to what extent the Ealing adolescents and children show the same kinds of variation in these diphthongs as have been attested in East London (Cheshire et al., 2011; Kerswill et al., 2008; S. Fox, 2015; Gates, 2019). The current chapter, rather than taking East London youth language as its starting point, is motivated by a discourse-pragmatic feature that was noticed during fieldwork in Ealing.

One of the research objectives for this project is to find out whether there are independent multiethnolectal language changes taking place in Ealing, potentially due to differences in the composition of Ealing's feature pool (Mufwene, 2003). Early on in fieldwork it became apparent that the youth centre attendees used the Arabic borrowing *wallah* (literally 'I swear to God'). This feature was of particular interest because while *wallah* has been attested in various European multiethnolects, it had not previously been found in London English. In this chapter, a form-based/derivational approach will be taken to *wallah* as, in line with other studies (Opsahl, 2009; Lehtonen, 2015), it will be analysed as part of a group of related phrases with the underlying semantic meaning of swearing the truth of something. Following Lehtonen (2015), these swear phrases are treated as grammaticalising epistemic phrases. After a qualitative analysis of the different functions of these features, I present a distributional analysis to suggest that different epistemic phrases are constrained (a) by interactional context and (b) by religious affiliation. Finally, some observations are made on the innovation-diffusion of these phrases.

8.2 Background

8.2.1 Discourse pragmatic features

The term "discourse-pragmatic features" is used here in the sense of "conventionalised, polyfunctional linguistic items and constructions" (H. Pichler, 2016b, p.1). Discourse-pragmatic features are not an empirically observable category – if anything, they are linked together by evading categorisation, as they "share neither a common set of formal linguistic properties nor an agreed upon macro-label" (H. Pichler, 2016b, p.3). Many in fact exist at the syntax-pragmatic interface (Cheshire, 2016). The term "discourse-pragmatic features" is therefore a heuristic which links together approaches to these heterogeneous features, and this is the term that is preferred in the current study.

One key feature that links discourse-pragmatic features together is their multifunctionality. These features may have several different available functions, potentially including their prescribed dictionary meaning. They may serve several different functions simultaneously. These functions can be pragmatic, information-structuring, discourse-structuring, or at the interactional and interpersonal level. Some examples include: intensification (Tagliamonte, 2016b); marking a direct quote (Levey, 2016; Tagliamonte & D'Arcy, 2009); gaining and/or holding the floor (H. Pichler, 2016c); marking information as new or old (Cheshire, 2005); topic marking (H. Pichler, 2016c); indexing speaker stance (Drager, 2016); indexing group membership (Drager, 2016); seeking listener response (H. Pichler, 2016c; Torgersen et al., 2011).

8.2.1.1 Approaches

How to analyse discourse-pragmatic features from a variationist perspective is a topic of much debate (see contributions in H. Pichler, 2016a), given that these features do not fit the traditional definition of the sociolinguistic variable, i.e. two ways of saying the same thing (Tagliamonte, 2012, p.2).

Waters (2016) delineates previous variationist studies of discourse-pragmatic features as taking one of three types of approach to the accountable analysis of these features: form-based, function-based, or position-based. Form-based approaches take as their starting point "either an individual lexical item or an underlying multi-word construction" (Waters, 2016, p.46). Function-based approaches begin with a function that can be served in the discourse; this is commonly the approach taken in the study of quotatives. Finally, analyses can begin from the position of the feature of interest in the utterance, e.g. at the left periphery, or as a sentence tag. This latter is the approach taken by Tagliamonte (2016b) in her work on "sentence starters".

At the same time, Cheshire (2016, p.265) has warned against shoe-horning discourse pragmatic features into being sociolinguistic variables, on the grounds that doing so almost invariably involves prioritising one function over others – "I would prefer to find a way of incorporating multifunctionality into an analysis rather than devising innovative ways of excluding it". Many discourse-pragmatic features, especially those that are grammaticalising, serve multiple functions simultaneously at different levels of the discourse, and in many instances there will be ambiguity as to which function prevails. This is necessarily difficult to incorporate into a variationist analysis.

8.2.1.2 Grammaticalisation

Grammaticalisation is defined as the process by which content words gradually take on grammatical functions, and how already grammatical items develop new grammatical functions (Hopper & Traugott, 2003, p.1; Heine & Song, 2011, p.590). But grammaticalisation is usually not understood as a process that applies only at the level of grammar: rather, grammaticalising forms can take on discourse-pragmatic functions that give them scope over wider areas of discourse than just the sentential context in which they occur (e.g. Cheshire, 2007, 2013). The meanings that grammaticalising forms have at different stages of the grammaticalisation process are not consecutive but rather, continue to co-exist alongside one another (Hopper & Traugott, 2003, p.3).

Heine and Song (2011, p.591) define the following "parameters" of grammaticalisation – intended as diagnostic tools for identifying instances where grammaticalisation is occurring: (1) extension, "i.e. when linguistic expressions are extended texts that invite the rise of grammatical functions (context-reinterpretation)"; (b) desemanticisation or semantic bleaching – when the original semantic meaning that a form had becomes less important to its use; (c) decategorialisation, "i.e. loss in morphosyntactic properties characteristic of lexical or other less grammaticalized forms"; and (d) erosion, i.e. phonetic reduction. In addition to these criteria, Cheshire (2007) uses co-occurring discourse markers to assess discourse-pragmatic function, as this is proposed by Aijmer (2002) to be a way in which speakers deal with the ambiguity of grammaticalising forms. Cheshire (2007) finds that the quantitative results of phonetic reduction and decategorisation support each other, i.e. if a form looks far grammaticalised in terms of phonetic reduction, it will also appear frequently with no discernible syntactic category. Tagliamonte (2016a) employs two measures of grammaticalisation for the case of general extenders: the syntagmatic length of the general extender, e.g. *and things* vs. *and things like that*; and co-occurrence with other discourse-pragmatic variables.

8.2.2 Epistemic phrases

Following Lehtonen (2015), swearing phrases such as *I swear* and *wallah* will be treated in the current analysis as epistemic phrases. This section will explain what epistemic phrases are, the importance of the intonation unit in determining the scope of epistemic phrases in use, and signs of the grammaticalisation of epistemic phrases.

8.2.2.1 Definition of epistemic phrases

Kärkkäinen (2003, p.1) defines epistemicity as comprising "linguistic forms that show the speakers' commitment to the status of the information that they are providing, most commonly their assessment of its reliability".

Speakers take an epistemic stance by showing their commitment to the status of the information they are providing. According to (Kärkkäinen, 2003), the expression of epistemic stance can be assessed at different levels of interaction: at the level of the linguistic form chosen; at the level of the intonational unit; and at the level of the turn-taking structure. Speakers show a tendency to mark their epistemic stance at the beginnings of intonation units (Kärkkäinen, 2003, p.4).

While in semantic research epistemics concerns the truth of propositions, this is not the most productive angle from which to approach epistemics in interaction (cf. Lehtonen, 2015, p.181). Kärkkäinen (2003, p.18) points out:

> The notion of truth of propositions has almost been a *sine qua non* in semantic research on modality, but it is not necessarily helpful for a more interactionally-based study such as the present one. My data do not show any clear orientation of the participants to the truth of what they are saying, but rather that they assess something as more or less reliable, or express their belief that such and such is the case, to name the two most common semantic meanings expressed in the database

Whether the proposition is true or not is not the primary issue: speakers are more concerned with showing how confident they feel in the truth or reliability of what they are telling. Similarly, Prieto and Borràs-Comes (2018, p.564) choose to define speaker epistemic commitment as encompassing expressions of both belief and evidentiality: speaker epistemic commitment is to be treated as "equivalent to speaker epistemic stance, which includes interrelated but separate notions of epistemicity (or speaker certainty or belief about the proposition expressed) and evidentiality (or the source or evidence that the speaker has to back up the proposition expressed)."

In a corpus of American speech, Kärkkäinen (2003, p.53) finds that the most common semantic meanings for expressions of epistemic stance in the data are, in descending order: reliability (of the information being expressed); belief (strength of the speaker's belief in what they are saying); hearsay evidence; mental construct e.g. *I imagine*, *I thought*; deduction; induction; sensory evidence. Reliability and belief are far and away the most common, followed by hearsay evidence. Kärkkäinen (2003, p.54) observes that within the categories of reliability and belief, speakers " tend to express *low* rather than *high* reliability, and a *weak* rather than *strong* belief, and thus generally express a low degree of confidence".

Expression of high confidence or reliability, and low confidence and reliability, are not relegated to separate epistemic phrases, but can be indexed by the same epistemic phrase, dependent on context. Kärkkäinen (2003, p.112) refers to Biber and Finegan (1989, p.110) who "observe [...] that in conversation certainty and doubt are sometimes expressed side by side". Moreover, Kärkkäinen (2003) finds that *I think*, too, can express either a strong conviction, or doubt – effectively operating at opposite ends of a continuum of epistemicity.

This will become relevant in the current analysis. Later on, it will be argued that *I swear* has grammaticalised far enough that it can be used either to do intensification and indicate strong conviction (the "lack of doubt/certainty") end of the continuum, or, perhaps more frequently, to show doubt.

8.2.2.2 The intonation unit (IU)

Kärkkäinen (2003, p.9) stresses the importance of the intonation unit (IU) as "the fundamental unit of discourse production." This is crucial in determining where the epistemic phrase occurs in relation to the information being provided in the utterance. For example, it means that although it may look like an epistemic phrase such as *I know* is used clause-medially, a closer analysis of the way in which the sentence is spread over intonation groups will reveal that the phrase actually appears at the beginning of a new IU (Kärkkäinen, 2003, p.32).

8.2.2.3 Grammaticalisation of epistemic phrases

Kärkkäinen (2003) also pays attention to the grammaticisation of epistemic phrases, particularly *I think*. Epistemic phrases can grammaticalise from acting primarily at the sentence level, to acting primarily at the discourse or interpersonal level. Kärkkäinen (2003) argues that this is occurring with phrases that include *say* e.g. *I said*. Signs of the grammaticalisation of forms such as *I think* into epistemic phrases include the absence of complementiser *that*, and phonological reduction. Traugott (1989) has argued that as epistemic forms grammaticalise, they become increasingly focused on the speaker's subjective beliefs: with the ongoing grammaticalisation process, linguistic forms gradually proceed from having less subjective to more subjective meaning. Kärkkäinen (2003, p.184) argues that this is the case with hearsay evidential epistemic phrases, while this process has already reached completion with *I think*.

8.2.3 Previous studies of *wallah*

This section will review studies that have attested the borrowing of *wallah* into various European majority languages.

Wallah is an Arabic construction comprising the particle *waw* 'by', *Allah*, and the genitive ending *-i* which may variably be dropped, so that it can be pronounced either *wallahi* or *wallah* (Al-Khawaldeh, 2018, p.115). Although *wallahi* is therefore the full form, and is reportedly more common than the *wallah* variant (at least in Jordanian Arabic; Al-Khawaldeh, 2018, p.115), in this chapter, *wallah* will be used to refer to both the *wallah* and *wallahi* forms. This is for the sake of continuity with other studies of European multiethnolects, which typically refer to *wallah* rather than *wallahi* – variably spelled as *wolla* (Opsahl, 2009) or *wallah* (Freywald et al., 2011; Quist, 2008).

In the data examples provided, participants variably use *wallah* and *wallahi*. The transcriptions make it clear which variant has been pronounced. It is beyond the scope of the current chapter to examine what conditions the variation between these two forms.

8.2.3.1 Attestations of *wallah*

The earliest study of multiethnolects in Europe, Kotsinas (1988), cited *wallah* as an example of an Arabic loanword that was used in Rinkeby Swedish. Since then, *wallah* has been attested in multiethnolectal speech in Danish (Quist, 2008), Norwegian (Opsahl, 2009; Svendsen & Røyneland, 2008), Finnish (Lehtonen, 2015), German (Freywald et al., 2011; Kallmeyer & Keim, 2003) and Dutch (Nortier, 2001). More recently, *wallahi* has been noted in Toronto English (Denis, 2019).

8.2.3.2 Style & shibboleth status

It is known that in Scandinavia, *wallah* has shibboleth status. Opsahl (2009) finds that there are tokens of *wallah* in speech from adolescents who are not from the multicultural Eastern part of Oslo, but that these tokens are instances of *wallah* being used metonymically to refer to the typical users of this phrase (Opsahl, 2009, p.226). However, Lehtonen (2015, p.195) states that *wallah* in Finnish does not have the kind of shibboleth status that it does in other Nordic countries.

Many of the aforementioned studies converge in seeing *wallah* and related phrases as markers of a style: "Conversations among adolescents in multiethnic areas in Oslo seem to be characterized by the use of a set of discourse markers which emphasize the truth value of utterances, thus contributing to an extended degree of epistemic focus" (Opsahl, 2009, p.221). Similarly, Lehtonen (2015) sees the use of *wallah* and Finnish counterparts *ma vannon* ('I swear') and *ma lupaan* ('I promise') as defining a style. Quist (2005, 2008) identified style clusters among students at a high school in Copenhagen, and one component of the "Cool" style cluster was the use of lexical items such as *wallah, jalla* (see also Karrebæk, 2015, p.24).

Opsahl relates the prevalence of epistemic phrases to the popularity of hip-hop, a culture in which there is an emphasis on authenticity: "I venture to say that a conversational style with emphasis on truthfulness and epistemic focus fits well with a culture where authenticity is a key concept" (Opsahl, 2009, p.239). Yet Lehtonen (2015, p.181) disagrees with Opsahl's assertion that a whole conversational style can be characterised by an emphasis on epistemicity, writing that she would not want to suggest that epistemicity should be more central to multiethnic youths' discussions than it is to other people's conversations. This issue will be returned to in the discussion at the end of the current chapter (Section 8.6).

8.2.3.3 Related phrases

Lehtonen (2015) relates the use of *wallah* and *wallahi* to the Finnish phrases *ma vannon* ('I swear') and *ma lupaan* ('I promise'). Lehtonen (2015) believes that the multilingualism of speakers and communities has precipitated grammatical changes in the Finnish phrases *ma vannon* ('I swear') and *ma lupaan* ('I promise') because of the influence of *wallah/wallahi*. Opsahl (2009, p.231) acknowledges that treating phrases such as those meaning 'I swear' as having developed from the borrowing of *wallah* is "speculative".

8.2.3.4 Distribution

Opsahl (2009, p.226) finds that boys are the chief users of *wallah*, accounting for 87% or 119 out of 137 tokens in a recent corpus of Oslo youth speech. Of these 119 tokens, 111 are from boys with two foreign-born parents. Meanwhile, in Helsinki, *wallah* is used mostly by Muslim kids, by boys more than by girls, and by those who have multiethnic friendship circles more than those who do not (Lehtonen, 2015, p.194).

8.2.3.5 Functions

So far, we have left aside the issue of what *wallah* actually does when used in interaction. Svendsen and Røyneland (2008, p.71) describe *wallah* as "intensifier and emphasizer". Similarly, Quist (2008, pp.47–48) states that *wallah* can be spoken with rising intonation to mean 'Is it true?', or can be used "as an intensifier to underline the importance or value of a statement." Opsahl (2009) treats *wallah* as a discourse marker. Her criteria for classifying some of its uses this way are that the constructions analysed "do not enter into a fixed expression, serve as an adjective or take on an obligatory syntactic function" (p.233) – i.e. *wallah* is becoming bleached of semantic content and appears in positions that are peripheral to the sentence syntax. When used in argumentative contexts, "the emphasizer *wolla* seems to upgrade an assertion or an assessment (cf. Pomerantz 1984:65), and appears to be an efficient verbal device for winning an argument" (Opsahl, 2009, p.228). (Incidentally, Lehtonen (2015, p.176) also notes *wallah* appearing in argumentative contexts.) Of *si wolla* ('say wallah'), Opsahl (2009, p.229) states that "It seems in many of the cases to be an automatic minimal response rather than an actual request for the performance of a specific speech act ('to swear by God')."

Lehtonen (2015) approaches *wallah* somewhat differently, by looking at its relationship to the whole interaction. Lehtonen (2015) treats *wallah* as marking a particular storytelling ritual. According to Lehtonen (2015), the scope of *wallah* is not just the clause it occurs in; rather, the use of this phrase indexes the speaker's stance towards the entire interaction. Lehtonen (2015) sees *wallah* and associated phrases as sign-posting a "sensational news" genre (p.183). This genre is partly a collaborative creation: other participants must ratify the story as being newsworthy. Lehtonen specifies that there are two worlds that are relevant at the moment of sensational news telling: the world in which the animator (Goffman, 1981) of the story does the telling; and the storyworld – relevant concepts are *figure* (Schiffrin, 1990; Georgakopoulou, 2007) and *telling* vs. *taleworld* (Georgakopoulou, 2007, 2008). The narrative will often be preceded by an introductory sequence (Routarinne, 1997) in which the narrator offers a story and the listeners promise to give up the floor for the duration of the telling. Yet the narrative is constructed in cooperation with the listeners. Lehtonen asserts that in her data, there is often an introductory sequence in which the narrator offers a story preface, and the listeners react. Lehtonen (2015, p.187) argues that *wallah* or equivalent often appears at the onset of sensational news telling, and signals the narrator's stances towards the entire story, justifying the novelty and tellability of the events to be told, as well as communicating the narrator's responsibility for the story.

Lehtonen (2015) also describes 'say wallah' sequences as a form of ritual, after Rampton (2006, pp.174–175). Rituals involve a suspension of ordinary interaction, as they incorporate "traditional material" in "relatively rigid patterns". Ritual generates a sense of collectivity among participants and involves an element of performance.

Finally, in the sensational news genre, it is not necessary for the story to be true – it is more important for the story to be worth telling (Lehtonen, personal communication).

8.2.3.6 Grammaticalisation of *wallah*

Opsahl (2009) raises the question of the grammaticalisation of *wallah*. She presents evidence that *wallah* is not treated by adolescents as a literal swear, and that this has been overtaken by its function as emphasiser: "there seems to have been a typical loss of some of its original semantic content paralleled by a pragmatic strengthening" (Opsahl, 2009, p.230). Opsahl also notes increased flexibility of position, i.e. movement from the left periphery to utterance-final position or status as an independent utterance, as signs of grammaticalisation (p.237), and phonological reduction (p.237). She proposes a cline of grammaticalisation as follows: (1) an independent syntactic unit with propositional content; (2) "fixed expression with propositional content and relatively free placement options"; (3) "discourse marker with weakened propositional content and fixed placement, typically utterance-initial"; (4) "morpho-phonological reduction"; (5) "Renewal by addition of reinforcing elements with propositional content (*wolla Koran* 'wolla Quran') or by combining several discourse markers (*wolla jeg sverger* 'wolla I swear', *wolla helt ærlig* 'wolla quite honestly')".

8.2.3.7 Co-occurrence

In both Danish and Norwegian, *wallah* can co-occur with *Koran*; this latter is not a noun denoting the Islamic religious text, but rather, functions in a similar way to *wallah*, giving emphasis to an utterance. (Swearing by the Q'uran is also possible in Jordanian Arabic; Almutlaq, 2013.) Opsahl (2009, p.230) gives the example *Wolla de er farlig wolla Koran' (.) e heftig farlig (.) jævla farlig* ('Wolla they are dangerous wolla Quran' (.) e severely dangerous (.) fuckin' dangerous'). Karrebæk (2015, p.35) describes how in Danish, *Koran* "is used to emphasize the truth-value of the proposition" and is a shibboleth, alongside *wallah*, of urban youth speech. Opsahl (2009, p.237) calls phrases meaning 'I swear', 'quite honestly' etc "combinatory or alternative markers". Quist (2008, p.48) notes that *wallah* "often appears in fixed phrases like *wallah jeg sværger* (English: 'wallah I swear') and *wallah billa* (= Arabic, *billa* is a contraction of *bi ism Allah*, meaning in the name of God)."

8.2.4 Theoretical perspective and approach

Following from what we have learned in the previous subsections, the following approach will be taken to analysing *wallah* in the current analysis. Taking a form-based approach to this discourse-pragmatic feature (Waters, 2016), and following from previous studies of *wallah*, phrases with an underlying semantic meaning of *I swear* will be included in the analysis. The first person subject does not need to be present and nor does the swearing verb – for example, *on [my/your] mother's/mum's life* will be included, which sometimes appears with a first-person subject and the verb *swear*, i.e. *I swear on my mum's life*, but can also appear in the form *on my mum's life* or *mum's life*. Examples from the current data will be transcribed in a way that chiefly adheres to Conversation Analytic transcription conventions, but that also pays attention to Intonational Units (IUs) (Kärkkäinen, 2003, p.9). These phrases will primarily be treated as epistemic phrases (Lehtonen, 2015; Kärkkäinen, 2003), and there will be a focus on whether they show high or low speaker commitment. At the same time, attention will be paid to the discourse-structuring and interpersonal functions of these phrases, as have been described by previous research (Opsahl, 2009; Lehtonen, 2015), and as such they will be treated as multifunctional discourse-pragmatic features, having functions at several levels of the discourse simultaneously.

In the analysis of the functions that different epistemic phrases have in the current data, comparisons will be made between the current, Ealing data and older data from Hackney (Cheshire et al., 2011). This allows us to see which epistemic phrases seem to be widespread in London, and which ones have not been described before (although care will also be taken to avoid the "recency fallacy" (G. Andersen, 2016, p.40)). It also allows us to compare the functions that the different epistemic phrases have served at two different points in time, in two different boroughs of London. Some of the epistemic phrases to be examined in the present chapter were also found in the MLE corpus from Cheshire et al. (2011) and these are listed with accompanying examples in Table 8.1. Comparisons will also be made between the current data and Ilbury's (2020) recent data from Hackney – the same fieldsite as the MLE project (Cheshire et al., 2011).

Table 8.1: Examples of epistemic phrases found in the MLE corpus

Variant	Example
I swear	Najib: yeah. when you read it yourself as well **I swear** it sounds really nice
swear?	Dexter: have you had that done? Aimee: yeah . Dexter: **swear?**
on X life	Stacey: no it ain't **I swear on my mum's life**
swear down	Dexter: I wanted to I don't care <Aimee laughs> I I wanted to jump on her . **I swear down**
swear to god	Omar: I was just walking **I swear to god** I was just going to see my grandma ...

8.3 Forms and functions of the different epistemic phrases

This section will provide examples of the epistemic phrases that are in use and examine their function and their position – e.g. whether they can stand alone, whether they must appear at the left periphery or at the right periphery of an utterance, what intonation contours they host.

8.3.1 *I swear*

We begin with *I swear*. Just as *I think* can be used to show either a high degree of speaker confidence, or at the opposite extreme, to show doubt, so too can *I swear* in the current data (Kärkkäinen, 2003).

8.3.1.1 *I swear*: Showing high commitment

In Example 72, Sqara has told a story about how, during the Syrian uprising, he was chased and someone shot at him. Speaker commitment is directly challenged in this instance: the interviewer observes that Sqara is smiling (and his laughter is audible in line 2) and asks one of his friends if Sqara is telling the truth. The interviewer thus expresses doubt about the truthfulness of Sqara's story. Sqara's two friends, Karim and Ahmed, both immediately answer in the affirmative, and Sqara uses *I swear* apparently to convince the interviewer of the truth of his story. In this way, *I swear* is used to index speaker commitment at a moment when speaker commitment is being challenged. *I swear* is used in conjunction with repetition of the phrase "it's serious".

(72) Sqara: I swear it's serious

1	RO	why were they chasing you though?
		(0.48)
2	Sqara	trying take out my l(h)ife hu hu
		(0.34)
3	RO	he's smiling is he te- is this a real story?
		(0.23)
4	Ahmed	[yeah] it is
5	Karim	[yeah]
6	Sqara	[**I swear** it's serious]
7	Ahmed	[yeah I'm just] laughing cos:
		(0.28)
8	Sqara	is a serious story.

Comparable usage of *I swear* is also found in the MLE corpus. In Example 73, there appears to be an issue between Dumaka and Uzay, where Uzay is maintaining one version of what Dumaka has just done, and Dumaka is refuting Uzay's account as being not true. In line 3, *I swear* co-occurs with repetition of *I didn't*. In this respect, *I swear* appears to be used to contribute intensification or emphasis (cf. Opsahl 2009). Just like in the previous example, *I swear* is used in a context where one participant, Uzay, shows doubt in Dumaka's story, and Dumaka uses *I swear* at the same time as he repeats the proposition "I didn't say".

(73) MLE corpus: I swear

1	Dumaka:	#1 no xxx xxx I didn't say . I didn't I didn't say #
2	Uzay:	#2 [Arfaan: ah ah ah] ¡laughs¿ and [Arfaan: whoa whao whoa] ey xxxxxx he say like this to #
3	Dumaka:	I didn't I didn't I /didn't **I swear/**
4	Arfaan:	[Uzay: uh uh] okay okay
5	Uzay:	and he's doing like this .
6	Dumaka:	I didn't ¡laughter¿ . how could I do that? .. liar liar pants on fire

8.3.1.2 *I swear*: Showing low commitment

Yet *I swear* can also be used to index a low degree of speaker belief, or the unreliability of knowledge. This can be seen in Examples 74, 75 and 76, which all show the phrase *I swear* being used to indicate speaker uncertainty. This is in fact the more commonly served function of *I swear* in the current data.

In Example 74, CB has just seen a friend, who he believed had been banned from attending the youth centre, enter the centre. *I swear* co-occurs with *oh* and *shit*: this utterance expresses surprise. The utterance as a whole can be taken as expressing doubt about the speaker's own prior knowledge, i.e. that the friend was banned, which does not fit with the current evidence, i.e. that the friend is attending the youth centre. In sum, *I swear* indicates incongruence between epistemicity and evidentiality: the speaker had a high degree of belief in one state of affairs, but new evidence suggests this knowledge to be unreliable. In this respect, *I swear* can be seen as prefacing a confirmation-seeking question (Prieto & Borràs-Comes, 2018).

Example 75 also shows *I swear* being treated as prefacing a confirmation-seeking question. After Lucy has uttered line 2, Jessica's response offers Jessica's own knowledge on the matter. Lucy has been explaining how Islam's teachings are contradictory to her own religious faith, Catholicism. Lucy utters the second IU in line 1 with a high-rising terminal (Levon, 2016; Fletcher & Harrington, 2001; Britain, 1992). In line 1, Lucy presents information that may be new to her recipients. Lucy expresses a lower degree of belief in her next statement, “Muslims believe in more than one god”: she utters this statement with a slight rise/continuing intonation as opposed to final intonation; after a pause, she follows it with *I think*.

Finally, in Example 76, *I swear* co-occurs with *innit*, which in this instance appears to be used as a canonical neg-tag (H. Pichler, 2016c). That Khadir’s proposition “they moved” contradicts knowledge that has been proferred prior to this in the discourse is indicated by the particle “though”. There is a micropause inbetween *though* and *innit*, such that *innit* can be seen as a turn increment (Kärkkäinen, 2003, p.30). The preferred response is confirmation, yet there is initially a pause of more than half a second, then Ali replies “nah” (l.2) which is not sufficiently informative in terms of Gricean politeness. Khadir reformulates the first pair part, which again includes the base proposition “they moved”, but this time, there are two items in the left periphery, *I thought* and the hearsay evidential *you said*, and on the proposition “they moved” there is a rising contour.

(74) CB: I swear you’re banned

1	CB	oh shit Sami >>**I ↑swear** you’re << ↓banned,

(75) Lucy: I swear Muslims believe in more than one god

1	Lucy	like with like the ten commandments and stuff, (.)
		it says that you should only be↑lieve in one god?
		(0.31)
2	Lucy	>>**I swear**<< muslims believe in more than one god,
		(0.81)
3	Lucy	[>>°I think° <<]
4	Jessica	[they ha]ve like more prophets and

(76) Khadir: I swear they moved though, innit

1	Khadir	>>**I swear**<< they moved >>though (.) ↑innit<<¿
		(0.71)
2	Ali	nah.
		(0.18)
3	Khadir	>>I thought you said<< they moved?

This use of *I swear*, i.e. to express doubt and seek confirmation, is also found in the MLE corpus. In Example 77, Roshan is trying to guess the age of the interviewer, Sue. In line 1, in a similar fashion to Example 76, a proposition with *innit* is used. *I swear* is then used together with the hearsay evidential *you told* in the utterance “I swear you already told me how old you was". When Sue responds to the hearsay evidential, and not to the proposition "you're like fifty", Roshan re-formulates this confirmation-seeking question as an information seeking question, "did you say you was about fifty?", and then resorts to a directive, "ah just tell me the age group".

(77) MLE corpus: I swear ...

1	Roshan:	I bet you're something . no you're like fifty innit <Robert laughs> **I swear** you told me already how old you was .
2	Sue:	did I?
3	Roshan:	did you say you was about fifty? .. ah just tell me the age group ..

8.3.2 *(I) swear down*

8.3.2.1 *(I) swear down*: Showing high commitment

Both in the current corpus and in the MLE corpus, *swear down* can be used to show a high level of speaker commitment, i.e. high epistemicity.

Of the tokens of *swear down* in the current data that are associated with high speaker commitment, most of these involve the speaker managing others' perception of him/herself, or saving face in some way. In Example 78, Tariq, Sami and the interviewer are discussing men's and women's roles in marriage. Tariq is apparently managing others' perception of himself and trying to convince others that he is not sexist. *Swear down* here co-occurs with two other devices associated with emphasis: increased loudness in l.11, and repetition (of the statement "I'm not being sexist").

(78) I'm not being sexist

1	Tariq	I'm not being <u>sex</u>ist or a(h)ny hh
2	Tariq	**I SW(h)(h)EAR D(h)Own** .hh
		(0.15)
3	Tariq	I'm not being sexist,

In the MLE corpus, *swear down* frequently appears in what might be described as "sensational news" stories (Lehtonen, 2015). In Example 79, *swear down* (l.3) co-occurs with *I'm serious* (l.1) in the same narrative. He exaggerates, e.g. "untold units" and "I just carried on going... just carried on going on and on". In Example 80, the story is not exactly "sensational news" but again involves exaggeration. In this instance, Roshan uses intensification devices in a story about how boring the village was, and how many push ups he had to do to stave off the boredom. In Example 81, *I swear down* occurs twice in a story about "the worst arrest" William ever had.

(79) MLE corpus: I swear down

1 William: I don’t remember . I’m serious. after hitting my head
I couldn’t remember anything
2 Sue: were you drinking then?
3 William: yep [Sue: yeah] untold units going through my body.
I was drink. **I swear down** I started with thirteen shots at the start.
and i just carried on going <Drew sighs> just carried on going on
and on .

(80) MLE corpus: swear down (A)

1 Roshan: just boring . was like a flipping village where I am blud
<Robert laughs> ..
swear down it was . over dead .. me and my brother yeah
do you know what we used to do?
we used to do push ups to get by <Sue laughs>..
swear down imagine that? .
we was doing like a hundred push ups for like two weeks ..
hundred push ups a day

(81) MLE corpus: swear down (B)

1 Sue: you’ve been arrested have you? [William: huh <laughs>]
how many times?
2 William: eight
3 Sue: eight times
4 William: #1 but they on. only warnings they give me though they
arrested me and give me warnings .
what’s the point of arresting me? . you could just
give me a warning on the spot but then oh **I swear down**
I think the worst arrest was when I got chucked into the cage .
they chucked me . the cage door wasn’t even open.
they just chucked me into the cage
I smashed my face open and everything .
and I was like dizzy . on the stairs .
they just “get up” they booted me to try
and get me up . and I think one of my ribs broke .
so they chucked me in there
I was on the floor /face/ #
5 Roshan: #2 /your ribs/ didn’t break blud your ribs were cool #

6	William:	no bruv **I swear down** cos like there's a indent in my ribs now . here . and they got me up. I was in the police station face bleeding and everything I'm thinking "fuck this for a . clean me up before my mum and dad come"

8.3.2.2 *(I) swear down*: a news-marking response token

Many instances of *swear down* in the current data could be described as response tokens (M. McCarthy, 2003; Schegloff, 1982), because in these instances the phrase seems to be used in a way analogous to *really*. M. McCarthy (2003, p.42) has described *oh* and *really* as news-marking response tokens and in particular, says of *really* that it "is of special interest in that it invites continuation by the previous speaker, or at least some indication of confirmation before the talk can continue, and before the full affective reaction occurs" (p.51). In this section I will show how *swear down* is used by participants to acknowledge new information and to invite continuation. *Swear down* often gets used when the interlocutor is in a position of greater knowledge than the speaker.

In Example 82, a group of boys have been talking about creative locations for shooting a music video. Raphael brings up a place that he apparently heard Ben and another friend talking about as a potentially relevant location for shooting a video. However, Raphael cannot remember exactly what this place was called, and refers to it as "that ting" in an information-seeking question in line 1, then as "West [place]" in line 4. Ben recalls the place in question, laughs and says that they can't go there. In the pause between lines 13 and 14, there may be some gaze or gesture between Kai and Ben that leads Ben to begin a lengthier explanation in l.14; Kai self-selects as respondent by back-channelling in l.15, suggesting that Ben has been addressing this explanation specifically to Kai. When Ben pauses at the end of l.14, Kai expresses interest by asking "where", and Ben overlaps with Kai in his reply. At this point, l.18, Kai uses *swear down* for the first time.

Both times that Kai utters *swear down*, it is in rapid speech and with question intonation. The fact that it stands alone, the with no subject adjacent to the verb, and that it is spoken so rapidly, speak in favour of it being a new phase of grammaticalisation that follows on from the high commitment function that we saw in the previous section. This use of *swear down* may have grammaticalised from an intermediate function that would challenge the epistemic status of another's utterance, i.e. *You swear down?*, as there is still a residual meaning of "Is that true?". We will see examples similar to this hypothetical intermediate stage later on, in the context of the *say* phrases.

Challenging epistemic status, however, is not prioritised as a meaning of *swear down* in this instance. Rather, it appears to be marking new information in what Ben has said

– in the second instance, at l.25, *swear down* co-occurs with the change-of-state token *oh* (Heritage, 1984). Both times, Ben treats Kai's use of *swear down* as an invitation to keep the floor and carry on telling. At line 19, Ben acknowledges the preceding turn with *yeah* (Tagliamonte, 2016b). He then continues "you see where um:", i.e. giving Kai more information to help him locate this house within the local area. Kai's "swear down" appears not interpreted by Ben as casting doubt on Ben's previous turn, but as a request for more information and/or an invitation to continue his telling. Indeed, Ben carries on supplying information that is relevant to Kai's direct question "where" in line 16. Both instances of *swear down* occur at a transition relevance place (TRP) and appear to act as an invitation to carry on holding the floor. Both instances occur after a brief but maximally informative and discourse-new utterance from Ben: "west [place]" in the first instance, and "right next to it" in the second.

(82) Some freemason shit

01	Raphael	((t)) what was that ting that you and {name} were talking about though, the one in u:m (0.19)
02	Raphael	((t))
03	Raphael	u::mm
		(0.71)
04	Raphael	West {place}
		(1.05)
05	Ben	what's in West {place}.
		(0.93)
06	Ben	((t)) oh nah, we can't go in there bruv: hh ts:
		(0.58)
07	Ben	#(are you) NUTs(h)s .hh hu hu
		(0.29)
08	Ben	that's some freeMAson shit
09	Ben	hhn hhn
10	Ben	w(h)e c(h)an't go in th(h)ere(h) .hh
11	Ben	and shoot a video f(h)am
12	Ben	hhn they will kill us
13	Ben	ha hu he he
		(1.03)
14	Ben	er men found some [freemason house] like,
15	Kai	[phhhh]
16	Kai	[where]
17	Ben	[west] {PLACE} fam

18	Kai	>>**swear down**?<<
19	Ben	yeah you see where um:
		(0.64)
20	Ben	((t)) {name},
21	Ben	st- [st-] um
22	Kai	[mm]
23	Ben	(the) youth club is¿
		(0.58)
24	Ben	ri::ght next to it fam
		(0.27)
25	Kai	>>↓oh **swear down**?<<
26	Ben	there's one –
27	Ben	there's one TING (as in like) (.) the HOUse is written in LA- it's got LAtin WRIting [on it]
28	Kai	[yeah]
		(0.57)
29	Ben	yeah
30	Kai	and w- and w- and,
		(0.43)
31	Kai	and wha's the WRIting innit wha's it say
		(0.55)
32	Ben	ah, for the god who will destroy us

Example 83 is slightly different in that the speaker who uses *swear down* is addressing an uninformative recipient. Again, there is a knowledge differential: Sqara has had his hair cut, and Omar and Ibrahim want to know which barber he went to. Omar and Ibrahim have to pose their questions several times before Sqara tells them the desired information, i.e. where he got his hair cut. But again, like in Example 82, *swear down* is used immediately after the requested information has been given: Omar asks where Sqara got his hair cut in lines 4, 5 and 12, and Sqara gives the relevant answer in line 13. Omar responds with *swear down* in line 15. And again, the use of *swear down* co-occurs with *oh* (line 15). Omar's use of *swear down* is not treated by Sqara as an invitation to tell more. This could possibly be because Omar's use of *swear down* does not have the question intonation that Kai uses in Example 82. Omar follows his turn in line 15 with an information-seeking question; he does not leave any pause for Sqara to respond, so it may be that he also perceives his use of *swear down* as not mandating any explicit response.

(83) A fresh cut

01	Omar	# ah is that a fresh cut and that ¿#

		(0.09)
02	Omar	normal
03	Sqara	ten pound yeah [hu hu hu hu hh?]
04	Omar	[(dan no). from where,] eagle's?
		(0.47)
05	Omar	where d'you [get it]
06	Ibrahim	[get it from] legal's))?
		(0.20)
07	Omar	# man said [legal's] #
08	Ibrahim	[**say w-**]
		(0.13)
09	Omar	a [hu hh]
10	Ibrahim	[**say wallah** that] was ten pound from eagle's?
		(0.36)
11	Sqara	no
		(0.11)
12	Omar	[where d'you get it from]
13	Sqara	[it was hassan]
		(0.19)
14	Ibrahim	[((you said)) –
15	Omar	[o:h, >>**swear down**<<]
16	Omar	how much
17	Ibrahim	>>**swear** you just said ten [pound]<<
18	Sqara	[thirteen] pound
		(0.37)
19	Omar	f- oh cool
20	Ibrahim	thirteen pound?
		(1.18)
21	Ibrahim	looks hard still.

No comparable tokens of *swear down* were found in the MLE corpus. In that corpus, *swear down* mostly occurs with the 1st-person pronoun *I*, and is generally in a declarative/narrative context. It is hard to be certain without listening to the recordings. The closest thing to *swear down* as described in Examples 82 and 83 is the use of *swear* with question intonation in the MLE corpus. In Example 84, Dexter appears to either acknowledge new information and/or request confirmation, by using *swear* with question intonation. Aimee's response is affirmative, "yeah", similar to Ben's responses in Example 82. Similarly, in Example 85, Dexter and Aimee disagree as to when an event occurred, and when Dexter uses *swear* in l.3, Aimee responds by giving a clarification.

(84) Dexter: swear?

1	Aimee:	the woman helped her
2	Dexter:	**swear?** .
3	Aimee:	yeah

(85) Dexter: swear

1	Dexter:	that was time ago though
2	Aimee:	because . no it wasn't you know
3	Dexter:	**swear**
4	Aimee:	it was like last year .

8.3.3 *Wallah*

8.3.3.1 Showing high commitment

Wallah has been described by Opsahl (2009); Quist (2008) as intensifying/emphasising. In this respect, *wallah(i)* may be understood in Norwegian and Danish as indexing high commitment by the speaker to their utterance. This function of *wallah* also exists in the current data. Example 86 shows *wallah* being used when the truth status of an utterance is challenged: Lola offers a story preface in l.1–3 and Khadir challenges the truth of this story in a "bald on record" way by saying "you're lying" (l.5). Lola repeats her utterance from l.3 verbatim, but prefaces it with *wallah*. Given that the only lexical difference between l.3 and l.6 is *wallah*, *wallah* may be seen as indexing epistemic stance and upgrading the truth status of the proposition in l.6.

(86) Lola: *Wallah* I slipped

01	Lola	oh ( ) y'know yesterday yeah,
		(0.25)
02	Lola	when (.) I was walking home >>it ws<< raining so much,
		(0.15)
03	Lola	#I slipped, and I f(h)ell on the floor#
04	Ali	[hh hh hh hh] [.hh]
05	Khadir	[you're lying]
06	Lola	**[wa]llah** [I]slipped and I fell on the floor

8.3.3.2 Sensational news stories

Wallah is also known in other languages to serve interpersonal and discourse-structuring functions. This section gives examples of *wallah(i)* being used in sensational news stories in the English data.

Lehtonen (2015) describes sensational news stories as being collaborative events: the events of the story will be highly implausible, and recipients must agree to suspend disbelief and give the speaker the floor for the duration of the telling. Lehtonen (2015) gives examples of narratives that appear to be structured by *I swear* phrases, used at particular moments in the telling as stand-alone IPs. Lehtonen (2015, p.187) states that the use of these *I swear* phrases maintain the narrator's stance towards the story of the whole narrative, and justify the newsworthiness and tellability of the events. Lehtonen (2015) suggests that this means that at another level, *wallah(i)* acts as an index of genre, i.e. of the sensational news genre.

In the following extract, the interviewer has put the question to Khadir, Ali, Amanda and Lola: "Have you ever been stopped by the police?" and Khadir has told a story about one of his experiences. Ali has then related to the interviewer and to his friends how he once got arrested for climbing the scaffolding on the local town hall. As the extract begins, he says that he was lucky that he was young enough not to receive a criminal charge.

Ali's utterances of *wallahi* show the overlap between *wallah* as discourse- and narrative-structuring, and *wallah* as used when the epistemic status of what the speaker has just said has been challenged (Opsahl, 2009). The telling of Ali's story is both collaborative and combative at the same time. To put this story in context, Khadir, Amanda and Lola (who is also present) have spent much of the interview teasing Ali, and he complies with the role they assign him. In an earlier narrative, the friends recount how when Ali first moved to the area, other kids stole his sliders from his feet. In the current narrative, Ali presents himself in a more serious light: he didn't just get stopped and searched, he got arrested; it was for trespassing on the town hall; at that age, he used to like climbing buildings and scaffoldings. Amanda deflates some of the drama of the narrative by revealing that while Ali got arrested, his friend escaped (l.4). Ali needs to both hold the floor and complete telling his story, and also maintain face against his friends' alternative version of events.

The moments at which Ali uses *wallah* are at moments of high drama, and also at moments when Ali's control over the floor and over his narration are threatened. At l.04, Amanda says "Diego got away though" with surpressed laughter audible in her voice. Ali tries to claim in l.07–09 that Diego only got away because Ali had saved him, but Amanda again turns this into a tease, saying "so you let him go and let yourself get bagged?" It is at this point, when Ali's narrative and his portrayal of himself are under threat again, that he uses *wallah(i)*. Lehtonen (2015) has described how in the telling of "sensational news stories", two worlds are important: the world of the telling, in which the speaker is animator, author and principal simultaneously (Goffman, 1981); and the story-world, in which the speaker is usually the main character. Ali's status at this point both as animator and author of his story, and as central character in that narrative, are being threatened. Ali's use of *wallah(i)* at l.15 can be seen as "attention getting" (Tagliamonte, 2016b) and also as intensifying (cf. Opsahl, 2009) in so far as it marks a moment of high drama in the story: events had become so scary that Diego was crying. This means that *wallah* is also showing high speaker commitment and attesting to the reliability of the speaker's version of events: Amanda has hinted at an alternative story in which Diego manages to get away, but Ali is not quick enough to escape; Ali's account, that actually, Diego was scared, contradicts his friends' version of events and *wallah(i)* perhaps prepares his recipients for an implausible turn in the story. This is indeed in line with Lehtonen's (2015) description of *wallah(i)*.

After his first use of *wallah(i)*, Ali is able to carry on his story uninterrupted for the duration of l.15–31. At l.31 there is a potential TRP and Amanda and Khadir both laugh – general extenders can signal the end of a turn (Cheshire, 2007). Ali again says *wallah(i)* and continues on the same topic. Similarly, at l.39, Ali reveals that he was caught by a dog, and this appears to be met by disbelief by Khadir in l.30, and as his Khadir starts laughing (l.41), Ali again says *wallah* before continuing. Both of these latter two instances of *wallah* are dialogic: they appear at moments when other speakers might be about to take the floor, and signal that there may be more to come of the story; they highlight particular moments in the story as being particularly dramatic, while also looking ahead to recipients' potential disbelief in these moments. Khadir's responses in l.33 and l.40–41 are also in line with Lehtonen's (2015) claim that *wallah(i)* indexes a narrative genre: Khadir's response weakly expresses disbelief in l.40, but his laughter and Amanda's in l.32–33 show affiliation with the telling. It seems that what is wanted by the group of friends is a sensational story, and Ali's *wallah* indicates to the others that he is about to provide sensational news.

(87) I got caught by a dog

01	Ali	but I was YOUng so I was LUcky
		(0.89)
02	Rosie	[yeah]
03	Ali	[I was]n't even 'fifteen I was like 'fourteen
		(1.21)
04	Amanda	# Diego got away though #
		(0.24)
05	Ali	mmyeah
		(0.91)
06	Amanda	hhh[hh]
07	Ali	[if it was]n't for ME he got

08	Ali	. he
		(0.23)
09	Ali	he would NOT ’ve got away with it
		(0.76)
10	Amanda	# so you let him go and let yourself get bagged? #
		(0.58)
11	Amanda	[fth]hh:
12	Ali	[y:eah]
13	Ali	that’s no
14	Ali	c-
		(0.26)
15	Ali	[cos DIEgo] **wallA(hi)** was cry-
16	Amanda	[((claps))]
		(0.25)
17	Ali	he was STRESSing was SO DEEP
		(0.56)
18	Ali	so – so they were –
19	Ali	the police (were) like is there anyone else inside?
		(0.30)
20	Ali	and Diego musta been the:re with: Tom
		(0.42)
21	Ali	I said nah nah the th- –
		(0.26)
22	Ali	I >>ws like<< ’s only: (.)
23	Ali	cos another guy got
24	Ali	(w) e:r
25	Ali	caught with me
		(0.41)
26	Ali	I was like ’s ONly me NO ONE else (0.09)
27	Ali	there
		(0.15)
28	Ali	and I was
		(0.40)
29	Ali	then they
		(0.19)
30	Ali	they CHECKed but they c-
		(0.32)
31	Ali	they made it a ↑big ↑thing they got HElicopters and EVerything

		(0.29)
32	Amanda	hhhhh [.hh]
33	Khadir	[hh .hh]
34	Ali	**[wa]llAH(i)**
		(0.20)
35	Ali	[LOOKing for] them
36	Khadir	[hh]
		(0.59)
37	Ali	and DOGs and (that)
38	Ali	that’s why I got caught by
		(0.48)
39	Ali	[I got] caught by a ↑dog [yeah]
40	Khadir	[a dog?]
41	Khadir	[hhh] hu hu
42	Ali	# **wallAH** #
		(0.79)
43	Ali	but ME I HATE dogs ((door opens)) so I STOPPed
		(0.13)
44	Ali	I couldn’t even carry on running

Analysing *wallah* as indexing a narrative genre in this way also gives us a different perspective on Example 86. Example 86 is repeated and expanded as 88. Previously, we analysed Khadir’s utterance in l.5 as a challenge to the epistemic status of what Lola is saying, but while this is indeed the referential content of the utterance, this seems an unlikely interpretation of its function: why would Lola lie about having fallen over on the way home? It seems more likely that Khadir’s utterance “you’re lying” (l.5) constitutes a challenge to Lola to tell a better story. If this is the case, her use of *wallah* in l.6 can be seen as projecting a more sensational version of events. This indeed appears to be what she does in l.6–9, saying she was in shock, and adding emphasis on the /k/ in *shock* and on the intensifier *so* in l.9.

(88) *Wallah* I slipped (B)

01	Lola	oh ( ) y’know yesterday ˜yeah˜,
		(0.25)
02	Lola	when (.) I was walking home >>it ws<< raining so much,
		(0.15)
03	Lola	#I slipped, and I f(h)ell on the floor#
04	Ali	[hh hh hh hh] .hh]
05	Khadir	[you’re lying]

06	Lola	**wa]llah** I] slipped and [I fell] on the [floor]
07	Ali	[hhh] (.) [hhh]
		(0.16)
08	Lola	’n’ I was in shocK
		(0.44)
09	Lola	an’ it was raining <u>so</u> much

8.3.4 *On X life*

8.3.4.1 Showing high commitment

In the current data, *on my mum’s life/mother’s life* can show high speaker commitment. It gets used to contribute force to threats, as in Example 89. *On X life* also gets used when one party’s past or future actions are at issue. For example, in Example 90, Chantelle uses *mother’s life* when attempting to persuade a youth worker that someone has cheated at pool by picking up the white ball.

(89) CB: **on my mum’s life** Ima fuck you up

(90) Chantelle: **mother’s life** he did, he picked up the white ball (ennit)

This use of *on X life* is also found in the MLE corpus. In Example 91, for example, *on my life* co-occurs with “I’m not even lying to you”, indicating that the speaker is showing high commitment, and emphasising the truth of what she is saying.

(91) Zaida: ... it was like a machete I’m not even lying to you **I swear on my life** it was like a machete

8.3.5 *I swear to God*

It should be mentioned that *I swear to God*, which would be the literal translation of *wallah(i)* into English, appears in both the current data and in the MLE corpus, although it is not frequent in either – there are only two tokens in the current data. Intriguingly, the most frequent user in the MLE corpus (contributing 3 out of 5 tokens) comes from a Somali teenager. In his interview, he and his friend have been talking about how they are afraid of going to Camden because of the “Somalian boys up there”. Omar tells a story about how he was once confronted by a man or boy in Camden in Example 92.

(92) Omar: the trust me man they're not just like they're not just Somalian they just a lot of our people [David: I know bruv] as well . but . you thinking like . I was Somalian I was never be hold by Somalian people cos I I am Somalian I look Somalian . but the thing was . I was just there .. I was just walking **I swear to god** I was just going to see my grandma ... and all I find out all of a sudden it just all I find out was just someone kicked my bag and I was . as I looked back .. xxx just bigger than me like ..

This form was also found in a recent study in Hackney (Ilbury, 2020), along with *I swear on the Holy Bible*.

8.3.6 Phrases with *say*: *say mum's*, *say swear*, *say wallahi*

8.3.6.1 Telling someone to swear

The semantic meaning of any of the phrases with *say* appears to be telling someone to swear the truth of something. This basic meaning is in evidence in Example 93.

Ahmed asks twice, in l.1 and l.3, "where's the rizla?", which may be interpreted pragmatically as a request: Ahmed's reaction in l.9 when he finds out that CB does not have the rizla papers suggests that he was expecting CB to have the rizla. Ahmed's response to CB in l.6 is indeterminate between showing disbelief, i.e. challenging CB to tell the truth, and asking CB for confirmation of what he has just said. This use of *say mums* could be glossed as 'Is it true?', similar to Quist's (2008, p.47) interpretation of *Wallah?* in Copenhagen multiethnolect.

(93) Where's the rizla?

1	Ahmed:	where's the rizla
		(3.15)
2	Ahmed:	where's the rizla
3	CB:	I dashed it
		(0.08)
4	Ahmed:	**say mums**
5	CB:	yeah
6	CB:	why, dyou wanna bill another ting
7	Ahmed:	**on my mum's life** Ima fuck you up, **say on my mum's life** you d- you dashed it
8	CB:	listen bro I dashed it, don't say you're gonna fuck me up cos I'll fuck you up right now

8.3.6.2 Response token

Above, we discussed how *swear down* can function as a news-marking response token. This function also seems to be held by *say wallah*.

In Example 94, Ibrahim and Omar have been explaining local postcode wars to the interviewer. Omar says that one local gang "got" a person known to both Ibrahim and Omar. Ibrahim treats this as newsworthy by uttering *say wallah* with weakly rising intonation. Omar treats this as an invitation to continue: in his next turn, he continues the topic and adds the additional information "they're the ones that got me as well".

(94) Ibrahim:say wallah

1	Omar	they got [name]
		(0.07)
2	Ibrahim	**say wallah¿**
		(0.28)
3	Omar	they're the ones that got me as well ¿

The next example shows *say wallah* being used when a potential story preface has been given. The interview asks Sqara and Ahmed how they started smoking (l.2) and Ahmed then asks the question to Sqara (l.3). Sqara also repeats the question in l.4. It seems that both Sqara and Ahmed interpret the interviewer's question as fishing for a story, and that both are casting around for a story to tell. Sqara's answer in l.5, "it was a dare", is a potential story preface. Rosie's repetition of "a dare" at l.6, with weakly rising intonation, constitutes an invitation to continue. However, Sqara gives the minimal response "yeah" in l.7, which appears to be dispreferred. Ahmed indexes affiliation with two quiet laughter particles, but does not take the floor, while Rosie replies "okay", both conceding the floor to Sqara. Ahmed then says "say wallah" and Sqara replies "wallah", and after a pause of almost one second, Ahmed in l.12 and 14 gives his own story preface: he began smoking by smoking Shisha.

A story is the preferred response to the interviewer's original question and Ahmed makes a story from Sqara relevant by repeating the question to him. Sqara gives a story preface but then pauses for 0.61 seconds. Rosie's repetition of the end of Sqara's previous turn can be seen as inviting Sqara to continue his turn, but Sqara gives a minimal response rather than expanding on his previous turn. Ahmed's *say wallah* may be a simultaneously affiliative and more forceful invitation to Sqara to tell his narrative. Sqara replies "wallah"; this seems to be intended by Sqara as sequence closing (in the next section, we will discuss *say wallah/wallah* as a routine for closing a sequence) but not necessarily treated as such by Ahmed, because a pause of 0.85 seconds follows. Then Ahmed self-selects, but indicates a small topic shift by prefacing his turn with "me", showing that the focus will be on him rather than on Sqara, and gives the preface to his own story.

(95) Ahmed: say wallah

01	Sqara	yeah what (bout)
		(0.25)
02	Rosie	how did you start smoking
		(0.21)
03	Ahmed	how did you start smoking?
		(0.36)
04	Sqara	how did I start smo-
		(0.08)
05	Sqara	it was a dare
		(0.61)
06	Rosie	a dare¿
		(0.16)
07	Sqara	yeah
08	Ahmed	[hh .hh]
09	Rosie	[okay]
		(0.14)
10	Ahmed	**say wall:ah**
		(0.15)
11	Sqara	**wall:ah**
		(0.85)
12	Ahmed	me: I don't even remember (how to) start smoking
13	Sqara	it [was -]
14	Ahmed	[I was sm]oking shisha?

8.3.6.3 Routine & sequence closing

Opsahl (2009, p.229) describes *say wallah* as a "routine practice":

> as many as 16 times throughout their conversation, Farid utters *si wolla* as a minimal response to Samir's utterances. It seems in many of the cases to be an automatic minimal response rather than an actual request for the performance of a specific speech act ('to swear by God') (Opsahl, 2009, p.229)

In Opsahl's (2009) analysis, *say wallah* in Norwegian is routinised, and not a literal command to swear to God.

However, Lehtonen (2015) treats *sano wallahi/wallahi* adjacency pairs ('say wallahi'/ 'wallahi') as interactional ritual, following the definition of ritual from Rampton (2006). According to Lehtonen, the use of such *say wallahi/wallahi* pairs can be a way to cement a commitment to a promise. In such ritual contexts, Lehtonen sees a hierarchy whereby *wallahi* creates the most binding promise, and *swear* and *promise* show less commitment. 'I swear' is felt by one of Lehtonen's informants to be "lighter" than *wallahi*. Lehtonen writes that the swearing phrases show a continuum of performativity, which at one end has ritualised uses of *wallahi* that create a moral obligation between participants; at the other end, the use of such a phrase shows the speaker's commitment, but does not involve such moral obligations. According to Lehtonen, Opsahl's and Quist's characterisation of *wallah* as intensifying/emphatic sits at this end of the continuum – the one that shows epistemic commitment, but does not involve moral obligations between participants.

In the current data, *say wallah* never appears to incur the seriousness of ritual as described by Lehtonen (2015). The instances of *say wallah*, especially those that are responded to simply with *wallah*, appear to be routinised, as described by Opsahl (2009), and involving none of the moral obligations that go with ritual. The difference between *say wallah* as a response token and the *say wallah–wallah* adjacency pair seems to be that *say wallah–wallah* closes a sequence; whereas *say wallah* can also be treated by the first speaker as a response token, and they can continue as speaker. Thus the recipient can choose to continue the topic, or can turn it into a sequence-closing routine.

In Example 96, Khadir and Ali's use of *say wallah* and *wallah* in l.32–33 appears to be an adjacency pair (NB. lines 24–27 have already been quoted as Example 76). Ali's immediate response of *wallah*, with some emphasis, in l.33 seems very close to what Opsahl (2009, p.229) describes – "an automatic minimal response". This adjacency pair in lines 32–33 seems to bookend the interaction and could even be said to be a framing device. After line 33, there is a pause and then the interviewer asks a follow-up question: as far as Ali and Khadir are concerned, the *say wallah/wallah* adjacency pair seems to have put the current interactional frame to bed. The interaction involves no obligations between participants: Ali has just provided a clarification, on request from Khadir, about where a particular family known to both boys now lives. Once the clarification has been provided, the *say wallah/wallah* pair seems to mark the clarification activity as completed. After this extract finishes, these is a pause, and then the interviewer asks another question.

(96) Where do you live?

01	Lola	where do you <u>live</u>
02		(0.12)
03	Khadir	[that's a ( )]
04	Ali	[you know (place)] the park
05		(0.07)
06	Lola	yeah
07		(0.20)

08	Lola	[right there?]
09	Ali	[there.]
10		(0.67)
11	Ali	yknow,
12		(0.08)
13	Ali	yknow H–
14		(0.36)
15	Ali	yknow [name]
16		(0.88)
17	Lola	yea::h¿
18	Ali	er that has above a (.) [[name] and that]
19	Lola	[yeahyeahyeahyeah]
20		(0.23)
21	Ali	leGIT opposite them
22	Lola	↑REally?
23		(0.38)
24	Khadir	>>**I swear** they moved though innit<<
25		(0.48)
26	Ali	nah
27	Khadir	I thought you said they move
28		(0.07)
29	Ali	they moved and they came back cos the house BURNT.
30	Ali	but they fixed it.
31		(0.28)
32	Khadir	>>**say wallah**<<=
33	Ali	**=wallah**

8.3.6.4 Target specific information for clarification

As we have seen, when used as a stand-alone IP, the *say* phrase can be a news-marking response token, inviting the other speaker to continue. Yet the *say* phrase can also attach to a proposition at the left periphery in order to request clarification on some specific piece of information.

In Example 97, Tariq uses *say on mum's life* successfully to initiate repair. The two boys have been talking about a racist incident they recently experienced, and from l.2 Tariq explains that this was a freak occurrence, and that he has not experienced racism before or since. At l.6 he begins explaining that his friendship group is not racist: most White people who have Black friends, he says, have permission to use the N-word. Sami contradicts this in l.16, saying he has never had permission. In l.17, Tariq uses *say on mum's life* at the left periphery, in such a way that the phrase appears to scope over the rest of his utterance. Even without *say on mum's life* at the left periphery, Tariq's partial repetition in l.17 of Sami's utterance in l.16 has the format of a "repeat" or "understanding check" type of other-initiated repair (Sidnell, 2010, pp.117–118). Yet Sami begins his next utterance in l.18 with *on my mother's life*, spoken with emphasis, suggesting that the epistemic phrase *mum's life* is treated as important by Sami. Sami then begins repeating his utterance from l.16 but stops and self-repairs, and a clarification sequence takes place.

(97) Tariq: say on mum's life

06	Tariq	[like most] (0.09) white people nowadays,
		(0.42)
08	Sami	[hh]
09	Tariq	[who: are] ha- who have black friends,
		(0.61)
10	Tariq	have got permission to say nigga
		(1.03)
11	Tariq	[>>°the N word° <<]
12	Rosie	[permi]ssion to what
		(0.06)
13	Tariq	to say the N word
		(0.41)
14	Rosie	oh: [ok↑ay]
15	Tariq	[yeah]
		(0.25)
16	Sami	I've never had permission
		(1.12)
17	Tariq	>>**say on mum's life**<< you NEVer said [it¿]
18	Sami	**[on my] MOTHer's life**
		I've never had p-
19	Sami	(wd-)
20	Sami	woah
		(0.08)
21	Sami	(w) I've said it yeah
22	Tariq	ye[ah]
23	Sami	[I s]ay I'm not er - I've NEVer AS[Ked for permiss]ion
24	Tariq	[but you never a-/]
		(0.10)

25	Tariq	>>yeahyeah<< [but]
26	Sami	[and no one has] ever DONE some[thing]
27	Tariq	[but you]
		GET what I mean¿
28	Sami	yeah cos I'm not white I don't know I don't know bout them tings

From Example 97, it was suggested that *say wallah/mum's/swear* can preface a clarification request, and can scope over the proposition about which there is doubt. This is also the case in Example 98, where *say wallah* is used to preface an "understanding check" type of repair (Sidnell, 2010, p.118).

Omar and Ibrahim are having their interview when Sqara walks into the room. Sqara has already done the interview and knows that participants are paid ten pounds for taking part in an interview. When Sqara walks in, the first thing that Omar and Ibrahim notice is that he has had a haircut, and Omar asks "is that a fresh cut and that". Sqara says "ten pound yeah" and laughs, looking at the interviewer. Sqara is most likely addressing the interviewer and making the joke that because he has entered the interview, he should be paid ten pounds. However, Omar and Ibrahim interpret Sqara's utterance as meaning that his haircut cost ten pounds. In lines 4, 5 and 6, Omar and Ibrahim ask Sqara where he got his hair cut, and apparently get no response. Ibrahim then uses *say wallah* at the left periphery of an "understanding check" type of other-initiated repair. *Say wallah* scopes over the proposition "[that haircut cost] ten pound from eagle's".

(98) A fresh cut and that

01	Omar	#ah is that a fresh cut and that¿#
		(0.09)
02	Omar	normal
03	Sqara	ten pound yeah [hu hu hu hu hh?]
04	Omar	[(dan no). from where,] eagle's?
		(0.47)
05	Omar	where d'you [get it]
06	Ibrahim	[get it from] legal's))?
		(0.20)
07	Omar	#man said [legal's]#
08	Ibrahim	[say w-]
		(0.13)
09	Omar	a [hu hh]
10	Ibrahim	[**say wallah** that] was ten pound from eagle's?
		(0.36)
11	Sqara	no

		(0.11)
12	Omar	[where d'you get it from]
13	Sqara	[it was hassan]
		(0.19)
14	Ibrahim	[((you said)) –
15	Omar	[o:h, >>swear down<<]
16	Omar	how much
17	Ibrahim	>>swear you just said ten [pound]<<
18	Sqara	[thirteen] pound
		(0.37)
19	Omar	f- oh cool
20	Ibrahim	thirteen pound?
		(1.18)
21	Ibrahim	looks hard still.

The next example shows *say swear* being used in a way that suggests it was intended to initiate a clarification sequence. However, in this instance, *say swear* is uttered as a stand-alone IP, rather than at the left periphery of an utterance that contains the information that requires clarification. Whether because of this or for other reasons, the *say swear* utterance needs to be repaired before the clarification sequence can take place.

Example 99 shows CB using *say swear* apparently to request further information or clarification from ZR. On the day of the boys' interview, a youth worker had given ZR a telling off, because ZR in turn had been criticising his girlfriend, Jessica, for wearing clothes that he thought were too revealing. The youth worker briefly came into the interview and the interviewer asked him if he wanted to take ZR aside again. Immediately after this, CB says "say swear?" with question intonation to ZR. However, ZR initiates repair with the question-term *huh* (Schegloff, 1997; Sidnell, 2010, p.117): this suggests that whatever CB is questioning with "say swear" is not retrievable to ZR. CB's utterance in l.4 is a confirmation-seeking question. ZR responds to CB's question in l.4 by offering further explanation in l.5 and l.7.

This example is different from the others we have seen in a number of ways: it is the only instance of *say swear*, and so perhaps *say swear* is an invention of CB's on the spot, and not part of the community grammar; *say wallah/mum's*, as we have seen, are typically response tokens, appearing after one speaker has given new information, yet in Example 99, CB is not responding to something ZR has said – it would be more accurate to say that he is changing the topic and requesting ZR to give more information on what has happened between ZR, Jessica and the youth worker; when we have seen *say mum's/say wallah* requesting clarification or doing "understanding check" other-initiated repair, the *say* phrase has been used at the left periphery of an utterance in which the speaker states what their present knowledge is, yet CB here uses *say swear* as a stand-alone IP. Any one of these could be the reason why *say swear* precipitates a momentary breakdown in the interaction. One further interpretation is that this occurrence of *say swear* is of the left-periphery type, but that CB does not spell out the matter on which clarification is needed because it is difficult for CB to spell out what he requires clarification on without risking face. The matter in question is a delicate one, because ZR has been humiliated by having a youth worker discipline him for his treatment of his girlfriend. CB has a number of false starts before doing an "understanding check" in l.4, suggesting careful online planning of his utterance – which would likely be because what he is about to allude to is potentially face-threatening.

(99) CB: say swear

1	CB	**say swear**?
		(0.45)
2	ZR	huh?
		(0.30)
3	CB	he t- he he
4	CB	he had to s- <u>chat</u> to you about what Jessica wears?
		(0.60)
5	ZR	ye↑a::h let her be happy ((xx))
		(0.23)
6	CB	hh
7	ZR	watch and one day I'm just gonna be like yknow what fuck off:

This is the only token of *say swear* in the current data. However, *say swear* (although not *say wallah*) is also found by Ilbury (2020) in Hackney.

8.3.7 Functions of the epistemic phrases: discussion

Section 8.3 has analysed the different epistemic and discourse-pragmatic functions of different *I swear* phrases in the data. A number of common threads emerge.

The basic, semantic meaning of *I swear* is to attest the truth of something. As we have seen, when speakers index an epistemic stance, they may show commitment by foregrounding the reliability of evidence or the strength of their own belief. *I swear*, *swear down*, *on X life* and *wallah* all retain this key function of indexing high commitment to a proposition by the speaker. This can occur when the epistemic phrase is IU-initial, or a standalone IU. It usually co-occurs with prosodic emphasis on content words in the proposition, and/or on the epistemic phrase itself. It also co-occurs with repetition of the proposition. A common discourse context in which epistemic phrases are used to show high speaker commitment is that of the sensational news story (Lehtonen, 2015).

I swear can be used at the left periphery of an utterance to do the opposite, i.e. to show that the speaker is uncommitted and doubts the truth of a proposition. It is known that *I think*, which has grammaticalised so far as to become a discourse marker, also has these twin functions of expressing either high commitment or doubt (Kärkkäinen, 2003). When *I swear* is used in this way, it is typically spoken fast and/or phonologically reduced (another diagnostic of grammaticalisation), and uttered with rising intonation on the IU as a whole. Only *I swear* has the function of indicating doubt.

Swear down can index high commitment in the current data, just as it did in the MLE corpus. Yet it also appears with a function that is not in evidence in the MLE corpus: the phrase *swear down* (with no first-person pronoun) can be a stand-alone IU, spoken after new information has been presented by the interlocutor. It concedes the floor and allows the interlocutor to continue their turn. As such, I have described it as a news-marking response token (M. McCarthy, 2003), much like *really*. It may co-occur with the change-of-state token *oh* (Heritage, 1984); it may or may not be spoken with rising intonation; and it is typically spoken fast and /or phonologically reduced. This additional function suggests that *swear down* could potentially be in the process of grammaticalising.

The *say* phrases are not found in the MLE corpus (Cheshire et al., 2011). Their development seems to be related to the borrowing of *wallah*, seeing as equivalent *say wallah* phrases are also found in other multiethnolects where *wallah* is used (Opsahl, 2009; Quist, 2008; Lehtonen, 2015). Aside from their basic semantic meaning, they show several other discourse-pragmatic functions besides. I would speculate that initially, they had the function of challenging the epistemic status of the information one participant has offered, and that the function of requesting clarification may represent the next stage in their grammaticalisation. The overlap between these two functions can be seen in example 97. These phrases can also act as a response token, showing similar functions to *swear down?* Finally, as a routinised adjacency pair, the *say wallah/wallah* pair in particular can close a sequence in talk: the first pair part *say wallah* gives the recipient the option to expand the topic, or to reply *wallah* and close the sequence.

8.4 Distributional analysis

A distributional analysis was conducted to ascertain the rates of use of the epistemic phrases, and who their users were. Because of the difficulties of treating discourse-pragmatic features as sociolinguistic variables in a variationist-type analysis (Cheshire, 2016), the quantitative analysis takes into account only the surface forms of the different features. This means that, for example, all tokens of *I swear* are lumped together, and not separated according to whether they indicate high speaker commitment (strong conviction) or low speaker commitment (uncertainty/doubt).

Interview data and self-recorded data are considered separately because it was expected that the use of epistemic phrases would be inhibited in the interviews compared to self-recorded data (Opsahl, 2009). In compiling word counts for each participant, reading passage and wordlist data were excluded.

The figures and tables below show participants' normalised frequencies per thousand words of the different epistemic variants. The bottom row of each table shows the normalised frequency per thousand words of each epistemic phrase across all participants for that data type (i.e. in the interview data, or in the self-recorded data).

Figure 8.1 shows different participants' rates of use of the various epistemic phrases in the interview data – the same data is shown in Table 8.2, with the addition of the raw token numbers in brackets and the total number of words contributed by each participant. The first thing to observe is that 12 out of 30 participants did not use any of these phrases in their sociolinguistic interview at all. The participants who did use the epistemic phrases in their interviews did so infrequently. All of the epistemic variants except for *say mum's* appear in the interview data.

As was expected, the epistemic phrases are more frequent in the self-recorded data – though the number of participants contributing self-recorded data is lower. Figure 8.2 shows the frequencies of the epistemic phrases for the participants who contributed to the self-recordings. The accompanying token counts and overall word counts are shown in Table 8.3. Ahmed is the most frequent user of these phrases, but this is somewhat misleading, as he only contributed 67 words – he appeared briefly on CB's self-recording, and in that time happened to use three different epistemic variants. Because his data may not be representative, Figure 8.3 shows the other participants' frequencies of the epistemic phrases in the self-recorded data with Ahmed's data excluded, so that it is easier to see the differences between the other participants' frequencies.

The major finding from the distributional analysis is that *wallah* and *say wallah* are only used by adolescents who identified as Muslim: Ahmed, Ali, Sami, Karim, Tariq, Khadir, Sqara, Ibrahim, Omar, ZR, Lola and Amanda. The exception is CB, who did not describe himself as Muslim but stated that his father was Muslim.

Meanwhile, the other phrases with *say* (*say mum's, say swear*) are used by CB, Tariq and Ahmed, but also by Shantel, Chantelle and Raphael. Shantel and Chantelle are girls who live locally (outcode 1) and are some of the youth centre's most frequent attendees – they are part of the Youthclub CofP. Raphael also lives on the estate (outcode 1) but is part of the Studio CofP. This will be discussed further below.

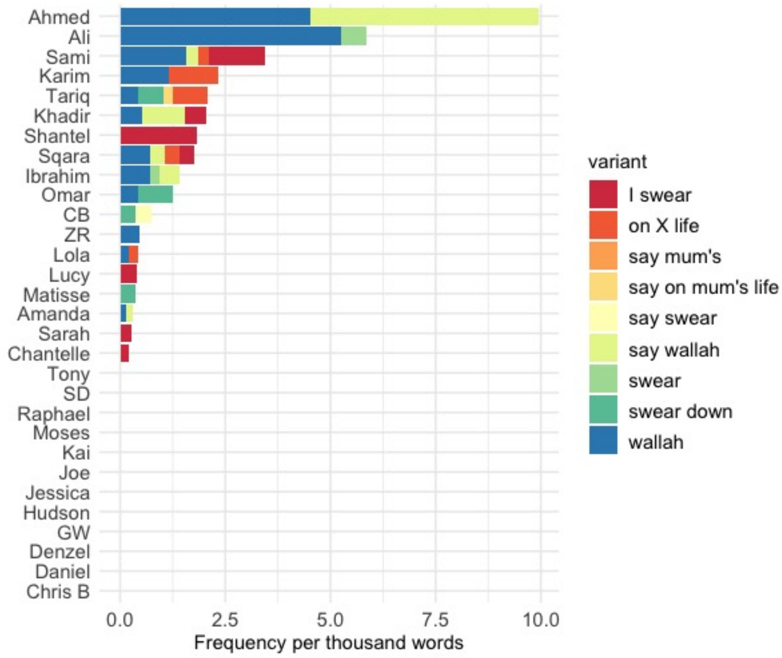


Figure 8.1: Participants' frequency of use of different epistemic phrases in the interviews

Table 8.2: Participants' total word counts in the interview data and normalised frequency of each epistemic phrase per thousand words, with raw token numbers in brackets

Participant	Total words	*I swear*	*on X life*	*say mum's*	*say on mum's life*	*say swear*	*say wallah*	*swear*	*swear down*	*wallah*
Ahmed	1107						5.42 (6)			4.52 (5)
Ali	1712							0.58 (1)		5.26 (9)
Sami	3783	1.32 (5)	0.26 (1)				0.26 (1)			1.59 (6)
Karim	1711		1.17 (2)							1.17 (2)
Tariq	4818		0.83 (4)		0.21 (1)				0.62 (3)	0.42 (2)
Khadir	1941	0.52 (1)					1.03 (2)			0.52 (1)
Shantel	1649	1.82 (3)								
Sqara	2827	0.35 (1)	0.35 (1)				0.35 (1)			0.71 (2)
Ibrahim	4285						0.47 (2)	0.23 (1)		0.70 (3)
Omar	2403								0.83 (2)	0.42 (1)
CB	2682					0.37 (1)			0.37 (1)	
ZR	2233									0.45 (1)
Lola	4567		0.22 (1)							0.22 (1)
Lucy	2448	0.41 (1)								
Matisse	5680								0.35 (2)	

Table 8.2 – *Continued from previous page*

Participant	Total words	*I swear*	*on X life*	*say mum's*	*say on mum's life*	*say swear*	*say wallah*	*swear*	*swear down*	*wallah*
Amanda	6504						0.15 (1)			0.15 (1)
Sarah	3797	0.26 (1)								
Chantelle	4680	0.21 (1)								
Total	**232,666**	**0.11 (26)**	**0.08 (18)**	**0 (0)**	**0.01 (2)**	**0.01 (2)**	**0.11 (26)**	**0.02 (4)**	**0.07 (16)**	**0.29 (68)**

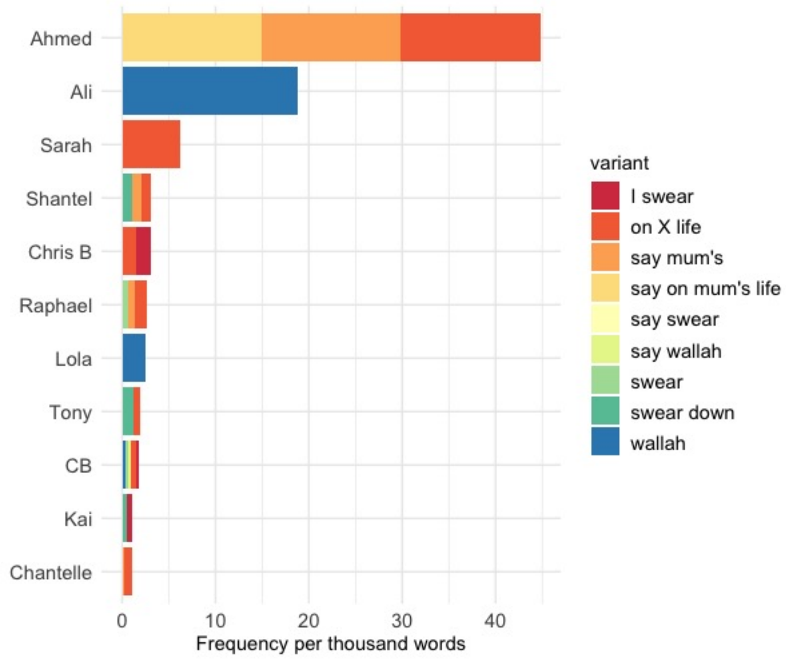


Figure 8.2: Participants' frequency of use of different epistemic phrases in the self-recordings

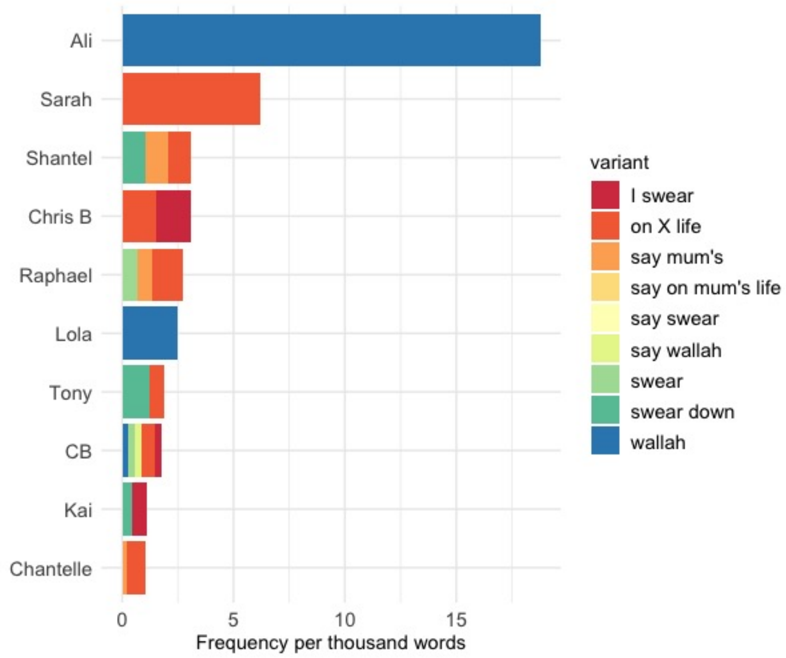


Figure 8.3: Participants' frequency of use of different epistemic phrases in the self-recordings, excluding Ahmed's data

Table 8.3: Participants' total word counts in the self-recorded data and normalised frequency of each epistemic phrase per thousand words, with raw token numbers in brackets

Participant	Total words	*I swear*	*on X life*	*say mum's*	*say on mum's life*	*say swear*	*say wallah*	*swear*	*swear down*	*wallah*
Ahmed	67		14.92 (1)	14.92(1)	14.92 (1)					
Ali	373									18.77 (7)
Sarah	161		6.21 (1)							
Shantel	968		1.03 (1)	1.03 (1)					1.03 (1)	
Chris B	652	1.53 (1)	1.53 (1)							
Raphael	1481		1.35 (2)	0.68 (1)				0.68 (1)		
Lola	402									2.49 (1)
Tony	1606		0.62 (1)						1.25 (2)	
CB	3445	0.29 (1)	0.58 (2)				0.29 (1)	0.29 (1)		0.29 (1)
Kai	4406	0.68 (3)							0.45 (2)	
Chantelle	4667		0.86 (4)	0.21 (1)						
Total	**18,228**	**0.27 (5)**	**0.71 (13)**	**0.22 (4)**	**0.05 (1)**	**0 (0)**	**0.05 (1)**	**0.11 (2)**	**0.27 (5)**	**0.49 (9)**

8.4.1 Distribution of the epistemic phrases: discussion

The distributional analysis found that overall, the epistemic phrases occur at higher rates in the self-recorded data compared to the interview data. This is similar to the findings of Opsahl (2009). In fact, when epistemic phrases did appear in the interview data, these tended not to occur in "typical" interview interaction, but either in situations where more than two interviewees were present, and the interview turned into more of a group discussion, and/or in "byplay" (Goffman, 1981, p.134), when the interviewees directly addressed one another and the interviewer was momentarily excluded. For example, Ahmed, Sqara and Karim did an interview together and passed the two microphones between the three of them, as well as to other friends who wandered in or out of the interview. Similarly, the tokens of *wallah, say wallah, on X life* contributed by Amanda and Lola occurred when they were present for other friends' interviews. Amanda sat in on Tariq and Sami's interview, and she and Lola were present for the interview with Ali and Khadir. When Amanda and Lola together were the sole interviewees, they did not use any of the epistemic phrases. This suggests that the use of epistemic phrases, especially the more innovative ones, is primarily an in-group phenomenon, and is inhibited by the presence of an outsider.

It was also found that *wallah* and *say wallah* are only used by Muslim young people in the current data. By contrast, Opsahl (2009) found instances of adolescents with two Norwegian-born parents (though she does not specify their religion) using *wallah*. Quist (2008) also mentions ethnically Danish boys who are users of multiethnolect, presumably including *wallah*. This indicates that *wallah* and *say wallah* in the current data are not muliethnolectal features, because unlike in the Norwegian and Danish data, their use does not appear to have spread to non-Muslim adolescents. Rather, they are better described as ethnolectal features.

However, the other *say* phrases – *say swear*, *say on your mum's life* and its contraction *say mum's* – show wider distribution, being used by girls from within the Youthclub CofP, and by Raphael, from the Studio CofP. This appears to offer some support for the theory that the borrowing of *wallah* leads to the development of majority-language-lexified epistemic phrases with analogous functions, as suggested by Opsahl (2009) and Lehtonen (2015). We could tentatively hypothesise that *say swear* and *say mum's* are first taken up by non-Muslim adolescents with Muslim friends – such as Shantel and Chantelle, who are integrated into a social group that has a large number of Muslim members – before spreading beyond. Raphael socialises with the boys who are part of the Studio CofP, and so potentially represents the next step in the diffusion of such phrases, in that he is part of a different social group, but is nonetheless connected to the local area and is a frequent attendee of the youth centre.

In fact, *say swear* is also attested by Ilbury (2020). Ilbury's data was collected in

Hackney, the same fieldsite as the Multicultural London English project (Cheshire et al., 2011). The attendees of the youth group were by and large Christian rather than Muslim (personal communication), and *wallah* and *say wallah* are not found in this data. We have no way of knowing whether in Ilbury's data, the use of *say swear* is at all related to *wallah* and *say wallah*; it may be the case that *say swear* is diffusing in London, but this is purely speculation.

With this in mind, it will be highly pertinent for future research to investigate: (a) whether *wallah* and *say wallah* get adopted by non-Muslim speakers; and (b) how the other epistemic phrases (particularly those with *say*) continue to develop in different parts of London.

8.5 Diffusion in progress

Finally, we turn to an example of innovation diffusion in progress.

The self-recordings that CB made offer a unique snapshot of innovation diffusion happening in real time. Example 100 is taken from a self-recording, and CB is at this point persuading a friend, X, to play pool with him and to bet money on the game. In response to his request, X uses the Arabic phrase *uqsimu billah*, which is a way of swearing to god, similar to *wallah* or *billah* (Almutlaq, 2013). CB speaks neither Arabic nor Somali. Yet when his friend tells him "say uqsem billah", CB without hesitation attempts to say it. CB pronounces it incorrectly (e.g. missing the initial glottal stop, and missing the /b/ at the beginning of *billah*), but this is not addressed as being an issue in the subsequent interaction.

(100) *uqsem billah*

1	CB	((raised voice)) you wanna put MONey on it?
		(0.16)
2	X	**say ʔuqsem billah=**
3	CB	**=ok[sumillah]**
4	X	[(. . . five]sheet)

Cheshire, Fox, Kerswill, and Torgersen (2008a) examine 7 speakers from the MLE corpus who had especially high usage of the MLE phonological innovations and who also use the innovative quotative *this is* + *speaker*. The authors suggest that aside from having highly multiethnic friendship networks, these speakers may be linguistic "brokers" (Eckert, 2000; S. Fox, 2015; Wenger, 1998). These speakers move between multiple friendship groups (cf. the findings of S. Fox (2015)) and are popular characters in these groups:

> The seven speakers who are the focus of this paper would seem to have more than multi-ethnic friendships as a common denominator. All are dominant characters within their friendship groups and highly regarded by their peers. Their friendship networks extend beyond the college grounds, giving them the opportunity for brokering. They are all involved in activities such as rapping and MCing either as participants or consumers, and these are highly valued resources in contemporary youth culture. These factors, together with the evidence from our analyses, lead us to conclude that these seven speakers are the leaders of change amongst the adolescent speakers in this study, and are representative of the social and personality types who are innovators within their group. (Cheshire et al., 2008a)

Cheshire et al. (2008a) also make the point that "some people are better at brokering than others and, as Wenger notes, some people even seem to thrive on being brokers, regularly creating connections and engaging in 'import-export' (Wenger 1998: 109)". According to this definition, CB seems likely to be a broker. He is an exuberant personality, primarily involved with the Youthclub CofP but also socialising with the Studio CofP (indeed, he opted to conduct his and ZR's interview in the recording studio). Many of the adolescents treated me, as an outsider, with suspicion when I was conducting fieldwork. CB, by contrast, was friendly to me from the beginning. Example 100 shows him willing to adopt and try an unfamiliar phrase on the fly.

8.6 Chapter summary and concluding remarks

Section 8.3 analysed the forms and functions of different epistemic phrases in the current data, and Section 8.4 presented data on their distribution between interview and self-recorded data, and across participants.

The following forms were identified: *I swear*; *swear?*; *swear down*; *on X life* (e.g. "on my life", "on my mother's life"); *I swear to god*; *wallah(i)*; *say wallah(i)*; *say swear*; *say on mum's life* and its abbreviation, *say mum's*. *I swear*, *swear?*, *swear down*, *on X life* and *I swear to god* are all found in the MLE corpus, while the others are not – although *swear down* was also shown to have an additional function in the current data that did not appear in the MLE corpus (section 8.3.2.2 above).

Commonalities were found between how *wallah* and *say wallah, say swear* are used in Norwegian and Finnish youth speech, and how they are used in the Ealing data. In particular, *wallah* is used as part of a "sensational news" genre of narrative (Lehtonen, 2015); and the pair *say wallah* – *wallah* can be used in a sequence-closing routine.

At the same time, unlike in Norway and Denmark, *wallah* and *say wallah* show no evidence of being multiethnolectal in the Ealing data – rather, they are only used by Muslim adolescents, as shown by the distributional analysis in Section 8.4. However, *say swear* and *say mum's* are used by adolescents without Muslim backgrounds, and *say swear* is also attested by Ilbury (2020). I have argued that whereas phrases containing *wallah* seem only to be used by young Muslims, *say swear* and *say mum's* are multiethnolectal features that are being adopted by non-Muslims.

8.6.1 Absence from child data

No instances of any of the epistemic phrases appear on the child recordings – neither the Diapix sessions, nor the self-recordings that were made during lessons and breaktimes.

However, *wallahi* was sometimes overheard during fieldwork in the primary school, indicating that it was used at least occasionally by some children. or example, early on in fieldwork, one child used *wallahi* while speaking with me. She was making playdough "cake" and offered me some. She told me to eat it and when I pretended to eat it, she was dissatisfied, apparently wanting me to really eat the playdough. When I said "I am eating it", she said "wallahi" and when asked to explain, she replied "It means you have to tell the truth".

The recording type could potentially be a factor here. The Diapix task was likely perceived by the children as a classroom activity – it is unsurprising that discourse-pragmatic MLE features and *wallah* do not appear in the Diapix task. During the Diapix task, the recorder was also visible on the table. The self-recordings were intended to elicit maximally un-selfconscious speech, but it's likely that they did not succeed in this respect. The children carried a highly visible large bag with them and were constantly being asked by other children what they were doing – there was little time for them to forget that their speech was being recorded.

The absence of these epistemic phrases from the child data will be returned to in Chapter 9.

8.6.2 A new epistemic mode?

In some cases, it has been suggested that discourse-pragmatic innovations lead to across-the-board pragmatic change. Opsahl (2009) suggests that as *wallah* rose to shibboleth status in Norwegian, this status also got carried over to Norwegian phrases with similar semantic meanings, and that this has led to the innovation of an entire epistemic style in Norwegian teenage speech, with this style being characterised by an array of different epistemic discourse markers. Similarly, Rodríguez Louro (2016) argues that in the case of quotatives in Australian English, the rise of non-*say* variants "is accompanied by an increase in self-revelations through reports of inner thoughts, feelings and attitudes". Outside discourse-pragmatics, Wiese (2009, 2013) has argued for the interdependence of changes taking place at different levels of German syntax in Kiezdeutsch. So there is good reason to look for connections between changes taking place at different levels of the grammar in MLE. We could relate the epistemic phrases described here to other attention-getters and response elicitors, i.e. *innit* and *(you) get me*. It would also be necessary to complement the analysis conducted in this chapter with a function-based approach, taking account of all the different tools adolescents have at their disposal for expressing epistemic stance.

However, we should not be hasty in drawing such conclusions. I agree with Lehtonen (2015, p.181) when she states "I would not seek an explanation that epistemicity as such should be more central to interaction among young people or multiethnic youth than it is in other people's discussions" (translation by Google). Although Opsahl (2009) suggests that the introduction of *wallah* triggered an increase in the use of epistemic phrases in young Norwegians' interactions, Opsahl's synchronic data provides no solid evidence for this suggestion. My data similarly has the drawback of offering only a synchronic snapshot of language, but as far as diachronic change can be inferred, my data indicates that *I swear* is more grammaticalised than *wallah*, leading me to believe that *I swear* and potentially *swear down* have actually been around longer than *wallah*. It could be argued that teenagers will inevitably be looking for ways to take and hold the conversational floor, claim attention for what they are about to say, make their narratives maximally sensational and intensify expressions of their beliefs (Tagliamonte, 2016b); in a situation of indirect language contact, they simply have more resources available to use for these functions.

Chapter 9

Summary, discussion & conclusions

9.1 Summary of thesis

To recap, the guiding research questions presented in Chapter 1 were:

1. Are MLE features used by adolescents in the Ealing fieldsite? What social and linguistic constraints govern their use?

2. Do the children appear to be acquiring the same speech variety as the adolescents?

3. Is there any evidence of multiethnolect development that has not been attested in East London?

Chapter 3 gave an in-depth description of the Ealing fieldsite. While the most populous languages among Ealing primary school pupils after English are Polish, Punjabi, Somali, Arabic and Urdu (Mangara, 2017), of these, Somali and Arabic are particularly key in the specific fieldsite and for the participants considered in this project. Many of the children and adolescents had caregivers who had emigrated from Arabic-speaking countries or the East African countries Somalia and Kenya. Groups of adolescent boys were overheard codeswitching into Arabic or Somali and the adolescent participants commented on the visibility and vitality of the local Somali community.

The participant observation led to the identification of two CofPs among youth club attendees. Table 3.2 from Chapter 3 is repeated here as Table 9.1, as it summarises the key characteristics of the two CofPs. In brief, the members of the Youthclub CofP came to the youth club for the primary purpose of socialising with their friends, whereas the members of the Studio CofP came to the youth club for the purpose of using the studio. It was suggested in Chapter 4 that this means that the Youthclub CofP have a locally-oriented outlook, while the Studio CofP is actually a sample of a larger CofP that spans London and beyond.

Table 9.1: Summary of distinguishing features of the two CofPs

	Youthclub	Studio
Gender	Boys and girls	Boys
Age	16–17	17 and over
Ethnicities	Arab, Somali, Black African, Black British, Black Caribbean, White British, White Irish	Black British, Black, African
Residence	Local	Local and non-local
Religion	Muslims, Christians	
Primary activities	Smoking outdoors, hanging out around the youth centre, playing Xbox, table tennis, pool	Recording music in the studio, discussing music and socialising in the studio
Street orientation	"active"	Orient away from road life
Music	Grime, drill – local influences	Hip-hop, rap – American influences

Chapter 4 gave an overview of the MLE features found in the Ealing data. Some features were found in both the adolescent and the child data. It was claimed that various phonetic features were shared between adolescents and children, though it was beyond the scope of the chapter to substantiate this claim with an accountable analysis – this was the task of Chapters 6–8. Indefinite article *a* before a word-initial vowel was evident in the speech of both children and adolescents. Similarly, conjoined verbs without *and* were used by both adolescents and children. Non-standard *was*, *wasn't*, *were* and *weren't* were found in the adolescent data. Non-standard *was* appeared in the child data, but none of the other non-standard past tense BE forms – but there were very few tokens of past tense BE in the child data in the first place. For many of these, further analysis would be required to identify whether the use of these features by the children is a consequence of developmental variation, or acquisition of sociolinguistic variation (Smith & Durham, 2019, pp.59–64).

Other features appeared in the adolescent data, but not in the child data. The *man* pronoun was used by some adolescent participants but did not appear anywhere in the child data. Similarly, various discourse-pragmatic features – the pragmatic markers *you* *get me* and *innit*, the interjection *rah*, the discourse marker *still*, and the intensifier *bare* – appeared in the adolescent data but not in the child data.

Finally, there were some MLE features not found in either the child or the adolescent data. There was no evidence of the quotative *this is + speaker* in the Ealing data; it was noted that Drummond (2018b) and Ilbury (2020) also did not find this feature in comparable fieldsites in Manchester and Hackney respectively.

Chapter 6 addressed RQ1, and presented the analysis of the adolescents' diphthongs. The descriptive overview suggested that overall, the adolescents show a vowel system that is similar to the emerging MLE one described by (Kerswill et al., 2008).

Variation in the adolescents' diphthongs was found to pattern with language-internal and social factors. In terms of differences between the two CofPs that had been identified in Chapter 4, it was found that the Studio CofP tend to have a more front PRICE onset (especially in *like*) and more back GOAT onset compared to the Youthclub CofP.

There were also sex differences, and these aligned with the common pattern of boys showing more advanced MLE features. Within the Youthclub CofP, the boys have more front and monophthongal PRICE and more back GOAT compared to the girls. In sum, boys seem to have more MLE-like realisations of PRICE and GOAT than the girls.

Outcode – whether an individual lived in the same postcode as the youth centre, lived in Ealing or a neighbouring borough, or came to the youth centre from further away – also had an effect. The Studio CofP members had more MLE-like realisations of the diphthongs GOAT and PRICE than the Youthclub CofP – though it was suggested that this may not be the case if it were not for the contributions of the girls within the Youthclub CofP – and within the Studio CofP, the boys who travel from further away to use the studio show the most backed realisation of GOAT. It was suggested that GOAT, more than PRICE, is indexically linked to HHSS and that possibly the sound symbolic associations of GOAT-backing that make it more useful than PRICE in indexing a stance of toughness and masculinity.

At the same time, the analysis also found evidence that GOAT-fronting seems to be in progress in this community. As has been found with studies of GOOSE- and GOAT-fronting, a preceding coronal favours a fronted realisation of the vowel. Also in line with other studies, girls lead in GOAT-fronting.

Variation in FACE did not correlate with any of the social factors, but did show some of the same language-internal conditioning found by Gates (2018): a preceding or following nasal favours a more open onset.

Chapter 7 presented the analysis of the children's diphthongs, addressing RQ2. It was found that the children had acquired the same onset qualities as the adolescents for all three diphthongs. For FACE, the children actually showed a more raised onset than the adolescents. For PRICE and GOAT, as we have seen, there were stark sex differences between the adolescents, and in both cases, the children seemed to favour an onset quality that was intermediate between that of adolescent boys and girls. However, the children differed from the adolescents in showing more diphthongal realisations of PRICE and GOAT. This suggests that monophthongisation of the diphthongs is an age-graded feature. We speculated as to the possible reasons for this in this community: this finding could be due to the formality of the school context in which the children were recorded, or because of the presence of diphthongs in the local Feature Pool. At the same time, it was pointed out that the children had acquired the same preceding environment constraints on monophthongisation as the adolescents for all three diphthongs, indicating that the children are participating in the same variable system – though they had not acquired the same following environment constraints.

Chapter 8 addressed RQ3: it investigated the adolescents' use of the Arabic borrowing *wallah*. This feature has been attested in other multiethnolects across Europe (e.g. Quist, 2008; Opsahl, 2009), but had not been picked up by prior sociolinguistic studies of British English. The analysis presented here treated *wallah* as one of a set of epistemic phrases, inspired by the work of Lehtonen (2015). The form-based analysis of *wallah* and forms based on *I swear* revealed the following epistemic phrases in the adolescent data: *I swear*; *wallah*; *swear down*; *on X life*; and a set of phrases with *say* – *say swear*, *say wallah* and *say mum's*.

As was found by Lehtonen (2015) in Helsinki data, *wallah* is used by adolescents as part of a sensational news genre of narrative in interaction. *Say wallah–wallah* sequences are used as an interactional routine, and get used in sequence closing, though the *say* phrases can serve another function too – they can target specific information for repair.

The distributional analysis showed that *wallah* and *say wallah* were used only by Muslim adolescents, plus one individual who had Muslim family. It was suggested that these two phrases could be better characterised as ethnolectal rather than multiethnolectal. *Say swear* and *say mum's*, meanwhile, did not appear to be restricted in the same way, and were used by adolescents who were not Muslim.

9.2 Discussion

9.2.1 Implications for future research

This section will discuss the implications that these findings have for future research, namely the connection between music and use of MLE, children's acquisition of MLE, and to what extent MLE is a youth style.

9.2.1.1 The role of music – rap, grime and hip-hop

The findings from Chapter 6 align with other studies of MLE and multiethnolect by indicating a link between involvement in the music scene and use of MLE features (Drummond, 2018a, 2018b; P. Pichler & Williams, 2016; Cutler & Røyneland, 2015). In particular, a connection was suggested between commitment to producing music and GOAT-backing. This finding is relevant given the suggestion by Cardoso et al. (2019) that young Londoners using backed GOAT may be perceived more negatively by outsiders. In Chapter 6 I suggested a connection between the sound symbolic associations of GOAT-backing and its heightened use by aspiring rappers. It remains for future studies to further investigate the interactional uses of GOAT-backing and -fronting by young Londoners.

This finding is worth emphasising because it is relevant beyond linguistics. P. Pichler and Williams (2016) comment on the use of a "thug life" persona in projecting a hip-hop identity, and similarly Drummond (2018b) describes the high-orientation end of his "urban/street-style" orientation scale as prototypically "an accomplished low-level grime artist", and "connected to people who would describe themselves as belonging to gangs" although "his actual participation is minimal". While neither of those studies assumes a connection between rap and violence, in the current study, as we saw in Chapter 3, the members of the Studio CofP were actually especially opposed to street violence. Yet these individuals were also those with the greatest degree of GOAT-backing, who would therefore be perceived most negatively by outsiders, if the findings of Cardoso et al. (2019) are to be believed. Further study of the perception of different MLE features is needed, like that of Cardoso et al. (2019), so as to better spread awareness and counteract bias triggered by their use.

9.2.1.2 RQ2: children's acquisition of MLE

One of the most important findings from this project is the replication of Cheshire et al.'s (2011) finding that children as young as 5 showed the same vowel system as adolescents in their community. This was found in Chapter 8 in the comparison of diphthong onset qualities between children and adolescents (Section 7.4.1). To this extent, the findings offer support for Cheshire et al.'s suggestion that in multilingual communities such as those found throughout London, where group second language learning is how many children acquire English, children orient to peers as their model in language acquisition at an earlier age than is found in monolingual communities.

Yet other features seem to be age-graded. At the level of phonetic variation, the children showed more diphthongal realisations of PRICE and GOAT (and potentially FACE too, though the evidence here is weaker) than the adolescents (Section 7.4.2). At the discourse-pragmatic level, the epistemic phrases that were in use among the adolescents did not appear in the child data – although occasional tokens of *wallahi* were overheard during fieldwork. Chapter 4 also reported no instances of the MLE pragmatic marker *you get me*, the neg-tag *innit*, discourse marker *still* or intensifier *bare* in the child data. Relatedly, Chapter 7 noted that 45% of the adolescent tokens of PRICE came from the lemma *like*, compared to 12% percent of the child tokens of PRICE, and suggested that this was because the children do not use discourse marker *like* as frequently as the adolescents do.

Of course, it is hard to know to what extent this discrepancy in the findings from the adolescent and child data is because of age-grading, or because of the different recording contexts. This was mentioned in Chapter 7 and is discussed further in Section 9.3 below.

Regarding the absence of epistemic phrases in the children's speech, to my knowledge there have been no studies of children's acquisition of epistemic phrases – although there have been a number of studies of children's acquisition of discourse markers. These studies suggest that while children may be able to produce the form of a given discourse marker by age 2–3, they may take much longer to master the range of functions that the form has in adult speech (Fox Tree, 2010). A study of four Turkish discourse markers found that children aged 9 did not use one form as frequently as adults, and had not acquired some of the adult functions of another form (Furman & Özyürek, 2007). Thus, even monolingual children would not necessarily have acquired adult-like use of discourse-pragmatic markers by age 5–7. The children in the current study have the added complication that many of them have non-native caregivers and may not hear forms such as discourse marker *like* in the input they receive at home – although many of them are likely to hear *wallah* in their heritage language. But in sum, it should not be surprising that the epistemic phrases were absent from the child data.

This still leaves unsolved the matter of why the adolescents have more monophthongal diphthongs than the children. The suggestions made in Chapter 8 included: the recording context; accommodation to the fieldworker's accent; and that the children's target in acquisition is not MLE, but the levelled form of SSBE found in the southeast of England generally; and the presence of diphthongal pronunciations of these vowels in the Feature Pool. The latter two options both point to age-grading in monophthongisation of the diphthongs, while the former two point to children's sensitivity to context and addressee, and their ability to style-shift.

It is also possible that the age-range selected – 5–7 – was not optimum for capturing child acquisition of MLE. In the MLE project, the 8-year-olds sounded somewhat more standard than the 5-year-olds (Eivind Torgersen, personal communication). Other studies have been mentioned that find that use of non-standard features drops in the first few years of school (Van Hofwegen & Wolfram, 2010; Youssef, 1991). It could be that the children who took part in this study acquired MLE in pre-school and their first year at primary school, but by the time at which they were recorded, were already on a trajectory of moving their language towards SSBE.

If there is age-grading, this in turn points to some features being part of an adolescent style.

9.2.1.3 Age-grading, variety vs. style, and MUBE

This point brings us back to the variety vs. style debate in research on multiethnolects.

As we have seen, studies of multiethnolects can be broadly divided into those taking a "structural variety" approach, and those taking a "stylistic practice" approach. I have related this debate to the issue in ethnolect studies of whether ethnolects are better treated as repertoires or as -lects: the methodological difference in approach reflects a difference in how these ways of speaking are conceptualised.

Quist's notion of "stylistic practice" is in turn inspired by Eckert's definition of style: "a clustering of linguistic resources and an association of that clustering with social meaning" (Eckert, 2001, p.123, quoted in Quist, 2008, p.51). Crucially, Eckert (1996, 2000) sees the production of style as an adolescent phenomenon: while linguistic style is relevant to adults, too, adolescence is the "hothouse" in which production of style is heightened, and pre-adolescence is the life stage at which children begin experimenting with style.

Another relevant distinction here is that between "off the shelf" and "under the counter" changes (L. Milroy, 2007). Cheshire et al. (2011) draw on this distinction to explain why an adolescent peak was found for GOOSE-fronting and quotative BE LIKE, but not for other changes. "Off the shelf" changes are "relatively freely available to appropriately positioned social actors as a stylistic and social resource, regardless of the structure and location of their primary social networks" (L. Milroy, 2007, p.152, quoted in Cheshire et al., 2011, p.179), while "under the counter" features require face to face transmission and may be more linguistically complex. According to Cheshire et al. (2011), the endogenous, "under the counter" changes had been acquired by adolescents and children alike, while adolescents were leading in the "off the shelf" changes. We could see this distinction as parallel to "variety" and "stylistic practice".

In Quist (2008), the stylistic practice half of the analysis divides Quist's adolescent participants into user and non-users of the multiethnolect: multiethnolect is one of a number of semiotic resources that the adolescents can use to position themselves in the social order. By contrast, Cheshire et al. (2011) claim that for at least some speakers in their Hackney sample, MLE has become the unmarked vernacular – the speakers' default style. Cheshire et al. (2011) point to the similarities between the London Jamaican data in Sebba (1993) and their Hackney data to say that for some young people, particularly the male Non-Anglos, MLE is the unmarked vernacular. Their apparent-time analysis also indicated that the features were developing endogenously, and being acquired in an "under the counter" way, as we have seen above.

However, there are already indications that for some young people outside London, adoption of MLE features occurs in an "off the shelf" way, and is a stylistic resource, in the sense that some speakers use many of the features, and some do not use them at all. This is the picture presented by Drummond (2016, 2018b) from his Manchester data. Like Quist (2008), Drummond (2016) is able to identify adolescents as having heavy, moderate, mild or no use of MUBE features. The speaker who is characterised as showing mild use of MUBE features is described as using "words rather than accent".

What about the Ealing fieldsite? To the extent that we have found age-graded features, there is evidence of some features being endogenous, as suggested by Cheshire et al. (2011), while others seem to behave in a stylistic-practice way. In Ealing, the age-graded developments include:

- Gender differentiation in the onset qualities of PRICE and GOAT, and the degree of monophthongisation in FACE
- Monophthongisation of PRICE and GOAT
- Increased use of discourse marker *like*
- Use of the following epistemic phrases: *swear down*, *on X life*, *wallah*, *say wallah*, *say swear*, *say mum's*
- Use of other MLE or MUBE pragmatic markers such as discourse marker *still*, *innit*, *you get me*; modifier *bare*; the *man* pronoun.

Some of these features could be part of MUBE, some may be specific to the Ealing community sampled here, and some may be more general – such as discourse marker *like*.

These fit Milroy's definition of "off the shelf" features in the sense of being easily adopted, and available as stylistic resources. Indeed, the qualitative analysis in Section 6.6 supports this idea to the extent that while the MLE diphthong variants seem to be part of some speakers' vernacular, they are used as a stylistic resource by others. We saw that with GOAT, for example, backing seems to be drawn on to shift footing or index a particular stance by speakers such as Amanda, but has become the habitual pronunciation of speakers such as Kai.

It has also been suggested that the "structural variety" vs. "stylistic practice" distinction is not just one of perspective, but one of chronology. Dorleijn and Nortier (2013) see multiethnolects as ephemeral youth styles, and at the same time, the "prestages or initial stages of stabilised contact languages". Similarly, Wiese (2020) sees youth language practices such as "translanguaging" as how varieties originate: "Over time, such practices can lead to new contact dialects as markers of a new, multiethnic urban generation. Findings on the distribution of these dialects so far suggest that they tend to emerge first in peer-group situations among young people, and can later spread to other age groups, becoming more general markers of social class, multiethnicity, or urbanness". In this light, innovations begin as in-group practices in adolescent peer groups, before eventually spreading beyond their original users to other peer groups and age groups.

Viewed from this perspective, we could hypothesise a continuum of features, with at one end, ones that have become "structural variety" features, and are acquired in an "under the counter" way by children: these would be the diphthong onset qualities. These seem to be enregistered features of MLE, and, while they are used in a "stylistic practice" way by some individuals, have become the habitual vernacular of others. At the same time, the conditions of the local Feature Pool support them through indirect language contact. This means that they are supported by various kinds of input that the children receive – accented input from caregivers, and also the language used by adolescents in the community.

Further along, there would be "stylistic practice" features, which would be those that seem to be used chiefly by adolescents – monophthongal diphthongs, and discourse-pragmatic MLE features. These are easily adopted by appropriately placed social actors – and so can be seen as "off the shelf" features – but are not supported by the local Feature Pool the way the diphthong onsets are. While they are used by adolescents, the input children receive from caregivers and teachers does not support them in the same way – so potentially children take longer to begin producing these variants.

Finally, there are some features used in an in-group way – in the current project, this seems to be the case for *wallah* and *say wallah*, which seem to be used by adolescents who identify as Muslim and who are part of a largely Muslim peer group. These have not even become "off the shelf" features yet, in that they are not available to just anyone, and rather, their use is very much dependent on "the structure and location of [the speaker's] primary social networks". Other epistemic phrases, however, such as *swear down* are "off the shelf" in that they are readily available for adoption by appropriately placed individuals.

9.2.2 Contributions of this project

Beyond the theoretical implications of the findings discussed here (see Section 9.2.1), I hope that this thesis will be useful to future researchers investigating MLE or MUBE, young people's language in London, epistemic phrases, or children's acquisition of sociolinguistic variation.

This thesis will be useful to future researchers because it gives a sociolinguistic description of young people's language in an understudied part of London. While Sharma's work focused on the South Asian community in Southall (Sharma, 2011; Sharma & Sankaran, 2011), many of the studies of MLE to date have been based in East London (e.g. Cheshire et al., 2011; S. Fox, 2015; Gates, 2019; Ilbury, 2020). It was noted that P. Pichler and Williams (2016) is an exception, but their study takes a discourse analytic perspective, rather than a variationist approach. In particular, I hope that Chapter 5 will be useful to those who want an overview of the nature of youth language in Ealing, but do not want to read the detailed analyses in Chapters 7–9. Chapter 5 is also intended to be a useful point of comparison for future researchers investigating the diffusion of MLE/MUBE features. Conversely, the sociophonetic analyses in Chapters 7 and 8 examined the language-internal constraints on variation in MLE diphthongs in this community. I hope that this will be useful to future researchers in the same way that the work of S. Fox (2015) and Gates (2018) was to me.

This project provides the first (to my knowledge) analysis of the functions of the borrowing *wallah* in English. It combines insights from Opsahl (2009), Quist (2008) and Lehtonen (2015). This will be useful to researchers working on *wallah* borrowed into other languages, and I hope it will also motivate linguists working on MLE/MUBE (or working on adolescent language in other varieties) to investigate epistemic phrases.

Finally, this project has contributed to the study of children's acquisition of sociolinguistic variation. In particular, most studies of children's sociolinguistic acquisition have focused on children growing up in monolingual communities (see Nardy, Chevrot, & Barbu, 2013). There have been very few variationist studies of children's acquisition of majority language variation in multilingual communities like those found in London, with the work by Cheshire et al. (2011) being a key exception (and see also Khattab (2007, 2013)). As mobility increases, sociolinguistic studies of people who have lived in one place all their lives, and/or have grown up acquiring one language or dialect, will become increasingly unrepresentative of the general population (Britain, 2016).

9.3 Limitations and future directions

9.3.1 Recording types

The biggest limitation in the current project is that different recording methods were used to elicit speech from children and adolescents, making it impossible to say whether differences found between the children and adolescents are real differences in their vernaculars, or whether it is an artefact of the different recording contexts.

As described in Chapter 3, having different types of speech data from the two age groups was inevitable firstly because of the specific challenges of collecting speech data from children, meaning that different recording methods needed to be used on the two age groups; and secondly because the adolescents were recorded in their youth centre, while the children were recorded at school. As described briefly in Chapter 8, the analysis of the child data included checking that the children did not have different pronunciations of the diphthongs in the Diapix keywords compared to their spontaneous speech, but arguably even their spontaneous speech may have been more formal than that of the adolescents.

9.3.2 Ages sampled

One key limitation is that no adults or elderly were surveyed. The Linguistic Innovators included elderly participants from Hackney and Havering as well as adolescent speakers (Cheshire et al., 2011). Meanwhile, in analysing their data, Kerswill et al. (2008) could also draw on a variety of studies of Cockney English conducted in locations around London's East End. To my knowledge there are no such historical studies of English in Ealing. This means the current project has no baseline with which to compare the speech of the adolescents and children. A potentially valuable direction for future study would be to survey the speech of elderly people in Ealing, and indeed elsewhere in London.

Similarly, it was beyond the scope of the current study to collect speech data from the caregivers of the children. Cheshire et al. (2011) interviewed the caregivers of the youngest children in their sample and this allowed them to state, for example, that the children's GOOSE pronunciations did not match those of their caregivers. Had caregiver speech been sampled in the current study, this could have given us insight into whether the children's more diphthongal diphthongs (compared to the adolescents) could be due to diphthongal variants in the English input they receive from caregivers.

9.3.3 MLE acquisition and language background

As described in Chapter 3, a language background questionnaire was administered to the caregivers of children participating in this project. However, not all caregivers responded, meaning that this data is missing for some children. For the adolescents, the level of language background information elicited was less detailed. Although it seems unlikely that the interspeaker differences among adolescents and children are due to L1 transfer, it would be helpful to be able to rule this out.

Relatedly, a fruitful area for future enquiry would be the potential influence of specific heritage language competence on MLE acquisition. While it is acknowledged that a single source for any of the MLE innovations is unlikely, it is also thought that MLE arises partly through group second language acquisition (Cheshire et al., 2011) – that being so, it would be fruitful to have more in-depth research into how MLE is acquired in pre-school, combining insights from K. McCarthy et al. (2014) and Nardy et al. (2014).

9.3.4 Ethnicity

A notable absence from this thesis has been the discussion of ethnicity, given that this usually plays a prominent role in discussions of multiethnolects. Moreover, recent work has shown links between the use of MLE features and ethnic orientation (Gates, 2019). There is also the valid criticism that "Multiethnolect scholarship [...] itself slips into nomadism in not itself sufficiently celebrating and taking account of the diverse ethnic backgrounds of its speakers" with the consequence that "the identities of individual speakers are lost as they are amalgamated into the multiethnic whole" (Britain, 2016, p.236).

It was not possible to include ethnicity in the statistical analysis because not all participants gave their self-defined ethnicity. At the same time, even if self-defined ethnicity had been given by all participants, this is arguably still a reductive approach in that it forces participants to categorise themselves. Including an ethnic orientation questionnaire like that of Hoffman and Walker (2010) was beyond the scope of the current study, given that investigating the link between ethnic identification and MLE use was not one of the research questions.

9.3.5 Language variation among children

This thesis has arguably taken too adult-centric an approach to analysing variation in the child data. For example, in Chapter 10 Epistemic phrases, a set of epistemic phrases was identified from examination of the adolescent data, and the child data was then searched for instances of these phrases. It was concluded that the use of epistemic phrases seems to be age-graded. However, epistemicity is surely not a feature only of adolescent and adult speech: the children probably have their own ways of expressing epistemic stance, but because the starting point for the analysis was the adolescent data, and then a form-based approach was taken, there was no room to examine how the children express epistemic stance. A goal for future research is to take a "bottom-up" approach, attending first to variation in child language and then exploring how it is structured, rather than beginning with features that vary in adult/adolescent speech and then investigating how they are acquired.

All of the 5–7 year olds made self-recordings in class and/or during lunchtime, and this data remains as yet unanalysed, simply due to constraints of time. One immediate goal for future work is to fully transcribe this data. This would allow, potentially, for future discourse-pragmatic analyses of the children's language.

9.4 Conclusion

In conclusion, regarding RQ1, the adolescents in Ealing do indeed use MLE features, though there is a great deal of inter- and intraspeaker variation. At the level of group averages, their FACE, PRICE and GOAT onsets resemble the MLE system described by Kerswill et al. (2008). The social factors influencing variation are also similar to those identified by similar studies (Cheshire et al., 2011; Drummond, 2018a): being male and involvement in rap and hip-hop both predict more MLE-like diphthongs.

Regarding RQ2, the children have acquired the MLE qualities of diphthong onsets, but not the MLE feature of monophthongisation of the diphthongs. There is also no evidence of the MLE discourse-pragmatic features in the children's speech.

As to RQ3, some of the adolescents sampled use the Arabic borrowing *wallah*, which has been attested in multiethnolects elsewhere in Europe, but had not previously been found in MLE. This seems to reflect differences in the Feature Pool between the Ealing fieldsite and the communities sampled by other studies of MLE, and the Muslim peer group within the adolescents. It was suggested that, in line with Norwegian and Finnish research (Opsahl, 2009; Lehtonen, 2015), the borrowing of *wallah* in MLE has triggered the development of *say swear* and *say mum's* by analogy with *say wallah*.

Appendix A

Additional tables from analysis of adolescents' diphthongs

A.1 Token numbers

Table A.1: Token numbers per variable by CofP, sex and outcode

CofP	sex	outcode	FACE	PRICE	*like*	GOAT
studio	M	1	360	248	199	354
studio	M	3	393	207	186	332
YC	F	1	526	433	409	571
YC	F	2	60	28	37	62
YC	M	1	127	73	24	110
YC	M	2	561	326	191	486
Total			2027	1315	1046	1915

Table A.2: FACE: token numbers by CofP, outcode, sex and preceding environment

CofP	sex	outcode	/l, w, ɹ, j/	coronal	labial	nasal	other	velar	**Total**
studio	M	1	77	192	21	31	7	32	360
studio	M	3	67	250	18	30	11	17	393
YC	F	1	103	261	45	49	27	41	526
YC	F	2	6	36	4	9	0	5	60
YC	M	1	16	71	8	6	8	18	127
YC	M	2	102	281	70	34	17	57	561
	Total		371	1091	166	159	70	170	2027

Table A.3: FACE: token numbers by CofP, outcode, sex and coda type

CofP	sex	outcode	nasal	word-final open	word-medial open	voiced obstruent	voiceless obstruent	**Total**
studio	M	1	47	125	76	41	71	360
studio	M	3	60	103	147	23	60	393
YC	F	1	54	216	123	51	82	526
YC	F	2	6	33	5	6	10	60
YC	M	1	28	42	25	16	16	127
YC	M	2	66	208	155	48	84	561
	Total		261	727	531	185	323	2027

Table A.4: PRICE: token numbers by CofP, outcode, sex and preceding environment

CofP	sex	outcode	/l, w, ɹ, j/	coronal	labial	nasal	other	velar	**Total**
studio	M	1	83	50	18	64	15	18	248
studio	M	3	98	44	22	36	3	4	207
YC	F	1	115	91	65	89	24	49	433
YC	F	2	13	3	5	7	0	0	28
YC	M	1	20	15	14	11	4	9	73
YC	M	2	123	56	46	63	17	21	326
	Total		452	259	170	270	63	101	1315

Table A.5: PRICE: token numbers by CofP, outcode, sex and coda type

CofP	sex	outcode	nasal	word-final open	word-medial open	voiced obstruent	voiceless obstruent	**Total**
studio	M	1	60	82	20	35	51	248
studio	M	3	44	70	21	38	34	207
YC	F	1	93	151	43	56	90	433
YC	F	2	2	12	2	7	5	28
YC	M	1	17	23	5	11	17	73
YC	M	2	59	110	37	57	63	326
	Total		275	448	128	204	260	1315

Table A.6: GOAT: token numbers by CofP, outcode, sex and preceding environment

CofP	sex	outcode	/l, w, ɹ, j/	coronal	labial	nasal	other	velar	**Total**
studio	M	1	29	120	29	100	28	48	354
studio	M	3	18	157	21	74	22	40	332
YC	F	1	44	214	27	166	35	85	571
YC	F	2	2	20	7	20	2	11	62
YC	M	1	7	30	6	39	10	18	110
YC	M	2	32	185	27	126	36	80	486
	Total		132	726	117	525	133	282	1915

Table A.7: GOAT: token numbers by CofP, outcode, sex and coda type

CofP	sex	outcode	nasal	word-final open	word-medial open	voiced obstruent	voiceless obstruent	**Total**
studio	M	1	85	178	50	17	24	354
studio	M	3	97	140	59	15	21	332
YC	F	1	139	314	55	39	24	571
YC	F	2	11	41	2	4	4	62
YC	M	1	23	56	16	3	12	110
YC	M	2	113	252	47	26	48	486
	Total		468	981	229	104	133	1915

A.2 Contrast matrices for sum-coded variables

Table A.8: Contrast matrix for the CofP variable in the models comparing CofPs

Studio	1
Youthclub	-1

Table A.9: Contrast matrix for the speaker sex variable in the Youthclub CofP models

F	1
M	-1

Table A.10: Contrast matrices for the outcode variable in the Youthclub CofP models and Studio CofP models

Youthclub CofP model		Studio CofP model	
Outcode 1	1	Outcode 1	1
Outcode 2	-1	Outcode 3	-1

Table A.11: Contrast matrix for the preceding environment variable

/l, w, ɹ, j/	1	0	0	0	0
coronal	0	1	0	0	0
labial	0	0	1	0	0
nasal	0	0	0	1	0
other	0	0	0	0	1
velar	-1	-1	-1	-1	-1

Table A.12: Contrast matrix for the coda type variable

nasal	1	0	0	0
word-final open syllable	0	1	0	0
word-medial open syllable	0	0	1	0
voiced obstruent	0	0	0	1
voiceless obstruent	-1	-1	-1	-1

A.3 Model summary tables

In these tables: $\hat{\beta}$ = the estimated regression coefficient; HDI = Highest Density Interval; PD = probability of direction; ROPE = region of practical equivalence; ESS = effective sample size; $\hat{R}$ = a measure of model convergence. For details on these, see Section 5.5.3.

A.3.1 FACE onset F1

Table A.13: Model summary: FACE onset F1: comparison between CofPs

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
Intercept	0.07	-0.16	0.34	0.72	0.52	831	1.00
Log(duration)	0.13	0.07	0.2	1	0.14	2146	1.00
CofP	0.07	-0.19	0.32	0.7	0.5	710	1.00
			Preceding environment				
/l, w, ɹ, j/	0.08	-0.05	0.24	0.87	0.59	2553	1.00
coronal	-0.14	-0.27	-0.02	0.99	0.24	2501	1.00
labial	0.1	-0.08	0.29	0.86	0.46	2847	1.00
nasal	0.41	0.19	0.62	1	0	2311	1.00
other	-0.08	-0.32	0.13	0.77	0.52	3078	1.00
			Following environment				
coda nasal	0.36	0.15	0.54	1	0	1784	1.00
word-final open syl.	0.1	-0.07	0.28	0.86	0.49	1590	1.00
word-medial open syl.	-0.12	-0.24	0.01	0.97	0.39	2735	1.00
coda voiced obs.	-0.3	-0.47	-0.12	1	0.01	2253	1.00
			CofP * preceding environment				
/l, w, ɹ, j/	-0.07	-0.2	0.06	0.85	0.65	2438	1.00
coronal	-0.06	-0.17	0.03	0.89	0.75	3059	1.00
labial	0.01	-0.16	0.16	0.53	0.78	3045	1.00
nasal	-0.01	-0.2	0.21	0.52	0.68	2852	1.00
other	0.1	-0.13	0.3	0.82	0.45	3331	1.00
			CofP * following environment				
coda nasal	-0.02	-0.18	0.16	0.58	0.76	2074	1.00
word-final open syl.	-0.08	-0.22	0.07	0.85	0.64	1789	1.00

Table A.13 – *Continued from previous page*

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
word-medial open syl.	0.07	-0.03	0.17	0.89	0.74	3303	1.00
coda voiced obs.	0.04	-0.09	0.18	0.70	0.79	2648	1.00

Table A.14: Model summary: FACE onset F1: Youthclub CofP

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
Intercept	0.23	-0.19	0.63	0.87	0.19	2594	1.00
Log(duration)	0.16	0.08	0.24	1	0.07	3014	1.00
outcode	-0.26	-0.67	0.13	0.91	0.17	3349	1.00
reading passage	-0.12	-0.32	0.07	0.90	0.38	4801	1.00
sex	0.16	-0.24	0.57	0.8	0.28	2574	1.00
Preceding environment							
/l, w, ɹ, j/	0.16	-0.01	0.33	0.97	0.23	4454	1.00
coronal	-0.07	-0.22	0.08	0.80	0.65	4787	1.00
labial	0.07	-0.14	0.28	0.76	0.54	4940	1.00
nasal	0.38	0.12	0.63	1	0.02	4910	1.00
other	-0.15	-0.4	0.1	0.89	0.3	5510	1.00
Following environment							
coda nasal	0.37	0.13	0.61	1	0.01	4155	1.00
word-final open syl.	0.19	0.01	0.39	0.98	0.15	4340	1.00
word-medial open syl.	-0.18	-0.34	-0.03	0.99	0.14	4251	1.00
coda voiced obs.	-0.36	-0.55	-0.18	1	0	4797	1.00
outcode * reading	-0.18	-0.41	0.04	0.94	0.22	4527	1.00
reading * sex	0.26	0.02	0.49	0.98	0.09	4278	1.00
Outcode * preceding environment							

Table A.14 – *Continued from previous page*

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
/l, w, ɹ, j/	0.01	-0.15	0.15	0.53	0.82	4696	1.00
coronal	0.01	-0.12	0.13	0.53	0.9	5091	1.00
labial	0	-0.16	0.16	0.51	0.78	6043	1.00
nasal	-0.01	-0.23	0.25	0.52	0.62	4183	1.00
other	-0.17	-0.38	0.08	0.92	0.28	6487	1.00
Outcode * following environment							
coda nasal	-0.16	-0.37	0.05	0.93	0.27	3315	1.00
word-final open	0.04	-0.1	0.2	0.72	0.74	4416	1.00
word-medial open syl.	0.08	-0.05	0.22	0.91	0.6	4643	1.00
coda voiced obs.	-0.04	-0.19	0.1	0.72	0.76	6648	1.00
outcode * sex	-0.16	-0.52	0.21	0.81	0.27	3311	1.00

Table A.15: Model summary: FACE onset F1: Studio CofP

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
Intercept	-0.01	-0.49	0.51	0.52	0.34	1839	1.00
Log(duration)	0.06	-0.04	0.17	0.87	0.81	3458	1.00
outcode	-0.02	-0.55	0.5	0.54	0.34	1544	1.00
reading passage	-0.02	-0.51	0.5	0.54	0.38	2424	1.00
Preceding environment							
/l, w, ɹ, j/	0.01	-0.25	0.25	0.52	0.59	2842	1.00
coronal	-0.23	-0.48	0.03	0.96	0.14	1953	1.00
labial	0.11	-0.29	0.54	0.73	0.34	2258	1.00
nasal	0.47	0.03	0.91	0.98	0.04	2489	1.00
other	-0.01	-0.48	0.48	0.51	0.32	2716	1.00
Following environment							

Table A.15 – *Continued from previous page*

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
coda nasal	0.41	0.14	0.7	1	0.02	2311	1.00
word-final open syl.	-0.02	-0.31	0.27	0.55	0.53	1943	1.00
word-medial open syl.	-0.04	-0.25	0.18	0.64	0.61	2669	1.00
coda voiced obs.	-0.27	-0.64	0.07	0.94	0.13	2182	1.00
outcode * reading	-0.16	-0.66	0.28	0.80	0.28	1978	1.00
Outcode * preceding environment							
/l, w, ɹ, j/	0.07	-0.14	0.29	0.75	0.56	2533	1.00
coronal	0.07	-0.14	0.28	0.76	0.58	1917	1.00
labial	-0.17	-0.55	0.27	0.82	0.27	3160	1.00
nasal	0.05	-0.37	0.48	0.61	0.41	2279	1.00
other	-0.01	-0.52	0.45	0.52	0.34	2485	1.00
Outcode * following environment							
coda nasal	-0.16	-0.35	0.05	0.93	0.27	3263	1.00
word-final open	0.01	-0.2	0.2	0.52	0.73	2724	1.00
word-medial open	0.07	-0.09	0.25	0.80	0.63	2999	1.00
coda voiced obs.	-0	-0.31	0.28	0.51	0.54	2278	1.00

A.3.2 FACE Trajectory Length

Table A.16: Model summary: FACE log(TL): comparison between CofPs

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
Intercept	-0.12	-0.32	0.09	0.86	0.42	952	1.00
Log(duration)	0.2	0.14	0.25	1	0	3452	1.00
CofP	-0.17	-0.39	0.06	0.93	0.24	819	1.01

Table A.16 – *Continued from previous page*

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
			Preceding environment				
/l, w, ɹ, j/	0.21	0.09	0.33	1	0.04	3418	1.00
coronal	0.02	-0.09	0.15	0.65	0.88	2968	1.00
labial	0.06	-0.12	0.28	0.73	0.58	2524	1.00
nasal	0.13	-0.04	0.29	0.93	0.37	3583	1.00
other	-0.1	-0.33	0.13	0.81	0.45	3416	1.00
			Following environment				
coda nasal	0.18	0.02	0.33	0.99	0.16	2519	1.00
word-final open	-0.09	-0.2	0.03	0.94	0.55	2958	1.00
word-medial open	-0.01	-0.12	0.09	0.61	0.92	3417	1.00
coda voiced obs.	-0.07	-0.21	0.1	0.80	0.64	3196	1.00
			CofP * preceding environment				
/l, w, ɹ, j/	0.01	-0.11	0.12	0.56	0.9	4037	1.00
coronal	0.05	-0.05	0.18	0.82	0.8	3277	1.00
labial	-0.07	-0.25	0.14	0.77	0.57	3690	1.00
nasal	0.03	-0.13	0.19	0.66	0.75	3405	1.00
other	-0.04	-0.27	0.18	0.64	0.59	3607	1.00
			CofP * following environment				
coda nasal	0.09	-0.06	0.23	0.88	0.56	3196	1.00
word-final open	-0.05	-0.15	0.05	0.84	0.83	3043	1.00
word-medial open	-0	-0.1	0.1	0.50	0.94	3241	1.00
coda voiced obs.	-0.02	-0.16	0.14	0.61	0.8	3267	1.00

Table A.17: Model summary: FACE log(TL): Youthclub CofP

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
Intercept	-0.05	-0.33	0.22	0.63	0.52	2077	1.00
Log(duration)	0.26	0.19	0.33	1	0	3797	1.00
outcode	0.04	-0.25	0.32	0.61	0.53	2058	1.00
reading passage	0.06	-0.22	0.31	0.69	0.5	3299	1.00
sex	0.17	-0.1	0.44	0.90	0.29	2098	1.00
Preceding environment							
/l, w, ɹ, j/	0.21	0.07	0.36	1	0.06	4584	1.00
coronal	-0.05	-0.19	0.11	0.74	0.74	3158	1.00
labial	0.14	-0.08	0.38	0.89	0.34	3308	1.00
nasal	0.11	-0.1	0.32	0.85	0.44	4164	1.00
other	-0.04	-0.3	0.23	0.61	0.51	4269	1.00
Following environment							
coda nasal	0.08	-0.12	0.26	0.79	0.57	3695	1.00
word-final open	-0.04	-0.17	0.08	0.72	0.82	3837	1.00
word-medial open	-0.01	-0.14	0.13	0.54	0.86	4636	1.00
coda voiced obs.	-0.06	-0.23	0.12	0.76	0.62	4889	1.00
outcode * reading	0.01	-0.3	0.31	0.52	0.48	3118	1.00
reading * sex	0.15	-0.18	0.45	0.84	0.32	3124	1.00
Outcode * preceding environment							
/l, w, ɹ, j/	-0	-0.14	0.14	0.52	0.84	4787	1.00
coronal	-0.04	-0.19	0.1	0.71	0.78	3352	1.00
labial	0.09	-0.13	0.3	0.80	0.5	4043	1.00
nasal	0.05	-0.14	0.25	0.68	0.64	4475	1.00
other	-0.06	-0.32	0.22	0.66	0.51	3820	1.00
Outcode * following environment							
coda nasal	0.06	-0.11	0.25	0.77	0.64	3480	1.00
word-final open	0.04	-0.06	0.14	0.76	0.89	5898	1.00

Table A.17 – *Continued from previous page*

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
word-medial open	-0.01	-0.14	0.12	0.58	0.88	5048	1.00
coda voiced obs.	-0.06	-0.22	0.11	0.76	0.67	4592	1.00
outcode * sex	0.11	-0.13	0.37	0.83	0.41	2723	1.00

Table A.18: Model summary: FACE log(TL): Studio CofP

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
Intercept	-0.11	-0.65	0.4	0.7	0.3	1817	1.00
Log(duration)	0.11	0.01	0.21	0.98	0.41	2893	1.00
outcode	-0.16	-0.65	0.36	0.77	0.27	1615	1.00
reading passage	0	-0.45	0.47	0.51	0.42	844	1.01
Preceding environment							
/l, w, ɹ, j/	0.15	-0.09	0.38	0.9	0.31	2122	1.00
coronal	0.05	-0.19	0.26	0.67	0.61	2496	1.00
labial	-0.04	-0.4	0.32	0.58	0.42	3720	1.00
nasal	0.15	-0.17	0.47	0.85	0.29	3113	1.00
other	-0.05	-0.55	0.47	0.58	0.33	3293	1.00
Following environment							
coda nasal	0.28	0	0.55	0.98	0.07	1750	1.00
word-final open	-0.13	-0.34	0.08	0.90	0.37	3208	1.00
word-medial open	-0.04	-0.23	0.13	0.68	0.69	4402	1.00
coda voiced obs.	-0.06	-0.36	0.24	0.67	0.49	3450	1.00
outcode * reading	-0.04	-0.51	0.4	0.59	0.42	2108	1.00
Outcode * preceding environment							
/l, w, ɹ, j/	0.04	-0.17	0.26	0.66	0.64	3085	1.00
coronal	0.01	-0.21	0.24	0.55	0.7	2853	1.00

Table A.18 – *Continued from previous page*

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
labial	-0.05	-0.42	0.32	0.62	0.45	3101	1.00
nasal	-0.14	-0.43	0.14	0.85	0.33	3117	1.00
other	0.26	-0.29	0.77	0.85	0.16	3035	1.00
Outcode * following environment							
coda nasal	-0.05	-0.27	0.19	0.68	0.6	2845	1.00
word-final open	0.09	-0.04	0.24	0.91	0.52	3395	1.00
word-medial open	-0.06	-0.21	0.09	0.81	0.68	3729	1.00
coda voiced obs.	-0.12	-0.39	0.15	0.83	0.37	3360	1.00

A.3.3 PRICE onset F2

Table A.19: Model summary: PRICE onset F2: comparison between CofPs

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
Intercept	0.28	-0.04	0.57	0.96	0.12	591	1.01
Log(duration)	-0	-0.08	0.07	0.51	0.99	1096	1.00
CofP	0.22	-0.07	0.53	0.94	0.18	431	1.00
Preceding environment							
/l, w, ɹ, j/	-0.1	-0.25	0.04	0.92	0.49	1749	1.00
coronal	0.15	-0	0.3	0.97	0.26	1909	1.00
labial	-0.61	-0.79	-0.43	1	0	1584	1.00
nasal	0.02	-0.17	0.22	0.59	0.67	1658	1.00
other	0.13	-0.12	0.37	0.84	0.38	2041	1.00
Following environment							
coda nasal	0.01	-0.15	0.18	0.54	0.78	1804	1.00
word-final open	-0.11	-0.29	0.07	0.88	0.46	1416	1.00

Table A.19 – *Continued from previous page*

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
word-medial open	0.17	0.04	0.31	0.99	0.15	2262	1.00
coda voiced obs.	0.01	-0.12	0.15	0.58	0.85	2057	1.00
CofP * preceding environment							
/l, w, ɹ, j/	0.05	-0.06	0.16	0.81	0.83	2066	1.00
coronal	0.1	-0.02	0.2	0.95	0.51	2778	1.00
labial	0.02	-0.11	0.15	0.64	0.84	2497	1.00
nasal	0.01	-0.15	0.17	0.57	0.79	1488	1.00
other	0	-0.18	0.19	0.51	0.71	2508	1.00
CofP * following environment							
coda nasal	0.04	-0.07	0.16	0.72	0.86	1916	1.00
word-final open	-0.04	-0.14	0.07	0.76	0.87	1871	1.00
word-medial open	0.08	-0.04	0.18	0.93	0.63	2619	1.00
coda voiced obs.	-0.06	-0.15	0.04	0.90	0.79	3181	1.00

Table A.20: Model summary: PRICE onset F2: Youthclub CofP

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
Intercept	0.33	-0.09	0.78	0.93	0.11	1766	1.00
Log(duration)	-0.05	-0.14	0.04	0.86	0.85	1558	1.00
outcode	-0.25	-0.68	0.19	0.87	0.18	1685	1.00
reading passage	-0.14	-0.44	0.2	0.81	0.33	2421	1.00
sex	-0.33	-0.81	0.08	0.94	0.11	1332	1.00
Preceding environment							
/l, w, ɹ, j/	-0.15	-0.33	0.01	0.95	0.29	2412	1.00
coronal	0.07	-0.11	0.23	0.79	0.62	2891	1.00
labial	-0.72	-0.91	-0.53	1	0	2139	1.00

Table A.20 – *Continued from previous page*

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
nasal	-0.09	-0.31	0.13	0.79	0.5	2618	1.00
other	0.18	-0.11	0.46	0.89	0.25	2528	1.00
Following environment							
coda nasal	-0.05	-0.24	0.14	0.69	0.65	2362	1.00
word-final open	-0.12	-0.3	0.08	0.88	0.43	1986	1.00
word-medial open	0.13	-0.02	0.28	0.96	0.33	2627	1.00
coda voiced obs.	0.08	-0.08	0.23	0.84	0.58	2190	1.00
outcode * reading	0.01	-0.39	0.36	0.52	0.41	2417	1.00
reading * sex	-0.06	-0.43	0.33	0.62	0.4	2360	1.00
outcode * sex	-0.08	-0.51	0.31	0.67	0.36	1823	1.00
Outcode * preceding environment							
/l, w, ɹ, j/	0.03	-0.1	0.17	0.7	0.82	2619	1.00
coronal	0	-0.12	0.12	0.51	0.89	3466	1.00
labial	0.06	-0.05	0.18	0.86	0.74	4239	1.00
nasal	0.06	-0.09	0.23	0.8	0.66	2645	1.00
other	-0.06	-0.25	0.16	0.71	0.6	3294	1.00
Outcode * following environment							
coda nasal	-0.06	-0.2	0.11	0.76	0.69	2181	1.00
word-final open	0.08	-0.05	0.21	0.91	0.59	2524	1.00
word-medial open	-0.05	-0.18	0.07	0.79	0.77	3225	1.00
coda voiced obs.	0	-0.1	0.1	0.51	0.94	3743	1.00

Table A.21: Model summary: PRICE onset F2: Studio CofP

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
Intercept	0.28	-0.3	0.91	0.85	0.16	1670	1.00
Log(duration)	0.13	-0.05	0.31	0.93	0.34	1767	1.00
outcode	-0.31	-0.9	0.35	0.87	0.14	1490	1.00
reading passage	0.03	-0.55	0.68	0.55	0.29	2850	1.00
Preceding environment							
/l, w, ɹ, j/	-0.24	-0.53	0.04	0.95	0.15	2719	1.00
coronal	0.16	-0.12	0.48	0.86	0.31	2966	1.00
labial	-0.64	-1.18	-0.11	0.99	0.01	1194	1.00
nasal	-0.11	-0.55	0.37	0.71	0.31	2148	1.00
other	0.17	-0.57	0.84	0.72	0.23	2344	1.00
Following environment							
coda nasal	0.08	-0.19	0.38	0.72	0.45	2958	1.00
word-final open	-0.18	-0.47	0.12	0.89	0.25	1839	1.01
word-medial open	0.18	-0.08	0.47	0.9	0.25	3496	1.00
coda voiced obs.	-0.02	-0.44	0.36	0.55	0.44	2313	1.00
outcode * reading	-0.15	-0.69	0.46	0.73	0.26	2425	1.00
Outcode * preceding environment							
/l, w, ɹ, j/	0.17	-0.06	0.4	0.93	0.25	2933	1.00
coronal	0.17	-0.1	0.41	0.91	0.25	2941	1.00
labial	0.11	-0.39	0.55	0.72	0.32	1276	1.00
nasal	0.16	-0.3	0.61	0.82	0.28	2399	1.00
other	-0.23	-0.99	0.54	0.76	0.18	2448	1.00
Outcode * following environment							
coda nasal	-0.06	-0.29	0.18	0.7	0.57	3113	1.00
word-final open	0.08	-0.11	0.26	0.82	0.54	3122	1.00
word-medial open	-0.03	-0.27	0.24	0.6	0.57	3681	1.00

Table A.21 – *Continued from previous page*

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
coda voiced obs.	-0.11	-0.49	0.28	0.75	0.37	2553	1.00

A.3.4 PRICE in *like* onset F2

Table A.22: Model summary: PRICE in *like* onset F2: comparison between CofPs

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
Intercept	0.34	0.03	0.64	0.98	0.07	459	1.01
Log(duration)	-0.02	-0.08	0.05	0.75	0.99	1489	1.00
CofP	0.48	0.21	0.77	1	0.01	652	1.00

Table A.23: Model summary: PRICE in *like* onset F2: Youthclub CofP

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
Intercept	0.27	-0.17	0.74	0.89	0.17	1929	1.00
Log(duration)	0.01	-0.1	0.12	0.57	0.92	1661	1.00
outcode	0.01	-0.45	0.44	0.51	0.37	1879	1.00
sex	-0.34	-0.8	0.08	0.94	0.11	1824	1.00
outcode * sex	0.09	-0.32	0.56	0.67	0.33	2145	1.00

Table A.24: Model summary: PRICE in *like* onset F2: Studio CofP

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
Intercept	0.13	-0.51	0.81	0.68	0.24	1481	1.00
Log(duration)	-0.09	-0.24	0.05	0.92	0.57	2095	1.00

outcode	-0.38	-1.03	0.31	0.9	0.1	1330	1.00

A.3.5 PRICE Trajectory Length

Table A.25: Model summary: PRICE log(TL): comparison between CofPs

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
Intercept	0.04	-0.11	0.22	0.7	0.71	1147	1.00
Log(duration)	0.45	0.38	0.52	1	0	2174	1.00
CofP	-0.06	-0.22	0.11	0.78	0.67	1080	1.00
Preceding environment							
/l, w, ɹ, j/	-0.13	-0.27	0	0.97	0.31	2233	1.00
coronal	-0.29	-0.48	-0.09	1	0.03	1899	1.00
labial	0.03	-0.18	0.24	0.61	0.64	2110	1.00
nasal	-0	-0.28	0.25	0.51	0.56	1743	1.00
other	0.35	0.05	0.66	0.99	0.05	2278	1.00
Following environment							
coda nasal	0.12	-0.04	0.27	0.93	0.41	2032	1.00
word-final open	-0.02	-0.19	0.12	0.63	0.78	1965	1.00
word-medial open	0.07	-0.08	0.25	0.82	0.6	2981	1.00
coda voiced obs.	-0	-0.16	0.16	0.52	0.77	2609	1.00
CofP * preceding environment							
/l, w, ɹ, j/	0.04	-0.07	0.16	0.78	0.83	2808	1.00
coronal	-0.1	-0.27	0.07	0.87	0.5	2293	1.00
labial	-0.03	-0.19	0.14	0.67	0.74	2949	1.00
nasal	0.01	-0.23	0.23	0.53	0.59	1802	1.00
other	0.03	-0.22	0.3	0.61	0.54	2711	1.00
CofP * following environment							
coda nasal	0.05	-0.1	0.17	0.77	0.76	2989	1.00
word-final open	0.04	-0.08	0.14	0.77	0.87	3044	1.00

Table A.25 – *Continued from previous page*

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
word-medial open	-0.02	-0.18	0.13	0.62	0.77	3216	1.00
coda voiced obs.	0.01	-0.14	0.16	0.56	0.81	2750	1.00

Table A.26: Model summary: PRICE log(TL): Youthclub CofP

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
Intercept	0.06	-0.2	0.33	0.68	0.52	2837	1.00
Log(duration)	0.51	0.43	0.58	1	0	4514	1.00
outcode	-0.05	-0.3	0.19	0.65	0.55	2455	1.00
reading passage	-0.06	-0.35	0.24	0.67	0.47	4442	1.00
sex	0.21	-0.01	0.46	0.96	0.16	2156	1.00
Preceding environment							
/l, w, ɹ, j/	-0.2	-0.36	-0.05	0.99	0.11	4057	1.00
coronal	-0.23	-0.47	0.01	0.97	0.12	3514	1.00
labial	0.04	-0.18	0.26	0.65	0.59	3379	1.00
nasal	-0.01	-0.29	0.26	0.52	0.51	3676	1.00
other	0.38	0.03	0.71	0.98	0.05	3525	1.00
Following environment							
coda nasal	0.09	-0.1	0.26	0.82	0.54	3487	1.00
word-final open	-0.06	-0.26	0.11	0.76	0.62	2448	1.00
word-medial open	0.1	-0.09	0.27	0.86	0.49	4698	1.00
coda voiced obs.	-0.03	-0.22	0.15	0.62	0.69	3383	1.00
outcode * reading	0.17	-0.17	0.5	0.85	0.28	4408	1.00
reading * sex	-0.07	-0.37	0.27	0.67	0.42	3942	1.00
Outcode * preceding environment							
/l, w, ɹ, j/	-0.03	-0.15	0.09	0.7	0.86	5052	1.00

Table A.26 – *Continued from previous page*

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
coronal	0.23	-0	0.43	0.98	0.1	3474	1.00
labial	-0.06	-0.22	0.11	0.79	0.67	4946	1.00
nasal	-0.01	-0.23	0.25	0.53	0.6	3450	1.00
other	-0.34	-0.62	-0.08	0.99	0.03	4008	1.00
Outcode * following environment							
coda nasal	-0.11	-0.26	0.04	0.92	0.45	4316	1.00
word-final open	0.12	0.02	0.23	0.99	0.32	6116	1.00
word-medial open	-0.05	-0.22	0.11	0.74	0.69	5372	1.00
coda voiced obs.	-0.06	-0.2	0.08	0.8	0.69	5668	1.00
outcode * sex	0.08	-0.14	0.33	0.77	0.52	3118	1.00

Table A.27: Model summary: PRICE log(TL): Studio CofP

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
Intercept	-0.1	-0.46	0.2	0.76	0.41	2730	1.00
Log(duration)	0.4	0.31	0.49	1	0	6582	1.00
reading passage	0.41	-0.1	0.9	0.94	0.08	2746	1.00
outcode	0.18	-0.12	0.54	0.89	0.25	1846	1.00
Preceding environment							
/l, w, ɹ, j/	0.05	-0.2	0.29	0.68	0.56	2771	1.00
coronal	-0.33	-0.64	-0.03	0.98	0.05	2766	1.00
labial	0.07	-0.3	0.47	0.65	0.41	2802	1.00
nasal	0.1	-0.71	0.79	0.63	0.24	1702	1.00
other	0.19	-0.35	0.78	0.77	0.22	3256	1.00
Following environment							
coda nasal	0.23	-0.03	0.51	0.95	0.14	3356	1.00

Table A.27 – *Continued from previous page*

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
word-final open	0.1	-0.11	0.3	0.84	0.47	3386	1.00
word-medial open	0.17	-0.22	0.62	0.83	0.26	2968	1.00
coda voiced obs.	-0.04	-0.39	0.4	0.59	0.43	1878	1.00
Outcode * preceding environment							
/l, w, ɹ, j/	-0.22	-0.46	0.02	0.97	0.14	2988	1.00
coronal	-0.26	-0.54	0.04	0.96	0.11	2703	1.00
labial	-0.01	-0.37	0.34	0.53	0.46	3176	1.00
nasal	0.1	-0.57	0.9	0.63	0.25	1784	1.00
other	0.43	-0.1	1.02	0.94	0.08	3530	1.00
Outcode * following environment							
coda nasal	-0.08	-0.34	0.18	0.75	0.48	2508	1.00
word-final open	-0.09	-0.27	0.09	0.85	0.51	4132	1.00
word-medial open	0	-0.43	0.38	0.51	0.44	2827	1.00
coda voiced obs.	0.07	-0.32	0.43	0.66	0.41	2160	1.00

A.3.6 PRICE in *like* Trajectory Length

Table A.28: Model summary: PRICE in *like* log(TL): comparison between CofPs

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
Intercept	-0.05	-0.26	0.16	0.68	0.62	772	1.01
Log(duration)	0.41	0.32	0.49	1	0	2119	1.00
CofP	0.02	-0.19	0.22	0.57	0.66	771	1.01

Table A.29: Model summary: PRICE in *like* log(TL): Youthclub CofP

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
Intercept	-0.11	-0.47	0.28	0.73	0.35	2141	1.00
Log(duration)	0.46	0.34	0.59	1	0	1433	1.00
outcode	0.09	-0.26	0.45	0.7	0.4	1705	1
sex	0.09	-0.28	0.44	0.7	0.39	1754	1.00
outcode * sex	0	-0.36	0.36	0.51	0.46	2333	1.00

Table A.30: Model summary: PRICE in *like* log(TL): Studio CofP

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
Intercept	0.04	-0.42	0.51	0.57	0.35	1381	1.00
Log(duration)	0.3	0.13	0.48	1	0.02	1746	1.00
outcode	0.01	-0.47	0.49	0.52	0.39	1570	1.00

A.3.7 GOAT onset F2

Table A.31: Model summary: GOAT onset F2: comparison between CofPs

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
Intercept	-0.13	-0.45	0.23	0.77	0.35	581	1.00
Log(duration)	-0.14	-0.17	-0.11	1	0.01	7032	1.00
CofP	-0.31	-0.63	0.02	0.97	0.1	672	1.00
reading passage	0.04	-0.07	0.15	0.76	0.87	5438	1.00
Preceding environment							
/l, w, ɹ, j/	-0.12	-0.27	0.03	0.94	0.38	2826	1.00
coronal	0.28	0.13	0.42	1	0.01	2007	1.00
labial	-0.25	-0.41	-0.09	1	0.04	2673	1.00
nasal	0.1	-0.05	0.25	0.9	0.51	2149	1.00

Table A.31 – *Continued from previous page*

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
other	0.08	-0.13	0.29	0.77	0.53	2085	1.00
			Following environment				
coda nasal	-0.07	-0.24	0.09	0.8	0.61	2145	1.00
word-final open	0.12	-0.02	0.27	0.95	0.4	1932	1.00
word-medial open	0.05	-0.08	0.19	0.77	0.75	2566	1.00
coda voiced obs.	0.06	-0.12	0.24	0.76	0.62	2700	1.00
CofP * reading	-0.05	-0.15	0.05	0.84	0.82	4705	1.00
			CofP * preceding environment				
/l, w, ɹ, j/	0.05	-0.07	0.18	0.77	0.76	3261	1.00
coronal	-0.06	-0.16	0.05	0.87	0.79	2403	1.00
labial	0.02	-0.11	0.15	0.6	0.85	3427	1.00
nasal	0.04	-0.06	0.16	0.78	0.85	2725	1.00
other	-0.07	-0.25	0.13	0.76	0.59	2104	1.00
			CofP * following environment				
coda nasal	0.1	-0.01	0.21	0.96	0.53	2717	1.00
word-final open	0.02	-0.08	0.12	0.7	0.94	2329	1.00
word-medial open	-0.02	-0.13	0.08	0.66	0.92	3115	1.00
coda voiced obs.	-0.02	-0.16	0.1	0.65	0.84	3327	1.00

Table A.32: Model summary: GOAT onset F2: variation within the Youthclub CofP

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
Intercept	0.03	-0.52	0.56	0.56	0.3	1502	1.00
Log(duration)	-0.15	-0.2	-0.11	1	0.02	3113	1.00
outcode	0.02	-0.51	0.52	0.53	0.31	1271	1.00
reading passage	0.08	-0.08	0.25	0.84	0.59	3111	1.00

Table A.32 – *Continued from previous page*

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
sex	0.43	-0.08	0.94	0.96	0.08	1200	1.00
Preceding environment							
/l, w, ɹ, j/	-0.21	-0.38	-0.05	0.99	0.1	2872	1.00
coronal	0.37	0.22	0.52	1	0	2445	1.00
labial	-0.26	-0.45	-0.07	1	0.05	2744	1.00
nasal	0.06	-0.11	0.24	0.76	0.63	2284	1.00
other	0.1	-0.11	0.31	0.84	0.47	2745	1.00
Following environment							
coda nasal	-0.17	-0.35	-0	0.97	0.2	1657	1.00
word-final open	0.13	-0.03	0.3	0.94	0.37	1566	1.00
word-medial open	0.08	-0.07	0.24	0.85	0.58	2525	1.00
coda voiced obs.	0.07	-0.15	0.27	0.74	0.56	2155	1.00
outcode * reading	0.05	-0.14	0.24	0.71	0.64	2679	1.00
reading * sex	0.05	-0.13	0.25	0.74	0.64	2462	1.00
outcode * sex	-0.06	-0.59	0.46	0.59	0.29	1462	1.00
Outcode * preceding environment							
/l, w, ɹ, j/	-0.01	-0.15	0.13	0.53	0.84	3355	1.00
coronal	-0.1	-0.19	-0.01	0.99	0.48	2752	1.00
labial	-0.05	-0.2	0.12	0.71	0.71	2864	1.00
nasal	-0.02	-0.13	0.09	0.65	0.91	2161	1.00
other	0.22	0.03	0.41	0.99	0.09	2663	1.00
Outcode * following environment							
coda nasal	0.03	-0.07	0.12	0.74	0.93	3248	1.00
word-final open	-0.02	-0.11	0.07	0.64	0.96	2621	1.00
word-medial open	-0.02	-0.13	0.1	0.64	0.9	2982	1.00
coda voiced obs.	-0.02	-0.18	0.13	0.61	0.77	2862	1.00

Table A.33: Model summary: GOAT onset F2: variation within the Studio CofP

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
Intercept	-0.2	-0.5	0.13	0.91	0.21	1880	1.00
Log(duration)	-0.13	-0.2	-0.06	1	0.18	4217	1.00
outcode	0.58	0.26	0.89	1	0.01	1674	1.00
reading passage	-0	-0.46	0.49	0.51	0.41	2442	1.00
Preceding environment							
/l, w, ɹ, j/	-0.11	-0.58	0.33	0.71	0.34	2136	1.00
coronal	0.28	-0.01	0.56	0.97	0.08	2616	1.00
labial	-0.3	-0.66	0.07	0.95	0.11	2184	1.00
nasal	0.25	-0.04	0.53	0.96	0.13	2695	1.00
other	0.02	-0.45	0.55	0.54	0.34	2313	1.00
Following environment							
coda nasal	0.1	-0.16	0.39	0.76	0.41	2275	1.00
word-final open	0.1	-0.15	0.34	0.79	0.45	2129	1.00
word-medial open	0.05	-0.27	0.39	0.64	0.48	2169	1.00
coda voiced obs.	0.06	-0.29	0.42	0.63	0.42	3184	1.00
outcode * reading	-0.05	-0.54	0.37	0.62	0.4	1927	1.00
Outcode * preceding environment							
/l, w, ɹ, j/	-0.01	-0.46	0.45	0.51	0.39	1914	1.00
coronal	0.12	-0.18	0.4	0.81	0.38	2485	1.00
labial	-0.02	-0.37	0.33	0.54	0.46	1947	1.00
nasal	-0.09	-0.37	0.18	0.77	0.46	2450	1.00
other	-0.14	-0.64	0.34	0.74	0.28	2141	1.00
Outcode * following environment							
coda nasal	0.06	-0.2	0.32	0.68	0.52	2784	1.00
word-final open	-0.11	-0.34	0.12	0.85	0.43	2747	1.00
word-medial open	0.15	-0.16	0.51	0.84	0.31	2443	1.00

Table A.33 – *Continued from previous page*

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
coda voiced obs.	-0.12	-0.47	0.24	0.76	0.36	3070	1.00

A.3.8 GOAT Trajectory Length

Table A.34: Model summary: GOAT log(TL): comparison between CofPs

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
Intercept	-0.1	-0.32	0.12	0.82	0.46	980	1.00
Log(duration)	0.19	0.13	0.25	1	0	2898	1.00
CofP	-0.01	-0.24	0.21	0.56	0.64	845	1.00
Preceding environment							
/l, w, ɹ, j/	0.02	-0.14	0.2	0.57	0.74	3626	1.00
coronal	0.02	-0.1	0.13	0.61	0.9	3435	1.00
labial	-0.07	-0.26	0.14	0.76	0.56	2997	1.00
nasal	0.3	0.14	0.46	1	0.01	2552	1.00
other	-0.02	-0.19	0.14	0.61	0.75	4005	1.00
Following environment							
coda nasal	0.24	0.08	0.42	1	0.05	2651	1.00
word-final open	-0.02	-0.14	0.1	0.6	0.89	2957	1.00
word-medial open	-0.04	-0.17	0.09	0.71	0.79	3574	1.00
coda voiced obs.	-0.07	-0.26	0.12	0.76	0.57	3710	1.00
CofP * preceding environment							
/l, w, ɹ, j/	0.03	-0.15	0.19	0.65	0.72	3100	1.00
coronal	0	-0.11	0.13	0.53	0.91	2980	1.00
labial	-0.06	-0.24	0.16	0.73	0.59	3849	1.00
nasal	0.07	-0.09	0.22	0.82	0.6	2530	1.00
other	0.03	-0.13	0.19	0.62	0.75	3136	1.00

Table A.34 – *Continued from previous page*

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
			CofP * following environment				
coda nasal	0.12	-0.06	0.27	0.92	0.39	2347	1.00
word-final open	0.03	-0.09	0.14	0.71	0.87	3023	1.00
word-medial open	-0	-0.13	0.14	0.52	0.86	3476	1.00
coda voiced obs.	-0.06	-0.25	0.14	0.72	0.61	3226	1.00

Table A.35: Model summary: GOAT log(TL): variation within the Youthclub CofP

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
Intercept	-0.09	-0.35	0.18	0.76	0.45	2389	1.00
Log(duration)	0.26	0.18	0.35	1	0	2748	1.00
outcode	-0.02	-0.3	0.24	0.55	0.54	1626	1.00
reading passage	0.03	-0.18	0.24	0.6	0.62	5010	1.00
sex	0.04	-0.23	0.32	0.62	0.55	1401	1.00
			Preceding environment				
/l, w, ɹ, j/	-0	-0.21	0.22	0.51	0.64	4067	1.00
coronal	0	-0.12	0.16	0.53	0.84	3427	1.00
labial	-0.02	-0.29	0.26	0.57	0.53	3623	1.00
nasal	0.24	0.04	0.43	0.99	0.08	2554	1.00
other	-0.05	-0.26	0.18	0.66	0.59	4065	1.00
			Following environment				
coda nasal	0.14	-0.07	0.36	0.91	0.33	2824	1.00
word-final open	-0.05	-0.21	0.09	0.77	0.7	2950	1.00
word-medial open	-0.05	-0.25	0.16	0.69	0.63	3727	1.00
coda voiced obs.	-0.02	-0.26	0.23	0.57	0.58	3698	1.00

Table A.35 – *Continued from previous page*

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
outcode * reading	0.06	-0.21	0.35	0.66	0.49	3067	1.00
reading * sex	0.13	-0.14	0.41	0.81	0.37	2867	1.00
Outcode * preceding environment							
/l, w, ɹ, j/	-0.05	-0.27	0.17	0.68	0.59	4034	1.00
coronal	0.01	-0.11	0.14	0.57	0.88	4543	1.00
labial	0.27	0	0.55	0.97	0.1	3283	1.00
nasal	-0.1	-0.28	0.08	0.86	0.46	2216	1.00
other	-0.12	-0.33	0.08	0.87	0.41	4082	1.00
Outcode * following environment							
coda nasal	-0.09	-0.31	0.11	0.79	0.52	2975	1.00
word-final open	-0.05	-0.17	0.09	0.75	0.79	4107	1.00
word-medial open	0.13	-0.08	0.33	0.89	0.37	3145	1.00
coda voiced obs.	0.04	-0.19	0.27	0.62	0.58	4378	1.00
outcode * sex	0.12	-0.13	0.36	0.83	0.4	1895	1.00

Table A.36: Model summary: GOAT log(TL): variation within the Studio CofP

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
Intercept	-0.17	-0.9	0.54	0.72	0.2	1180	1.00
Log(duration)	0.08	0	0.17	0.98	0.67	2907	1.00
outcode	-0.15	-0.87	0.57	0.69	0.23	1135	1.00
reading passage	0.1	-0.23	0.43	0.74	0.39	2586	1.00
Preceding environment							
/l, w, ɹ, j/	0.01	-0.32	0.34	0.52	0.47	2299	1.00
coronal	0.02	-0.22	0.28	0.58	0.63	1739	1.00
labial	-0.13	-0.43	0.2	0.79	0.35	2204	1.00

Table A.36 – *Continued from previous page*

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
nasal	0.4	0	0.73	0.98	0.04	1660	1.00
other	-0.01	-0.34	0.33	0.53	0.48	1630	1.00
			Following environment				
coda nasal	0.31	0.04	0.59	0.98	0.06	2540	1.00
word-final open	0.01	-0.24	0.24	0.52	0.64	2644	1.00
word-medial open	-0.07	-0.27	0.16	0.74	0.57	2627	1.00
coda voiced obs.	-0.06	-0.45	0.37	0.62	0.4	1508	1.00
outcode * reading	0.06	-0.3	0.37	0.66	0.45	3236	1.00
			Outcode * preceding environment				
/l, w, ɹ, j/	0.07	-0.29	0.4	0.68	0.42	2155	1.00
coronal	-0.01	-0.27	0.2	0.55	0.69	1657	1.00
labial	-0.01	-0.29	0.29	0.53	0.52	2966	1.00
nasal	-0.13	-0.47	0.2	0.81	0.35	1721	1.00
other	0.04	-0.26	0.41	0.61	0.48	2057	1.00
			Outcode * following environment				
coda nasal	0.02	-0.23	0.26	0.56	0.62	2340	1.00
word-final open	0.04	-0.16	0.28	0.69	0.65	1718	1.00
word-medial open	0.02	-0.18	0.23	0.59	0.67	3169	1.00
coda voiced obs.	-0.09	-0.51	0.35	0.68	0.36	1201	1.00

Appendix B

Additional tables from analysis of children's diphthongs

B.1 Token numbers

Table B.1: Token numbers by vowel and age

	FACE	PRICE	PRICE in *like*	GOAT
Children	841	842	117	942
Adolescents	1590	1076	866	1546
Total	2431	1918	983	2488

Table B.2: FACE: token numbers by age and preceding environment

age	/l, w, ɹ, j/	coronal	labial	nasal	other	velar	**Total**
adolescent	290	829	137	129	56	149	1590
child	136	248	89	47	85	236	841
Total	426	1077	226	176	141	385	2431

Table B.3: FACE: token numbers by age and coda type

age	nasal	word-final open	word-medial open	voiced obstruent	voiceless obstruent	**Total**
adolescent	201	620	371	158	240	1590
child	109	303	177	30	222	841
Total	310	923	548	188	462	2431

Table B.4: PRICE: token numbers by age and preceding environment

age	/l, w, ɹ, j/	coronal	labial	nasal	other	velar	**Total**
adolescent	347	215	141	229	53	91	1076
child	211	91	182	266	38	54	842
Total	558	306	323	495	91	145	1918

Table B.5: PRICE: token numbers by age and coda type

age	nasal	word-final open	word-medial open	voiced obstruent	voiceless obstruent	**Total**
adolescent	223	370	100	166	217	1076
child	155	224	111	171	181	842
Total	378	594	211	337	398	1918

Table B.6: GOAT: token numbers by age and preceding environment

age	/l, w, ɹ, j/	coronal	labial	nasal	other	velar	**Total**
adolescent	112	554	90	446	106	238	1546
child	77	262	32	355	139	77	942
Total	189	816	122	801	245	315	2488

Table B.7: GOAT: token numbers by age and coda type

age	nasal	word-final open	word-medial open	voiced obstruent	voiceless obstruent	**Total**
adolescent	364	831	165	85	101	1546
child	144	546	161	37	54	942
Total	508	1377	326	122	155	2488

B.2 Contrast matrices for sum-coded variables

Table B.8: Contrast matrix for the age variable

Adolescent	1
Child	-1

Table B.9: Contrast matrix for the speaker sex variable

F	1
M	-1

Table B.10: Contrast matrix for the preceding environment variable

/l, w, ɹ, j/	1	0	0	0	0
coronal	0	1	0	0	0
labial	0	0	1	0	0
nasal	0	0	0	1	0
other	0	0	0	0	1
velar	-1	-1	-1	-1	-1

Table B.11: Contrast matrix for the coda type variable

nasal	1	0	0	0
word-final open syllable	0	1	0	0
word-medial open syllable	0	0	1	0
voiced obstruent	0	0	0	1
voiceless obstruent	-1	-1	-1	-1

B.3 Model summary tables

In these tables: $\hat{\beta}$ = the estimated regression coefficient; HDI = Highest Density Interval; PD = probability of direction; ROPE = region of practical equivalence; ESS = effective sample size; $\hat{R}$ = a measure of model convergence. For details on these, see Section 5.5.3.

B.3.1 FACE model summary tables

Table B.12: Model summary: FACE onset F1: adolescents and children

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
Intercept	-0.02	-0.22	0.17	0.60	68.95	835	1.00
log(duration)	0.22	0.18	0.26	1.00	0	5245	1.00
age	0.16	-0.03	0.36	0.95	24.9	902	1.01
Preceding environment							
/l, w, ɹ, j/	0.25	0.11	0.39	1.00	17	1733	1.00
coronal	-0.18	-0.33	-0.06	1.00	11.25	1651	1.00
labial	-0.04	-0.21	0.15	0.64	67.98	2172	1.00
nasal	0.15	-0.02	0.33	0.95	26.55	2338	1.00
other	0.07	-0.11	0.25	0.78	59.78	2847	1.00
Following environment							
coda nasal	0.06	-0.12	0.24	0.77	61.8	1791	1.00
word-final open	0.13	-0.01	0.28	0.96	33.42	1712	1.00
word-medial open	-0.01	-0.14	0.13	0.55	84.8	1787	1.00

Table B.12 – *Continued from previous page*

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
coda voiced obs.	-0.2	-0.4	-0.02	0.98	14	1492	1.00
sex	0.02	-0.15	0.2	0.58	73.6	833	1.00
Age * preceding environment							
/l, w, ɹ, j/	-0.15	-0.28	-0.01	0.99	0.229	1778	1.00
coronal	0.1	-0.02	0.22	0.95	50.1	1807	1.00
labial	0.11	-0.06	0.29	0.89	43.35	2361	1.00
nasal	0.21	0.05	0.39	0.99	9.48	2499	1.00
other	-0.24	-0.41	-0.06	0.99	6.22	2854	1.00
Age * following environment							
coda nasal	0.26	0.1	0.43	1.00	2.55	1821	1.00
word-final open	0.02	-0.13	0.15	0.60	83.9	1743	1.00
word-medial open	-0.09	-0.22	0.03	0.92	54.42	1834	1.00
coda voiced obs.	-0.13	-0.31	0.04	0.93	35.28	1923	1.00
age * sex	0.04	-0.14	0.2	0.66	71.85	652	1.01

Table B.13: Model summary: FACE log(TL): adolescents and children

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
Intercept	0.07	-0.06	0.19	0.87	0.6485	1322	1.00
log(duration)	0.29	0.25	0.34	1	0	6568	1.00
age	-0.06	-0.2	0.07	0.83	0.7315	1474	1.00
Preceding environment							
/l, w, ɹ, j/	0.28	0.18	0.39	1	0.0015	3904	1.00
coronal	-0.04	-0.13	0.05	0.80	0.9038	3348	1.00
labial	0.08	-0.07	0.23	0.87	0.5685	3799	1.00
nasal	-0.01	-0.14	0.13	0.57	0.8375	4424	1.00

Table B.13 – *Continued from previous page*

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
other	0.01	-0.14	0.16	0.57	0.801	4248	1.00
			Following environment				
coda nasal	0.05	-0.07	0.16	0.79	0.8125	3271	1.00
word-final open	-0.05	-0.13	0.04	0.85	0.8888	3299	1.00
word-medial open	0	-0.09	0.1	0.54	0.9508	3690	1.00
coda voiced obstruent	-0.14	-0.3	0.01	0.96	0.2838	2670	1.00
sex	0.11	-0.01	0.22	0.96	0.4363	1245	1.00
			Age * Preceding environment				
/l, w, ɹ, j/.	-0.06	-0.17	0.04	0.87	0.7515	3262	1.00
coronal	0.04	-0.06	0.14	0.80	0.8625	3353	1.00
labial	0.02	-0.12	0.17	0.60	0.8022	3583	1.00
nasal	0.1	-0.05	0.24	0.91	0.4918	3589	1.00
other	-0.06	-0.21	0.09	0.80	0.6695	4138	1.00
			Age * following environment				
coda nasal	0.08	-0.04	0.21	0.90	0.6132	2908	1.00
word-final open	-0.01	-0.1	0.09	0.60	0.964	3359	1.00
word-medial open	-0.03	-0.13	0.08	0.70	0.899	2803	1.00
coda voiced obs.	0.08	-0.09	0.23	0.83	0.5812	2731	1.00
age * sex	0.15	0.04	0.26	0.99	0.1825	1376	1.00

B.3.2 PRICE model summary tables

Table B.14: Model summary: PRICE onset F2: adolescents and children

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
Intercept	0.17	-0.01	0.37	0.96	22.95	1345	1.00

Table B.14 – *Continued from previous page*

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
log(duration)	-0.12	-0.16	-0.08	1	19.05	7502	1.00
age	-0	-0.2	0.19	0.51	71.4	1234	1.00
Preceding environment							
/l, w, ɹ, j/	-0.12	-0.25	0.01	0.96	0.3865	2870	1.00
coronal	0.16	0.01	0.32	0.98	0.1958	2689	1.00
labial	-0.71	-0.89	-0.55	1	0	2856	1.00
nasal	0.04	-0.13	0.23	0.69	0.6722	2763	1.00
other	0.33	0.08	0.56	1	0.031	3062	1.00
Following environment							
coda nasal	-0.01	-0.16	0.14	0.57	0.796	2532	1.00
word-final open	-0.04	-0.22	0.13	0.68	0.696	2347	1.00
word-medial open	0.14	0.01	0.27	0.97	0.2925	2953	1.00
coda voiced obs.	-0.09	-0.23	0.06	0.89	0.5435	2715	1.00
sex	-0.27	-0.45	-0.1	1	0.0255	1246	1.00
Age * preceding environment							
/l, w, ɹ, j/.	0.01	-0.09	0.12	0.60	0.9255	3279	1.00
coronal	-0.11	-0.24	0.02	0.95	0.4132	2838	1.00
labial	0.03	-0.1	0.16	0.70	0.8228	3472	1.00
nasal	-0.08	-0.21	0.07	0.86	0.6078	2649	1.00
other	-0.15	-0.35	0.07	0.91	0.309	3066	1.00
Age * following environment							
coda nasal	-0.01	-0.12	0.1	0.57	0.9155	2671	1.00
word-final open	-0.03	-0.14	0.09	0.68	0.8732	3000	1.00
word-medial open	-0.01	-0.14	0.11	0.59	0.875	3244	1.00
coda voiced obs.	0.12	0.01	0.24	0.98	0.3538	3190	1.00
age * sex	-0.25	-0.4	-0.08	1	0.0352	1220	1.00

Table B.15: Model summary: PRICE log(TL): adolescents and children

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
Intercept	0.13	0.02	0.23	0.99	30.2	2078	1.00
log(duration)	0.49	0.45	0.52	1	0	6018	1.00
age	-0.28	-0.38	-0.19	1	0.03	1786	1.00
Preceding environment							
/l, w, ɹ, j/	-0.14	-0.24	-0.04	1	20.67	3305	1.00
coronal	-0.18	-0.33	-0.04	0.99	13.02	3462	1.00
labial	0.07	-0.05	0.2	0.88	65.82	3145	1.00
nasal	0.01	-0.15	0.16	0.53	79.45	2927	1.00
other	0.35	0.18	0.54	1	0.32	3735	1.00
Following environment							
coda nasal	-0.06	-0.16	0.07	0.84	77.28	2944	1.00
word-final open	-0.12	-0.24	0	0.98	37	2821	1.00
word-medial open	0.13	0.03	0.23	1	26.3	3763	1.00
coda voiced obs.	0.01	-0.1	0.12	0.58	92.42	2876	1.00
sex	0.06	-0.03	0.14	0.91	82.02	1508	1.00
Age * preceding environment							
/l, w, ɹ, j/	-0.02	-0.1	0.06	0.67	97.18	4156	1.00
coronal	-0.01	-0.14	0.12	0.59	86.6	3131	1.00
labial	0.02	-0.07	0.12	0.67	93.25	4409	1.00
nasal	0.03	-0.11	0.14	0.65	85.68	3337	1.00
other	0.02	-0.14	0.18	0.61	76.22	4010	1.00
Age * following environment							
coda nasal	0.13	0.04	0.21	1	23.43	4325	1.00
word-final open	0.07	-0.01	0.15	0.95	79.05	4399	1.00
word-medial open	-0.1	-0.2	-0.01	0.97	51	4403	1.00
coda voiced obs.	0.03	-0.06	0.12	0.73	94.22	3849	1.00

Table B.15 – *Continued from previous page*

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
age * sex	0.12	0.04	0.21	1	31.85	1513	1.00

Table B.16: Model summary: PRICE in *like* onset F2: adolescents and children

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
Intercept	0.28	0.01	0.59	0.97	0.0993	1107	1.00
log(duration)	-0.07	-0.12	-0.03	1	0.862	3761	1.00
age	-0.18	-0.46	0.08	0.91	0.2505	1300	1.00
sex	-0.21	-0.49	0.06	0.93	0.1955	1289	1.00
age * sex	-0.3	-0.58	-0.02	0.98	0.0813	985	1.00

Table B.17: Model summary: PRICE in *like* log(TL): adolescents and children

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
Intercept	0.43	0.26	0.59	1	0	1288	1.00
log(duration)	0.4	0.36	0.45	1	0	4754	1.00
age	-0.55	-0.72	-0.39	1	0	1174	1.00
sex	0.07	-0.09	0.23	0.78	0.6312	1600	1.00
age * sex	0.05	-0.1	0.23	0.74	0.6848	1319	1.00

B.3.3 GOAT model summary tables

Table B.18: Model summary: GOAT onset F2: adolescents and children

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
Intercept	0.03	-0.23	0.26	0.58	0.5765	969	1.00
log(duration)	-0.16	-0.19	-0.13	1	0	6023	1.00
age	-0.08	-0.34	0.17	0.75	0.501	557	1.00
Preceding environment							
/l, w, ɹ, j/	-0.14	-0.3	0.03	0.95	0.3212	1909	1.00
coronal	0.2	0.04	0.35	0.99	0.1115	1967	1.00
labial	-0.19	-0.39	0.01	0.97	0.184	2036	1.00
nasal	0.17	0	0.36	0.97	0.2	2062	1.00
other	0.05	-0.15	0.23	0.67	0.6435	2112	1.00
Following environment							
coda nasal	-0.11	-0.34	0.13	0.83	0.4438	1834	1.00
word-final open	0.04	-0.13	0.21	0.68	0.713	1329	1.00
word-medial open	-0.04	-0.19	0.15	0.67	0.7245	1784	1.00
coda voiced obs.	0.23	0.06	0.43	0.99	0.085	2188	1.00
sex	0.16	-0.09	0.39	0.91	0.2872	781	1.01
Age * preceding environment							
/l, w, ɹ, j/	-0.06	-0.21	0.09	0.77	0.6862	2314	1.00
coronal	0.14	-0.01	0.3	0.96	0.3192	1107	1.00
labial	-0.06	-0.25	0.1	0.77	0.6285	2136	1.00
nasal	-0.12	-0.26	0.04	0.94	0.403	2161	1.00
other	0.08	-0.1	0.26	0.83	0.5495	2379	1.00
Age * following environment							
coda nasal	-0	-0.19	0.19	0.52	0.7035	1602	1.00
word-final open	0.06	-0.08	0.18	0.79	0.7495	1688	1.00
word-medial open	0.14	-0.01	0.29	0.97	0.265	2064	1.00
coda voiced obs.	-0.15	-0.32	0.02	0.96	0.27	2147	1.00

Table B.18 – *Continued from previous page*

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
age * sex	0.19	-0.03	0.39	0.96	0.2087	836	1.01

Table B.19: Model summary: GOAT log(TL): adolescents and children

Parameter	$\hat{\beta}$	95% HDI lower	95% HDI upper	PD	% in ROPE	ESS	$\hat{R}$
Intercept	0.03	-0.1	0.14	0.67	0.8678	1926	1.00
log(duration)	0.29	0.25	0.33	1	0	6253	1.00
age	-0.17	-0.29	-0.04	1	0.1475	1695	1.00
Preceding environment							
/l, w, ɹ, j/	0.04	-0.09	0.18	0.72	0.7832	4432	1.00
coronal	-0.07	-0.17	0.03	0.93	0.7042	3513	1.00
labial	0.03	-0.16	0.22	0.61	0.6862	2703	1.00
nasal	0.18	0.06	0.29	1	0.0958	2451	1.00
other	-0.01	-0.15	0.13	0.55	0.8235	3986	1.00
Following environment							
coda nasal1	0.15	-0	0.29	0.98	0.2545	2688	1.00
word-final open	-0.09	-0.18	0.01	0.96	0.613	3207	1.00
word-medial open	0.01	-0.12	0.14	0.58	0.8698	3931	1.00
coda voiced obs.	-0.08	-0.25	0.08	0.83	0.5728	3611	1.00
sex	0.03	-0.09	0.15	0.69	0.8435	1382	1.00
Age * preceding environment							
/l, w, ɹ, j/	0	-0.13	0.14	0.52	0.8502	3545	1.00
coronal	0.07	-0.04	0.17	0.90	0.7338	3955	1.00
labial	-0.07	-0.26	0.13	0.75	0.5908	3324	1.00
nasal	0.05	-0.09	0.18	0.77	0.7685	2834	1.00
other	-0.03	-0.17	0.12	0.67	0.7822	3739	1.00

Table B.19 – *Continued from previous page*

Parameter	$\hat{\beta}$	**95% HDI lower**	**95% HDI upper**	**PD**	**% in ROPE**	**ESS**	$\hat{R}$
			Age * following environment				
coda nasal	0.05	-0.1	0.2	0.75	0.7168	2657	1.00
word-final open	0.04	-0.08	0.15	0.75	0.8552	2678	1.01
word-medial open	-0.03	-0.15	0.11	0.65	0.835	3646	1.00
coda voiced obs.	-0.02	-0.18	0.16	0.59	0.7402	3270	1.00
age * sex	0.02	-0.1	0.13	0.62	0.8918	1294	1.00

Appendix C

Information sheet & consent form sent to parents

Information sheet

Multilingual children and London youth language: information for parents

Dear Parent,

We would very much like to invite your child to be part of a research project, which is about children's multilingualism and youth speech in London. Please read the following information very carefully. You should only agree for your child to take part if you and your child both want to, it is entirely up to you both. If you choose for your child not to take part there won't be any disadvantages for either of you and you will hear no more about it.

The following information will tell you why the research is being done and what your child will be asked to do if s/he takes part. Please ask if there is anything that is not clear or if you would like more information (see contact details below).

If you decide that your child can take part you will be asked to **sign the attached form** to say that you agree.

You are still free to withdraw your child from the study at any time and without giving a reason.

Multilingual children and London youth language

The study is about youth language in multilingual areas of London. In particular we want to know if children's language is strongly connected to youth language, or whether "youth speak" is a different language variety. The study is about, and aims to encourage, dialect diversity in London – so we are not judging whether or not your children speak "correct" English.

For this project, I would like to compare the speech of young children with the speech of teenagers from the same area, to see whether they speak in the same way or whether and in what ways children's speech changes as they become older.

What is involved?

If you decide to let your child take part, I will ask you to:

- Complete a questionnaire about which languages your child speaks, who they speak them with and how often they use them.
- Give me permission to take your child and a friend aside at school for a short conversation (30 minutes)
- Give me permission to audio-record your child and their friends interacting as they normally do during lessons and at playtime.

Anonymity

In transcripts of the recordings, your child's real name, the names of any other people, and any place names which could be used to identify individuals, will be replaced with fake names. The sound files will be kept in password-protected folders on the researcher's computer.

Confidentiality

Information about your child will be kept confidential, unless an issue pertaining to child safeguarding arises.

Use of the data

The information gathered may be published in an academic journal, including direct quotations from your child. If you do not wish for your child to be quoted, you should show this on the consent form attached. I am also asking your permission to be able to play small clips of sound in academic presentations: you have the option whether to consent to this or not. In such sound clips, names of people and places will be removed with sound editing software so that no one can be identified.

It is up to you to decide whether or not your child can take part. If you do decide to let your child take part you will be given this information sheet to keep and be asked to sign a consent form.

If you have any questions or concerns about the manner in which the study was conducted please, in the first instance, contact the researcher responsible for the study:

Rosie Oxbury, r.f.oxbury@qmul.ac.uk

If this is unsuccessful, or not appropriate, please contact the Secretary at the Queen Mary Ethics of Research Committee, Room W104, Queen's Building, Mile End Campus, Mile End Road, London or research-ethics@qmul.ac.uk.

Queen Mary
University of London

Consent form

Please complete this form after you have read the Information Sheet and/or listened to an explanation about the research.

Title of Study: Multilingual children and London youth language
Queen Mary Ethics of Research Committee Ref: ________________

Thank you for considering letting your child take part in this research. The person organizing the research must explain the project to you before you agree for your child to take part.

If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you decide whether your child can join in. You will be given a copy of this Consent Form to keep and refer to at any time.

I understand that if I decide at any other time during the research that I no longer wish for my child to participate in this project, I can notify the researchers involved and my child will be withdrawn from it immediately.

I consent to the processing of my child's personal information for the purposes of this research study. I understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.

Please indicate (✓) if:

- I **DO NOT** consent to my child being quoted (anonymously) in academic publications
- I **DO NOT** consent to audio clips of my child's voice being used for teaching purposes
- I **DO NOT** consent to audio clips of my child's voice being used for conference presentations
- I **DO NOT** consent to being contacted about research in the future

Participant's Statement:

I ____ agree that the research project named above has been explained to me to my satisfaction and I agree for my child to take part in the study. I have read both the notes written above and the Information Sheet about the project, and understand what the research study involves.

Signed: Date:

Investigator's Statement:

I ____ confirm that I have carefully explained the nature, demands and any foreseeable risks (where applicable) of the proposed research to the parent of the participating child

Signed: Date:

Appendix D

Language background questionnaire given to caregivers

Table D.1: Questionnaire given to parents

About Child
Where was Child born?
What is his/her date of birth?
Where else has s/he lived apart from London?
About you:
What is your name?
What is your gender?
Where do you live now?
When and where were you born?
Where else have you lived in your life?
How long have you been:
in the UK?
in London?
Who else lives with you and Child?
Languages

What languages do you know?

How well can you speak and understand them?

What language(s) do you use to talk to Child?

What language(s) does Child use to talk to you?

What languages does Child know?

How well can s/he speak and understand them?

Did Child speak English before starting school at [[School]]?

Other caregiver (e.g. Child's mother/father, your partner/husband/wife):

Does Child live with or see regularly another parent or caregiver?

What is your relation to the other parent or caregiver (e.g. married)?

When and where was his/her second caregiver born?

Where else has he or she lived?

Where does he or she live now?

What languages does he or she know?

How well can he or she speak and understand them?

How long has he or she been:

in the UK?

in London?

What language(s) do YOU speak with HIM/HER?

What language(s) does HE/SHE use to talk to Child?

What language(s) does Child use to talk to HIM/HER?

Appendix E

Information sheet & consent form given to adolescent participants

Queen Mary
University of London

Information sheet

Multilingual London's youth language: information for participants

We would like to invite you to be part of this research project, if you would like to. You should only agree to take part if you want to, it is entirely up to you. If you choose not to take part there won't be any disadvantages for you and you will hear no more about it.

Please read the following information carefully before you decide to take part; this will tell you why the research is being done and what you will be asked to do if you take part. Please ask if there is anything that is not clear or if you would like more information.

If you decide to take part you will be asked to sign the attached form to say that you agree.

You are still free to withdraw at any time and without giving a reason.

Multilingual London's youth language

The study is about youth culture and youth language in multicultural areas of London. In particular we want to know if youth language is different in different parts of the city, and also whether there is a big difference between the language used by children and the language used by young adults. The study is about, and aims to encourage, dialect diversity in London.

<u>What is involved?</u>

If you volunteer to take part I will ask you to:

1. complete a questionnaire telling me more about your background, including what languages you speak
2. take part in an interview/group conversation, lasting approximately one hour

This will take place in one of the rooms at the youth centre. You can choose who you take part with – or you can do the interview by yourself if you prefer. I will ask you questions about life in Ealing, your experiences growing up in London, the friendships you have and what music you listen to. You will be rewarded £10 for taking part in this. I may follow up the interviews later, asking you about things that you said in the earlier interviews, but you will only receive a reward for the first interview.

3. record yourself talking with your friends for approximately 30 minutes

I would like you to choose some part of your normal routine that you do with your friends (e.g. catching the bus, hanging out after school) and make an audio-recording approximately 30 minutes long of you doing this, using your phone. This is so that I have an idea of the kind of conversations you have when there aren't adults around! Before you start recording, you should tell everyone around you what you're doing, and give them my contact details and/or a copy of the "Bystanders' information" sheet. If anyone objects to you audio-recording near them, please respect what they say and leave the recording for another time.

It is up to you which of these you take part in. For example, you may complete the questionnaire and nothing else. You might record yourself on your phone, but not take part in the interview or complete a questionnaire. Of course, I am hoping that you will do all three!

Anonymity

The sound files will be kept in password-protected folders on the researcher's computer. In transcripts of the data, your real name, the names of any other people, and any place names which could be used to identify individuals, will be replaced with fake names.

Use of the data

The information gathered may be published in an academic journal, including short direct quotations from you. If you do not wish to be quoted, you should show this on the consent form attached. I am also asking your permission to be able to play small clips of sound in academic presentations: you have the option whether to consent to this or not. In such sound clips, names of people and places will be removed with sound editing software so that no one can be identified.

It is up to you to decide whether or not to take part. If you do decide to take part you will be given this information sheet to keep and be asked to sign a consent form.

If you have any questions or concerns about the manner in which the study was conducted please, in the first instance, contact the researcher responsible for the study:

Rosie Oxbury, r.f.oxbury@qmul.ac.uk

If this is unsuccessful, or not appropriate, please contact the Secretary at the Queen Mary Ethics of Research Committee, Room W104, Queen's Building, Mile End Campus, Mile End Road, London or research-ethics@qmul.ac.uk.

Queen Mary
University of London

Consent form

Please complete this form after you have read the Information Sheet and/or listened to an explanation about the research.

Title of Study: Multilingual London's youth language
Queen Mary Ethics of Research Committee Ref: ________________

Thank you for considering taking part in this research. The person organizing the research must explain the project to you before you agree to take part.

If you have any questions arising from the Information Sheet or explanation already given to you, please ask the researcher before you decide whether to join in. You will be given a copy of this Consent Form to keep and refer to at any time.

I understand that if I decide at any other time during the research that I no longer wish to participate in this project, I can notify the researchers involved and be withdrawn from it immediately.

I consent to the processing of my personal information for the purposes of this research study. I understand that such information will be treated as strictly confidential and handled in accordance with the provisions of the Data Protection Act 1998.

Please indicate if:

- I **DO NOT** consent to being quoted (anonymously) in academic publications
- I **DO NOT** consent to audio clips of my voice being used for teaching purposes
- I **DO NOT** consent to audio clips of my voice being used for conference presentations
- I **DO NOT** consent to being contacted about research in the future

Participant's Statement:

I ____ agree that the research project named above has been explained to me to my satisfaction and I agree to take part in the study. I have read both the notes written above and the Information Sheet about the project, and understand what the research study involves.

Signed:

Date:

Investigator's Statement:

I ____ confirm that I have carefully explained the nature, demands and any foreseeable risks (where applicable) of the proposed research to the volunteer

Signed: Date:

Appendix F

Consent form given to bystanders during self-recordings

Queen Mary University of London

Bystanders' information sheet & consent form

Multilingual London's youth language

Title of Study: Multilingual children and London youth language
Queen Mary Ethics of Research Committee Ref: QMERC2016/71

Your friend/family member is helping to carry out research for Queen Mary University of London.

The aim of the project is to find out about youth culture and youth language in London, and how these differ between different areas of the city. To do this, we are asking young people such as your friend/family member to record themselves having an ordinary conversation with friends or family.

We would like to be able to transcribe your speech to analyse as part of the project, if you are aged 16 or over. If you are happy for this to happen, and you are aged 16 or older, you can return this form to the researcher.

If you do not return this form, or if you are younger than 16, your voice will not be transcribed, although we will keep the sound file in order to use the rest of the recording for our research.

If you want your voice to be destroyed entirely from the sound file, you should contact the researcher directly to say so.

If you want to know more about the study, or if you have concerns about the manner in which the study was carried out, please contact Rosie Oxbury at r.f.oxbury@qmul.ac.uk, who is carrying out the research. If this is unsuccessful, or not appropriate, please contact the Secretary at the Queen Mary Ethics of Research Committee, Room W104, Queen's Building, Mile End Campus, Mile End Road, London or research-ethics@qmul.ac.uk.

If you do not choose to take part there will be no disadvantages for you and you won't hear any more about the study.

Participant's Statement:

I ____ agree that the research project named above has been explained to me to my satisfaction and I agree to take part in the study. I have read both the notes written above about the project, and understand what the research study involves.

Signed: Date:

References

Aarsæther, F., Marzo, S., Nistov, I., & Ceuleers, E. (2015). Indexing locality: Contemporary urban vernaculars in Belgium and Norway. In J. Nortier & B. A. Svendsen (Eds.), *Language, youth and identity in the 21st century: linguistic practices across urban spaces* (pp. 249–270). Cambridge: Cambridge University Press.

Aasheim, S. (1995). *Kebab-norsk: fremmedspråklig påvirkning på ungdomsspråket i Oslo. [KebabNorwegian: influence of foreign languages on the youth language in Oslo]* (Unpublished master's thesis). University of Oslo, Oslo.

Adams, Z. (2018). 'I don't know why man's calling me family all of a sudden': Address and reference terms in Grime music. *Language and Communication*, *60*, 11–27.

Adams, Z., & Cheshire, J. (2013). *An analysis of 'still' as a new discourse marker in Multicultural London English.* (Unpublished manuscript, QMUL)

Aijmer, K. (2002). *English discourse particles: evidence from a corpus*. Philadelphia: John Benjamins.

Alam, F. (2015). *"Glaswasian"? a sociophonetic analysis of Glasgow-Asian accent and identity* (Unpublished doctoral dissertation). University of Glasgow, Glasgow.

Al-Khawaldeh, A. (2018). Uses of the discourse marker *wallahi* in Jordanian spoken Arabic: A pragma-discourse perspective. *International Journal of Humanities and Social Science*, *8*(6), 114–123.

Almutlaq, H. (2013). A sociolinguistic study of terms of oaths in Jordanian Arabic. *International journal of Humanities and Social Science*, *3*(21), 225–228.

Andersen, E. S. (1990). *Speaking with style: the sociolinguistic skills of children*. London: Routledge.

Andersen, G. (2001). *Pragmatic markers and sociolinguistic variation: a relevance-theoretic approach to the language of adolescents*. Amsterdam: John Benjamins.

Andersen, G. (2016). Using the corpus-driven method to chart discourse-pragmatic change. In H. Pichler (Ed.), *Discourse-pragmatic variation and change in English: new methods and insights* (pp. 19–40). Cambridge: Cambridge University Press.

Anderson, B. R. O. (2006). *Imagined communities: reflections on the origin and spread of nationalism* (Rev. ed.). London; New York: Verso.

Appel, R., & Schoonen, R. (2005). Street language: A multilingual youth register in the Netherlands. *Journal of multilingual and multicultural development*, *26*(2), 85–117.

Arai, T. (2005). Comparing tongue positions of vowels in oral and nasal contexts. In *6th INTERSPEECH*.

Aylett, M., & Turk, A. (2006). Language redundancy predicts syllabic duration and the spectral characteristics of vocalic syllable nuclei. *The Journal of the Acoustical Society of America*, *119*(5), 3048–3058.

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

Bamberg, M., & Georgakopoulou, A. (2008). Small stories as a new perspective in narrative and identity analysis. *Text & Talk*, *28*(3), 377–396.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

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

Becker, K. (2013). The sociolinguistic interview. In C. Mallinson, B. Childs, & G. Van Herk (Eds.), *Data collection in sociolinguistics* (pp. 91–117). New York: Routledge.

Bell, A. (1984). Language style as audience design. *Language in society*, *13*(2), 145–204.

Benor, S. B. (2010). Ethnolinguistic repertoire: Shifting the analytic focus in language and ethnicity. *Journal of Sociolinguistics*, *14*(2), 159–183.

Bernard, H. R. (2011). *Research methods in anthropology: qualitative and quantitative approaches* (5th ed.). Lanham, Md: AltaMira Press.

Biber, D., & Finegan, E. (1989). Styles of stance in English: Lexical and grammatical marking of evidentiality and affect. *Text*, *9*(1), 93–124.

Blommaert, J., & Rampton, B. (2011). Language and superdiversity. *Diversities*, *13*(2), 1–20.

Bodén, P. (2010). Pronunciation in Swedish multiethnolect. In P. Quist & B. A. Svendsen (Eds.), *Multilingual Urban Scandinavia: new linguistic practices* (pp. 65–78). Bristol: Multilingual Matters.

Boersma, P., & Weenink, D. (2019, January). *Praat: doing phonetics by computer* (software No. 6.0.46). Computer program. Retrieved from `http://www.praat.org/`

Brent, C. (2012). 'First step: Dress cool': Young people's representations of locality. *Youth & Policy*(109), 46–59.

Briggs, C. L. (1986). *Learning how to ask: A sociolinguistic appraisal of the role of the interview in social science research.* Cambridge: Cambridge University Press.

Britain, D. (1992). Linguistic change in intonation: The use of high rising terminals in New Zealand English. *Language Variation and Change*, *4*(1), 77–104.

Britain, D. (1997). Dialect contact and phonological reallocation: "Canadian Raising" in the English Fens. *Language in Society*, *26*, 15–46.

Britain, D. (2005). Where did New Zealand English come from? In A. Bell, R. Harlow, & D. Starks (Eds.), *Languages of New Zealand* (pp. 156–193). Wellington: Victoria University Press.

Britain, D. (2016). Sedentarism and nomadism in the sociolinguistics of dialect. In N. Coupland (Ed.), *Sociolinguistics: Theoretical debates* (pp. 217–241). Cambridge: Cambridge University Press.

Brookes, J., Hall, D., Cheshire, J., & Adger, D. (2017). Causal interrogative variation in Multicultural and traditional varieties of London English. *Queen Mary's OPAL*, *36*.

Brunstad, E., Røyneland, U., & Opsahl, T. (2010). Hip hop, ethnicity and linguistic practice in rural and urban Norway. In M. Terkourafi (Ed.), *The languages of global hip hop* (pp. 223–255). London: Continuum International.

Buerkner, P.-C. (2017). Brms: An R package for Bayesian multilevel models using Stan. *Journal of Statistical Software*, *80*(1), 1–28.

Börjesson, M. (2014). *The pronunciation of English by Somali L1 students in Sweden: Testing indications of phonetic transfer through Error Analysis and Contrastive Analysis.* Retrieved from `https://gupea.ub.gu.se/bitstream/2077/35054/1/gupea_2077_35054_1.pdf` (Unpublished BA thesis)

Cardoso, A. (2015). *Dialectology, phonology, diachrony: Liverpool English realisations of* PRICE *and* MOUTH (Unpublished doctoral dissertation). University of Edinburgh, Edinburgh.

Cardoso, A., Levon, E., Sharma, D., Watt, D., & Ye, Y. (2019). Inter-speaker variation and the evaluation of British English accents in employment contexts. In *Proceedings of the 19th International Congress of Phonetic Sciences.* Melbourne.

Carignan, C., Shosted, R., Shih, C., & Rong, P. (2011). Compensatory articulation in american english nasalized vowels. *Journal of Phonetics*, *39*(4), 668–682.

Chambers, J. (2002). Dynamics of dialect convergence. *Journal of Sociolinguistics*, *6*(1), 117–130.

Chambers, J. (2003). *Sociolinguistic theory* (2nd ed.). Oxford: Blackwell.

Cheshire, J. (2002). Sex and gender in variationist research. In J. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The handbook of language variation and change* (1st ed., pp. 423–443). Oxford: Blackwell.

Cheshire, J. (2005). Syntactic variation and beyond: Gender and social class variation in the use of discourse-new markers. *Journal of Sociolinguistics*, *9*(4), 479–508.

Cheshire, J. (2007). Discourse variation, grammaticalisation and stuff like that. *Journal of Sociolinguistics*, *11*(2), 155–193.

Cheshire, J. (2013). Grammaticalisation in social context: The emergence of a new English pronoun. *Journal of Sociolinguistics*, *17*(5), 608–633.

Cheshire, J. (2016). Epilogue: the future of discourse-pragmatic variation and change research. In H. Pichler (Ed.), *Discourse-pragmatic variation and change in English: new methods and insights* (pp. 252–266). Cambridge: Cambridge University Press.

Cheshire, J. (2020). Taking the longer view: Explaining Multicultural London English and Multicultural Paris French. *Journal of Sociolinguistics*, *24*(3), 308–327.

Cheshire, J., Adger, D., & Fox, S. (2013). Relative *who* and the actuation problem. *Lingua*, *126*(1), 51–77.

Cheshire, J., & Fox, S. (2009). *Was/were* variation: A perspective from London. *Language variation and change*, *21*(1), 1–38.

Cheshire, J., Fox, S., Kerswill, P., & Torgersen, E. (2008a). Ethnicity, friendship network and social practices as the motor of dialect change : linguistic innovation in London. *Sociolinguistica*, *22*, 1–23.

Cheshire, J., Fox, S., Kerswill, P., & Torgersen, E. (2008b). *Linguistic innovators: the English of adolescents in London: full research report*. Swindon: ESRC. (RES-000-23-0680)

Cheshire, J., Fox, S., Kerswill, P., & Torgersen, E. (2013). Language contact and language change in the multicultural metropolis. *Revue française de linguistique appliquée*, *18*(2), 63–76.

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.

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

Clyne, M. (2000). Lingua franca and ethnolects in Europe and beyond. *Sociolinguistica*(14), 83–89.

Cobain, I. (2018, June). London drill rap group banned from making music due to threat of violence. *The Guardian*. Retrieved from `https://www.theguardian.com/uk-news/2018/jun/15/london-drill-rap-gang-banned-from-making-music-due-to-threat-of-violence`

Conway, E. A. (2008). *An analysis of Somali pronunciation errors* (Unpublished master's thesis). Hamline University, Saint Paul, Minnesota.

Core, C. (2012). Assessing phonological knowledge. In E. Hoff (Ed.), *Research methods in child language* (pp. 300–316). Chichester: Wiley-Blackwell.

Cornips, L. (2008). Loosing grammatical gender in Dutch: The result of bilingual acquisition and/or an act of identity? *International Journal of Bilingualism*, *12*(1–2), 105–124.

Coupland, N. (1980). Style-shifting in a Cardiff work-setting. *Language in Society*, *9*(1), 1–12.

Cutler, C. (2003). "Keepin' it real": White hip-hoppers' discourses of language, race, and authenticity. *Journal of Linguistic Anthropology*, *13*(2), 211–233.

Cutler, C., & Røyneland, U. (2015). Where the fuck am I from? hip-hop youth and the (re)negotiation of language and identity in Norway and the US. In *Language, youth and identity in the 21st century: linguistic practices across urban spaces* (pp. 139–164). Cambridge: Cambridge University Press.

Davies, M. (2004). *BYU-BNC*. Retrieved from `https://corpus.byu.edu/bnc/` (Based on the British National Corpus from Oxford University Press)

De Decker, P., & Nycz, J. (2013). The technology of conducting sociolinguistic interviews. In C. Mallinson, B. Childs, & G. Van Herk (Eds.), *Data collection in sociolinguistics* (pp. 118–130). New York: Routledge.

Denis, D. (2019, November). Enregisterment, resistance and the spread of linguistic alterity in the most multicultural city in the world. In *Urban-Rural Language Research Workshop.*

Desmeules-Trudel, F., & Brunelle, M. (2018). Phonotactic restrictions condition the realization of vowel nasality and nasal coarticulation: Duration and airflow measurements in Québécois French and Brazilian Portuguese. *Journal of Phonetics*, *69*, 43–61.

Deterding, D. (2006). The North Wind versus a Wolf: short texts for the description and measurement of English pronunciation. *Journal of the International Phonetic Association*, *36*(2), 187–196.

Dienes, Z. (2008). *Understanding psychology as a science: an introduction to scientific and statistical inference*. New York: Palgrave Macmillan.

Di Paolo, M., Yaeger-Dror, M., & Beckford Wassink, A. (2011). Analyzing vowels. In *Sociophonetics: a student's guide* (pp. 87–106). London; New York: Routledge.

Disner, S. F. (1980). Evaluation of vowel normalization procedures. *The Journal of the Acoustical Society of America*, *67*(1), 253–261.

Dodd, B. J., Holm, A., Hua, Z., & Crosbie, S. (2003). Phonological development: a normative study of British English-speaking children. *Clinical linguistics & phonetics*, *17*(8), 617–643.

Dodd, B. J., Hua, Z., Crosbie, S., Holm, A., & Ozanne, A. (2002). *Diagnostic evaluation of articulation and phonology (DEAP)*. London: Psychological Corporation.

D'Onofrio, A. (2015). Persona-based information shapes linguistic perception: Valley Girls and California vowels. *Journal of Sociolinguistics*, *19*(2), 241–256.

Dorleijn, M., Mous, M., & Nortier, J. (2015). Urban youth speech styles in Kenya and the Netherlands. In J. Nortier & B. A. Svendsen (Eds.), *Language, youth and identity in the 21st century: linguistic practices across urban spaces* (pp. 271–289). Cambridge: Cambridge University Press.

Dorleijn, M., & Nortier, J. (2013). Multiethnolects: Kebabnorsk, Perkerdansk, Verlan, Kanakensprache, Straattaal, etc. In P. Bakker & Y. Matras (Eds.), *Contact languages: a comprehensive guide* (pp. 229–272). Berlin: de Gruyter.

Drager, K. (2011). Sociophonetic variation and the lemma. *Journal of Phonetics*, *39*(4), 694–707.

Drager, K. (2016). Constructing style: phonetic variation in quotative and discourse particle *like*. In H. Pichler (Ed.), *Discourse-pragmatic variation and change in English: new methods and insights* (pp. 232–251). Cambridge: Cambridge University Press.

Drummond, R. (2016). (Mis)interpreting urban youth language: white kids sounding black? *Journal of Youth Studies*, *20*(5), 1–21.

Drummond, R. (2018a). Maybe it's a grime [t]ing: TH-stopping among urban British youth. *Language in Society*, *47*(2), 171–196.

Drummond, R. (2018b). *Researching urban youth language and identity*. Basingstoke: Palgrave Macmillan.

Duranti, A. (1997). *Linguistic anthropology*. New York: Cambridge University Press.

Duñabeitia, J. A., Crepaldi, D., Meyer, A. S., 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.

Ealing Grid for Learning. (2017, September). *Travellers in Ealing.* Retrieved from `https://www.egfl.org.uk/school-effectiveness/teaching-and-learning/equality-and-achievement/gypsy-roma-and-traveller-0`

Eckert, P. (1996). Vowels and nail polish: The emergence of linguistic style in the preadolescent heterosexual marketplace. In N. Warner, J. Ahlers, L. Bilmes, M. Oliver, S. Wertheim, & M. Chen (Eds.), *Gender and belief systems* (pp. 183–190). Berkeley: Berkeley Women and Language Group.

Eckert, P. (1998). Age as a sociolinguistic variable. In F. Coulmas (Ed.), *The handbook of sociolinguistics* (pp. 151–167). Oxford: Blackwell.

Eckert, P. (2000). *Linguistic variation as social practice: the linguistic construction of identity in Belten High*. Oxford: Blackwell.

Eckert, P. (2001). Style and social meaning. In P. Eckert & J. Rickford (Eds.), *Style and sociolinguistic variation* (pp. 119–126). Cambridge: Cambridge University Press.

Eckert, P. (2008). Where do ethnolects stop? *International Journal of Bilingualism*, *12*(1–2), 25–42.

Eckert, P. (2011). Where does the social stop? In J. Parrott (Ed.), *Language variation – European perspectives III* (pp. 13–29). Amsterdam: John Benjamins.

Eckert, P. (2012). Three waves of variation study: The emergence of meaning in the study of sociolinguistic variation. *Annual Review of Anthropology*, *41*, 87–100.

Eckert, P. (2013). Ethics in linguistic research. In R. Podesva & D. Sharma (Eds.), *Research methods in linguistics* (pp. 119–126). Cambridge: Cambridge University Press.

Eckert, P. (2016). Variation, meaning and social change. In N. Coupland (Ed.), *Sociolinguistics: theoretical debates* (pp. 68–85). Cambridge: Cambridge University Press.

Eckert, P., & McConnell-Ginet, S. (1992). Think practically and look locally:language and gender as community-based practice. *Annual Review of Anthropology*, *21*(1), 461–488.

Elan (version 5.9) [Computer software manual]. (2020). Nijmegen. Retrieved from `https://archive.mpi.nl/tla/elan`

Elvin, J., Williams, D., & Escudero, P. (2016). Dynamic acoustic properties of monophthongs and diphthongs in Western Sydney Australian English. *The Journal of the Acoustical Society of America*, *140*(1), 576–581.

Englund, K. T. (2018). Hypoarticulation in infant-directed speech. *Applied Psycholinguistics*, *39*(1), 67–87.

Evans, B. G., & Alshangiti, W. (2018). The perception and production of British English vowels and consonants by Arabic learners of English. *Journal of Phonetics*, *68*, 15–31.

Fabricius, A. H., Watt, D., & Johnson, D. E. (2009). A comparison of three speaker-intrinsic vowel formant frequency normalization algorithms for sociophonetics. *Language Variation and Change*, *21*(3), 413–435.

Filppula, M., Klemola, J., & Paulasto, H. (Eds.). (2009). *Vernacular universals and language contacts: Evidence from varieties of English and beyond*. London: Routledge.

Fletcher, J., & Harrington, J. (2001). High-rising terminals and fall-rise tunes in Australian English. *Phonetica*, *58*(4), 215–229.

Forrest, J., & Wolfram, W. (2019). The status of (ING) in African American language: A quantitative analysis of social factors and internal constraints. *American Speech*, *94*(1), 72–90.

Fought, C. (2003). *Chicano English in context*. London: Palgrave Macmillan.

Fought, C. (2006). *Language and ethnicity*. Cambridge: Cambridge University Press.

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

Fox, S. (2007). *The demise of Cockneys? Language change in London's "traditional" East End* (Unpublished doctoral dissertation). University of Essex.

Fox, S. (2015). *The new Cockney: new ethnicities and adolescent speech in the traditional East End of London*. New York: Palgrave Macmillan.

Fox, S., Khan, A., & Torgersen, E. (2011). The emergence and diffusion of Multicultural English. In F. Kern & M. Selting (Eds.), *Ethnic styles of speaking in European metropolitan areas* (pp. 19–44). Amsterdam: John Benjamins.

Fox Tree, J. (2010). Discourse markers across speakers and settings. *Language and linguistics compass*, *4*(5), 269–281.

Franke, M., & Roettger, T. B. (2019, July). Bayesian regression modelling (for factorial designs): a tutorial. *PsyArXiv*. doi: 10.31234/osf.io/cdxv3

Freywald, U., Cornips, L., Ganuza, N., Nistov, I., & Opsahl, T. (2015). Beyond verb second – a matter of novel information-structural effects? Evidence from Norwegian, Swedish, German and Dutch. In J. Nortier & B. A. Svendsen (Eds.), *Language, youth and identity in the 21st century: linguistic practices across urban spaces* (pp. 73–92). Cambridge: Cambridge University Press.

Freywald, U., Mayr, K., Özçelik, T., & Wiese, H. (2011). Kiezdeutsch as a multiethnolect. In F. Kern & M. Selting (Eds.), *Ethnic styles of speaking in European metropolitan areas* (pp. 45–73). Amsterdam: John Benjamins.

Fridland, V. (2003). 'Tie, tied and tight': The expansion of/ai/-monophthongization in African-American and European-American speech in Memphis, Tennessee. *Journal of Sociolinguistics*, *7*(3), 279–298.

Furman, R., & Özyürek, A. (2007). Development of interactional discourse markers: Insights from Turkish children's and adults' oral narratives. *Journal of Pragmatics*, *39*(10), 1742–1757.

Gahl, S., & Baayen, R. H. (2019). Twenty-eight years of vowels: Tracking phonetic variation through young to middle age adulthood. *Journal of Phonetics*, *74*, 42–54.

Gates, S. M. (2018). Why the long FACE?: Ethnic stratification and variation in the London diphthong system. *University of Pennsylvania. Working Papers in Linguistics*, *24*(2), 39.

Gates, S. M. (2019). *Language variation and ethnicity in a multicultural East London secondary school* (Unpublished doctoral dissertation). Queen Mary University of London, London.

Gelman, A., & Carlin, J. (2014). Beyond power calculations: Assessing Type S (sign) and Type M (magnitude) errors. *Perspectives on Psychological Science*, *9*(6), 641–651.

Georgakopoulou, A. (2007). *Small stories, interaction and identities*. Amsterdam: John Benjamins.

Georgakopoulou, A. (2008). "On MSN with buff boys": Self-and other-identity claims in the context of small stories. *Journal of Sociolinguistics*, *12*(5), 597–626.

Giegerich, H. J. (1992). *English phonology: an introduction*. Cambridge: Cambridge University Press.

Gnevsheva, K. (2020). The role of style in the ethnolect: Style-shifting in the use of ethnolectal features in first-and second-generation speakers. *International Journal of Bilingualism*, *24*(4), 861–880.

Goffman, E. (1981). *Forms of talk*. Philadelphia: University of Pennsylvania Press.

Granlund, S. (2015). *Speech communication strategies in older children: acoustic-phonetic and linguistic adaptations to a hearing-impaired peer* (Unpublished doctoral dissertation). University College London, London.

Greater London Authority. (2012). *2011 Census snapshot: Country of birth, passports held and national identity* (resreport No. CIS2012-13). GLA Intelligence Census Information Scheme.

Greater London Authority. (2013). *2011 Census Snapshot: Ethnic diversity indices for wards* (resreport No. CIS2013-02). GLA Intelligence Census Information Scheme.

Green, J., & Bloome, D. (2005). Ethnography and ethnographers of and in education: a situated perspective. In J. Flood, S. B. Heath, & D. Lapp (Eds.), *Handbook of research on teaching literacy through the communicative and visual arts* (pp. 181–202). New York: Macmillan Publishers.

Green, L. J. (2002). *African American English: a linguistic introduction*. Cambridge: Cambridge University Press.

Gumperz, J. J. (1964). Linguistic and social interaction in two communities. *American anthropologist*, *66*(6), 137–153.

Habib, R. (2014). Vowel variation and reverse acquisition in rural Syrian child and adolescent language. *Language Variation and Change*, *26*(1), 45–75.

Haddican, B., Foulkes, P., Hughes, V., & Richards, H. (2013). Interaction of social and linguistic constraints on two vowel changes in Northern England. *Language Variation and Change*, *25*(3), 371–403.

Hall-Lew, L. (2009). *Ethnicity and phonetic variation in a San Francisco neighborhood* (Unpublished doctoral dissertation). Stanford University.

Hansen, G. F., & Pharao, N. (2010). Prosody in the Copenhagen multiethnolect. In P. Quist & B. A. Svendsen (Eds.), *Multilingual urban Scandinavia: new linguistic practices* (pp. 79–95). Bristol: Multilingual Matters.

Heine, B., & Song, K.-A. (2011). On the grammaticalization of personal pronouns. *Journal of Linguistics*, *47*(3), 587–630.

Heritage, J. (1984). A change-of-state token and aspects of its sequential placement. In J. M. Atkinson & J. Heritage (Eds.), *Structures of social action* (pp. 299–345). Cambridge: Cambridge University Press.

Hoff, E., & Rumiche, R. (2012). Studying children in bilingual environments. In E. Hoff (Ed.), *Research methods in child language* (pp. 300–316). Chichester: Wiley-Blackwell.

Hoffman, M. F., & Walker, J. A. (2010). Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English. *Language Variation and Change*, *22*(1), 37–67.

Holm, J. A. (2000). *An introduction to pidgins and creoles*. Cambridge: Cambridge University Press.

Holmes-Elliott, S. (2015). *London calling: assessing the spread of metropolitan features in the southeast* (Unpublished doctoral dissertation). University of Glasgow, Glasgow.

Home Office. (2015). *Serious Crime Act 2015, c.9.* Retrieved from `https://www.legislation.gov.uk/ukpga/2015/9/contents/enacted`

Hopper, P. J., & Traugott, E. C. (2003). *Grammaticalization* (2nd ed.). Cambridge: Cambridge University Press.

Horvath, B., & Sankoff, D. (1987). Delimiting the Sydney speech community. *Language in Society*, *16*(2), 179–204.

Ilbury, C. (2020). *Beyond the offline* (Unpublished doctoral dissertation). Queen Mary University of London, London.

Jaspers, J. (2008). Problematizing ethnolects: Naming linguistic practices in an Antwerp secondary school. *International Journal of Bilingualism*, *12*(1–2), 85–103.

Kallmeyer, W., & Keim, I. (2003). Linguistic variation and the construction of social identity in a German-Turkish setting. A a case study of an immigrant youth-group in Mannheim, Germany. In J. Androutsopoulous & A. Georgakopoulou (Eds.), *Discourse constructions of youth identities* (pp. 29–46). Amsterdam: John Benjamins.

Karrebæk, M. S. (2015). Arabs, Arabic and urban languaging: Polycentricity and incipient enregisterment among primary school children in Copenhagen. In L. M. Madsen, M. S. Karrebæk, & J. S. Møller (Eds.), *Everyday languaging: collaborative research on the language use of children and youth* (pp. 19–48). Berlin: De Gruyter Mouton.

Kern, F. (2015). Turkish German. *Language and Linguistics Compass*, *9*(5), 219–233.

Kerswill, P. (1996). Children, adolescents, and language change. *Language variation and change*, *8*(2), 177–202.

Kerswill, P. (2014). The objectification of 'Jafaican': the discoursal embedding of Multicultural London English in the British media. In J. Androutsopoulos (Ed.), *The media and sociolinguistic change* (pp. 427–456). Berlin: De Gruyter.

Kerswill, P. (2016, December). *Multicultural London English: a new dialect, a style, or both?* Talk presented at Humanities Research Centre, Sheffield Hallam University.

Kerswill, P., Cheshire, J., Fox, S., & Torgersen, E. (2013). English as a contact language: the role of children and adolescents. In *English as a contact language* (pp. 258–282). Cambridge: Cambridge University Press.

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.

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

Kerswill, P., & Williams, A. (2000b). Dialect recognition and speech community focusing in new and old towns in England: The effect of dialect levelling, demography and social networks. In D. Long & D. Preston (Eds.), *A handbook of perceptual dialectology* (Vol. 2, pp. 178–207). Amsterdam: Benjamins.

Kerswill, P., & Williams, A. (2005). New towns and koineization: Linguistic and social correlates. *Linguistics: An Interdisciplinary Journal of the Language Sciences*, *43*(5), 1023–1048.

Khan, A. (2006). *A sociolinguistic study of Birmingham English: Language variation and change in a multi-ethnic British community* (Unpublished doctoral dissertation). University of Lancaster.

Khattab, G. (2007). Variation in vowel production by English-Arabic bilinguals. In J. Cole & J. L. Hualde (Eds.), *Papers in Laboratory Phonology IX* (pp. 383–410). Berlin: Mouton de Gruyter.

Khattab, G. (2013). Phonetic convergence and divergence strategies in English-Arabic bilingual children. *Linguistics: An Interdisciplinary Journal of the Language Sciences*, *51*(2), 439–472.

Kießling, R., & Mous, M. (2004). Urban youth languages in Africa. *Anthropological Linguistics*, *46*(3), 303–341.

Kirkham, S. (2013). *Ethnicity, social practice and phonetic variation in a Sheffield secondary school* (Unpublished doctoral dissertation). University of Sheffield, Sheffield.

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

Kirschner, S., & Tomasello, M. (2010). Joint music making promotes prosocial behavior in 4-year-old children. *Evolution and Human Behavior*, *31*(5), 354–364.

Koffi, E. (2012). Intelligibility assessment and the vowel space: an instrumental phonetic account of the production of English lax vowels by Somali speakers. In J. Levis & K. LeVelle (Eds.), *Social factors in pronunciation acquisition: proceedings of the 3rd annual pronunciation in second language learning and teaching conference* (pp. 216–232).

Kotsinas, U.-B. (1988). Immigrant children's Swedish: A new variety? *Journal of Multilingual and Multicultural Development*, *9*(1-2), 129–140.

Kotsinas, U.-B. (2001). Pidginization, creolization, and creoloids in Stockholm, Sweden. In N. Smith & T. Veenstra (Eds.), *Creolization and contact* (pp. 125–155). Amsterdam: John Benjamins.

Kruschke, J. (2015). *Doing Bayesian data analysis: a tutorial with R, JAGS, and Stan* (2nd ed.). Boston: Academic Press.

Kuhl, P. K., Tsao, F.-M., & Liu, H.-M. (2003). Foreign-language experience in infancy: Effects of short-term exposure and social interaction on phonetic learning. *Proceedings of the National Academy of Sciences*, *100*(15), 9096–9101.

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

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26.

Kärkkäinen, E. (2003). *Epistemic stance in English conversation: a description of its interactional functions, with a focus on 'I think'*. Amsterdam: John Benjamins.

Labov, W. (1966). *The social stratification of English in New York City*. Washington, DC: Center for Applied Linguistics.

Labov, W. (1972). Some principles of linguistic methodology. *Language in Society*, *1*(1), 97–120.

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

Labov, W. (1994). *Principles of linguistic change* (Vol. 1: internal factors). Oxford: Blackwell.

Labov, W. (2001a). The anatomy of style-shifting. In P. Eckert & J. Rickford (Eds.), *Style and sociolinguistic variation* (pp. 85–108). Cambridge: Cambridge University Press.

Labov, W. (2001b). *Principles of linguistic change* (Vol. 2: social factors). Oxford: Blackwell.

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

Labov, W. (2013). Preface: The acquisition of sociolinguistic variation. *Linguistics*, *51*(2), 247–250.

Ladefoged, P. (2005). *Vowels and consonants: an introduction to the sounds of languages* (2nd ed.). Oxford; Malden, MA: Blackwell.

Lambert, B. (2018). *A student's guide to Bayesian statistics*. London: SAGE Publications.

Lanza, E. (1992). Can bilingual two-year-olds code-switch? *Journal of Child Language*, *19*(3), 633–658.

Leeser, R. (2016). *English indices of deprivation 2015* (Tech. Rep.). GLA Intelligence. Retrieved from `https://airdrive-secure.s3-eu-west-1.amazonaws.com/london/dataset/indices-of-deprivation/2019-10-09T14:23:00/indices-deprivation-2015.pdf?X-Amz-Algorithm=AWS4-HMAC-SHA256&X-Amz-Credential=AKIAJJDIMAIVZJDICKHA/20201113/eu-west-1/s3/aws4_request&X-Amz-Date=20201113T101913Z&X-Amz-Expires=300&X-Amz-Signature=9b3caa6506828ea480d58bf03384f08c1075bbe62fa41704d22f4499c3facc6e&X-Amz-SignedHeaders=host`

Lehtonen, H. (2011). Developing multiethnic youth language in Helsinki. In F. Kern & M. Selting (Eds.), *Ethnic styles of speaking in European metropolitan areas* (pp. 291–318). Amsterdam: John Benjamins.

Lehtonen, H. (2015). *Tyylitellen: Nuorten kielelliset resurssit ja kielen sosiaalinen indeksisyys monietnisessä helsingissä* (Unpublished doctoral dissertation). University of Helsinki, Helsinki.

Le Page, R. B., & Tabouret-Keller, A. (1985). *Acts of identity: Creole-based approaches to language and ethnicity*. Cambridge: Cambridge University Press.

Levey, S. (2016). The role of children in the propagation of discourse-pragmatic change: insights from the acquisition of quotative variation. In H. Pichler (Ed.), *Discourse-pragmatic variation and change in English: new methods and insights* (pp. 160–182). Cambridge: Cambridge University Press.

Levon, E. (2013). Ethnography and recording interaction. In *Research methods in linguistics* (pp. 195–214). Cambridge: Cambridge University Press.

Levon, E. (2016). Gender, interaction and intonational variation: The discourse functions of high rising terminals in london. *Journal of Sociolinguistics*, *20*(2), 133–163.

Maddieson, I. (1984). *Patterns of sounds*. Cambridge: Cambridge University Press.

Madsen, L. M. (2011a). Late modern youth style in interaction. In F. Kern & M. Selting (Eds.), *Ethnic styles of speaking in European metropolitan areas* (pp. 265–290). Amsterdam: John Benjamins.

Madsen, L. M. (2011b). Social status relations and enregisterment in contemporary Copenhagen. *King's College London working papers in urban language and literacies*, *72*.

Madsen, L. M., Karrebæk, M. S., & Møller, J. S. (Eds.). (2015). *Everyday languaging: collaborative research on the language use of children and youth*. Berlin: De Gruyter Mouton.

Mangara, M. (2017). *Ealing JSNA 2017: population characteristics.* Retrieved from `https://www.ealing.gov.uk/download/downloads/id/13887/`

Marzo, S., & Ceuleers, E. (2011). The use of Citétaal among adolescents in Limburg: the role of space appropriation in language variation and change. *Journal of Multilingual and Multicultural Development*, *32*(5), 451–464.

Matras, Y. (2009). *Language contact*. Cambridge: Cambridge University Press.

Mayr, R., Morris, J., Mennen, I., & Williams, D. (2017). Disentangling the effects of long-term language contact and individual bilingualism: The case of monophthongs in Welsh and English. *International Journal of Bilingualism*, *21*(3), 245–267.

McCarthy, K., Mahon, M., Rosen, S., & Evans, B. G. (2014). Speech perception and production by sequential bilingual children: A longitudinal study of Voice Onset Time acquisition. *Child development*, *85*(5), 1965–1980.

McCarthy, M. (2003). Talking back: "small" interactional response tokens in everyday conversation. *Research on Language and Social Interaction*, *36*(1), 33–63.

McGregor, K. K., Friedman, R. M., Reilly, R. M., & Newman, R. M. (2002). Semantic representation and naming in young children. *Journal of Speech, Language, and Hearing Research*, *45*(2), 332–346.

Mendoza-Denton, N. (2002). Language and identity. In J. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The handbook of language variation and change* (1st ed., pp. 475–499). Oxford: Blackwell.

Mendoza-Denton, N. (2008). *Homegirls: language and cultural practice among Latina youth gangs*. Malden, MA: Blackwell.

Milroy, J., Milroy, L., Hartley, S., & Walshaw, D. (1994). Glottal stops and Tyneside glottalization: Competing patterns of variation and change in British English. *Language variation and change*, *6*(3), 327–357.

Milroy, L. (2007). Off the shelf or under the counter? on the social dynamics of sound changes. In C. M. Cain & G. Russom (Eds.), *Managing chaos: strategies for identifying change in English. Studies in the history of the English language III* (pp. 147–172). Berlin: Mouton de Gruyter.

Møller, J. S. (2009). *Poly-lingual interaction across childhood, youth and adulthood* (Unpublished doctoral dissertation). University of Copenhagen, Copenhagen.

Morris, J. (2017). Sociophonetic variation in a long-term language contact situation: /l/-darkening in Welsh-English bilingual speech. *Journal of Sociolinguistics*, *21*(2), 183–207.

Mott, B. (2012). Traditional Cockney and Popular London Speech. *Dialectologia*, *9*, 69–94.

Mufwene, S. S. (2003). *The ecology of language evolution*. Cambridge: Cambridge University Press.

Munro, M. J. (1993). Productions of English vowels by native speakers of Arabic: Acoustic measurements and accentedness ratings. *Language and Speech*, *36*(1), 39–66.

Nance, C. (2015). 'New' Scottish Gaelic speakers in Glasgow: A phonetic study of language revitalisation. *Language in Society*, *44*(4), 553–579.

Nance, C. (2020). Bilingual language exposure and the peer group: Acquiring phonetics and phonology in Gaelic medium education. *International Journal of Bilingualism*, *24*(2), 360–375.

Nardy, A., Chevrot, J.-P., & Barbu, S. (2013). The acquisition of sociolinguistic variation: Looking back and thinking ahead. *Linguistics*, *51*(2), 255–284.

Nardy, A., Chevrot, J.-P., & Barbu, S. (2014). Sociolinguistic convergence and social interactions within a group of preschoolers: A longitudinal study. *Language Variation and Change*, *26*(3), 273–301.

Nassenstein, N., & Hollington, A. (Eds.). (2015). *Youth language practices in Africa and beyond*. Berlin: de Gruyter.

Newlin-Łukowicz, L. (2013). TH-stopping in New York City: Substrate effect turned ethnic marker? *University of Pennsylvania Working Papers in Linguistics*, *19*(2), 151–160.

Newlin-Łukowicz, L. (2015). Language variation in the diaspora: Polish immigrant communities in the U.S. and the U.K. *Language and Linguistics Compass*, *9*(8), 332–346.

Newlin-Łukowicz, L. (2016). Co-occurrence of sociolinguistic variables and the construction of ethnic identities. *Lingua*, *172–173*, 100–115.

Newman, M. (2010). Focusing, implicational scaling, and the dialect status of New York Latino English. *Journal of Sociolinguistics*, *14*(2), 207–239.

Nicenboim, B., Roettger, T. B., & Vasishth, S. (2018). Using meta-analysis for evidence synthesis: The case of incomplete neutralization in German. *Journal of Phonetics*, *70*, 39–55.

Nicenboim, B., & Vasishth, S. (2016). Statistical methods for linguistic research: Foundational ideas – Part II. *Language and Linguistics Compass*, *10*(11), 591–613.

Nicoladis, E. (2006). Cross-linguistic transfer in adjective-noun strings by preschool bilingual children. *Bilingualism*, *9*(1), 15–32.

Nortier, J. (2001). *Murks en Straattaal*. Amsterdam: Prometheus.

Nortier, J., & Dorleijn, M. (2008). A Moroccan accent in Dutch: A sociocultural style restricted to the Moroccan community? *International Journal of Bilingualism*, *12*(1–2), 125–142.

Office for National Statistics Annual Population Survey. (2015). *Percentage of people on low income, borough.* Retrieved from `https://data.london.gov.uk/dataset/percentage-people-low-income-borough`

Ohala, J. J. (1994). The frequency code underlies the sound-symbolic use of voice pitch. In L. Hinton, J. Nichols, & J. J. Ohala (Eds.), *Sound symbolism* (pp. 325–347). Cambridge: Cambridge University Press.

Opsahl, T. (2009). "Wolla I swear" this is typical for the conversational style of adolescents in multiethnic areas in Oslo. *Nordic Journal of Linguistics*, *32*(2), 221–244.

O'Shannessy, C. (2014). Researching children's acquisition of sociolinguistic competence. In J. Holmes (Ed.), *Research methods in sociolinguistics: a practical guide* (pp. 304–324). Hoboken: John Wiley.

Oxbury, R. (2016). *Situational style-shifting in Multicultural London English* (Unpublished master's thesis). Queen Mary University of London, London.

Payne, A. (1976). *The acquisition of the phonological system of a second dialect* (Unpublished doctoral dissertation). University of Pennsylvania, Philadelphia.

Pichler, H. (Ed.). (2016a). *Discourse-pragmatic variation and change in English: new methods and insights*. Cambridge: Cambridge University Press.

Pichler, H. (2016b). Introduction: discourse-pragmatic variation and change. In H. Pichler (Ed.), *Discourse-pragmatic variation and change in English: new methods and insights* (pp. 1–18). Cambridge: Cambridge University Press.

Pichler, H. (2016c). Uncovering discourse-pragmatic innovations: *Innit* in Multicultural London English. In H. Pichler (Ed.), *Discourse-pragmatic variation and change in English: new methods and insights* (pp. 59–85). Cambridge: Cambridge University Press.

Pichler, P., & Williams, N. (2016). Hipsters in the hood: Authenticating indexicalities in young men's hip-hop talk. *Language in Society*, *45*(4), 557–581.

Podesva, R. J. (2007). Phonation type as a stylistic variable: The use of falsetto in constructing a persona. *Journal of Sociolinguistics*, *11*(4), 478–504.

Podesva, R. J., D'Onofrio, A., Van Hofwegen, J., & Kim, S. K. (2015). Country ideology and the California Vowel Shift. *Language Variation and Change*, *27*(2), 157–186.

Pratt, T. (2018). Embodying toughness: LOT-raising, /l/-velarization, and retracted articulatory setting. *University of Pennsylvania Working Papers in Linguistics*, *24*(2), 107–116.

Prieto, P., & Borràs-Comes, J. (2018). Question intonation contours as dynamic epistemic operators. *Natural Language & Linguistic Theory*, *36*(2), 563–586.

Quist, P. (2005). *Stilistike praksisser i storbyens heterogene skole – en etnografisk og sociolingvistik undersøgelse af sproglig variation [Stylistic practices in the heterogeneous school in the big city – an ethnographic and sociolinguistic study of linguistic variation]* (Unpublished doctoral dissertation). University of Copenhagen, Copenhagen.

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

Quist, P., & Svendsen, B. A. (2010). Introduction. In P. Quist & B. A. Svendsen (Eds.), *Multilingual urban Scandinavia: new linguistic practices* (pp. xiii–xxiii). Bristol: Multilingual Matters.

R Core Team. (2020). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Rampton, B. (1995). *Crossing: language and ethnicity among adolescents*. London: Longman.

Rampton, B. (2006). *Language in late modernity: interaction in an urban school.* Cambridge: Cambridge University.

Rampton, B. (2011). Style contrasts, migration and social class. *Journal of Pragmatics*, *43*(5), 1236–1250.

Rampton, B. (2015). Contemporary urban vernaculars. In J. Nortier & B. A. Svendsen (Eds.), *Language, youth and identity in the 21st century: linguistic practices across urban spaces* (pp. 24–44). Cambridge: Cambridge University Press.

Rickford, J. (2014). Situation: Stylistic variation in sociolinguistic corpora and theory. *Language and Linguistics Compass*, *8*(11), 590–603.

Rickford, J., & McNair-Knox, F. (1994). Addressee- and topic-influenced style shift: a quantitative sociolinguistic study. In D. Biber & E. Finegan (Eds.), *Sociolinguistic perspectives on register* (pp. 235–276). Oxford: Oxford University Press.

Roberts, J. (1997a). Acquisition of variable rules: a study of (-t, d) deletion in preschool children. *Journal of Child Language*, *24*(2), 351–372.

Roberts, J. (1997b). Hitting a moving target: Acquisition of sound change in progress by Philadelphia children. *Language Variation and Change*, *9*(2), 249–266.

Roberts, J. (2002). Child language variation. In J. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The handbook of language variation and change* (1st ed., pp. 333–349). Oxford: Blackwell.

Rodríguez Louro, C. (2016). Quotatives across time: West Australian English then and now. In H. Pichler (Ed.), *Discourse-pragmatic variation and change in English: new methods and insights* (pp. 139–159). Cambridge: Cambridge University Press.

Rosenfelder, I., Fruehwald, J., Evanini, K., Seyfarth, S., Gorman, K., Prichard, H., & Yuan, J. (2014). *Fave (forced alignment and vowel extraction) program suite v1.2.2.* Retrieved from `https://github.com/JoFrhwld/FAVE` doi: 10.5281/zenodo.22281

Routarinne, S. (1997). Kertomuksen rakentaminen. In L. Tainio (Ed.), *Keskustelunanalyysin perusteet* (pp. 138–155). Tampere: Vastapaino.

Rudd, A. (2018, April). *Launch of the Serious Violence Strategy.* Speech. Retrieved from `https://www.gov.uk/government/speeches/home-secretary-announces-launch-of-the-serious-violence-strategy`

Saeed, J. (1999). *Somali.* Amsterdam: John Benjamins.

Schegloff, E. A. (1982). Discourse as an interactional achievement: some uses of "uh huh" and other things that come between sentences. In D. Tannen (Ed.), *Analyzing discourse, text and talk* (pp. 71–93). Washington, DC: Georgetown University Press.

Schegloff, E. A. (1997). Practices and actions: Boundary cases of other-initiated repair. *Discourse Processes*, *23*(3), 499–545.

Schiffrin, D. (1990). The management of a cooperative self in argument: the role of opinions and stories. In A. Grimshaw (Ed.), *Conflict talk* (pp. 241–259). Cambridge: Cambridge University Press.

Schilling, N. (2013). *Sociolinguistic fieldwork*. Cambridge: Cambridge University Press.

Schleef, E. (2013). Glottal replacement of /t/ in two British capitals: Effects of word frequency and morphological compositionality. *Language Variation and Change*, *25*(2), 201–223.

Schleef, E., Meyerhoff, M., & Clark, L. (2011). Teenagers' acquisition of variation: A comparison of locally-born and migrant teens' realisation of English (ing) in Edinburgh and London. *English World-Wide*, *32*(2), 206–236.

Sebba, M. (1993). *London Jamaican: language systems in interaction*. London: Longman.

Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics in Language Teaching*, *10*(1-4), 209–232.

Sharma, D. (2011). Style repertoire and social change in British Asian English. *Journal of Sociolinguistics*, *15*(4), 464–492.

Sharma, D. (2018). Style dominance: Attention, audience, and the 'real me'. *Language in Society*, *47*(1), 1–31.

Sharma, D., & McCarthy, K. (2018). Attentional load and style control. *University of Pennsylvania Working Papers in Linguistics*, *24*(2), 127–136.

Sharma, D., & Sankaran, L. (2011). Cognitive and social forces in dialect shift: Gradual change in London Asian speech. *Language Variation and Change*, *23*, 399–428.

Sidnell, J. (2010). *Conversation analysis: an introduction*. Chichester: Wiley-Blackwell.

Silva-Corvalán, C. (2014). *Bilingual language acquisition: Spanish and English in the first six years*. Cambridge: Cambridge University Press.

Silverstein, M. (2003). Indexical order and the dialectics of sociolinguistic life. *Language & communication*, *23*(3–4), 193–229.

Smith, J., & Durham, M. (2019). *Sociolinguistic variation in children's language: Acquiring community norms*. Cambridge: Cambridge University Press.

Smith, J., Durham, M., & Richards, H. (2013). The social and linguistic in the acquisition of sociolinguistic norms: Caregivers, children, and variation. *Linguistics: An Interdisciplinary Journal of the Language Sciences*, *51*(2), 285–324.

Smith, J., & Holmes-Elliott, S. (2018). The unstoppable glottal: tracking rapid change in an iconic British variable. *English Language and Linguistics*, *22*(3), 323–355.

Snell, J. (2010). From sociolinguistic variation to socially strategic stylisation. *Journal of Sociolinguistics*, *14*(5), 630–656.

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), 174–215.

Sóskuthy, M., Foulkes, P., Hughes, V., Hay, J., & Haddican, B. (2015). Word-level distributions and structural factors codetermine GOOSE fronting. In *Proceedings of the 18th International Congress of Phonetic Sciences.* Glasgow.

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

Stanford, J. N. (2008). Child dialect acquisition: New perspectives on parent/peer influence. *Journal of Sociolinguistics*, *12*(5), 567–596.

Starks, D., & Bayard, D. T. (2002). Individual variation in the acquisition of postvocalic /r/: Day care and sibling order as potential variables. *American Speech*, *77*(2), 184–194.

Stuart-Smith, J., Timmins, C., & Tweedie, F. (2007). 'Talkin' Jockney'? Variation and change in Glaswegian accent. *Journal of Sociolinguistics*, *11*(2), 221–260.

Svendsen, B. A. (2015). Language, youth and identity in the 21st century: content and continuations. In J. Nortier & B. A. Svendsen (Eds.), *Language, youth and identity in the 21st century: linguistic practices across urban spaces* (pp. 3–23). Cambridge: Cambridge University Press.

Svendsen, B. A., & Røyneland, U. (2008). Multiethnolectal facts and functions in Oslo, Norway. *International Journal of Bilingualism*, *12*(1–2), 63–83.

Szakay, A. (2012). Voice quality as a marker of ethnicity in New Zealand: From acoustics to perception. *Journal of Sociolinguistics*, *16*(3), 382–397.

Tagliamonte, S. (1998). *Was/were* variation across the generations: View from the city of York. *Language Variation and Change*, *10*(2), 153–191.

Tagliamonte, S. (2012). *Variationist sociolinguistics: change, observation, interpretation*. Malden, MA: Wiley-Blackwell.

Tagliamonte, S. (2016a). Antecedents of innovation: exploring general extenders in conservative dialects. In H. Pichler (Ed.), *Discourse-pragmatic variation and change in English: new methods and insights* (pp. 115–138). Cambridge: Cambridge University Press.

Tagliamonte, S. (2016b). *Teen talk: the language of adolescents*. Cambridge: Cambridge University Press.

Tagliamonte, S., & D'Arcy, A. (2009). Peaks beyond phonology: Adolescence, incrementation, and language change. *Language*, *85*(1), 58–108.

Tagliamonte, S., & Molfenter, S. (2007). How'd you get that accent?: Acquiring a second dialect of the same language. *Language in Society*, *36*(5), 649–675.

Tagliamonte, S., & Roberts, C. (2005). So weird; so cool; so innovative: The use of intensifiers in the television series *Friends*. *American Speech: A Quarterly of Linguistic Usage*, *80*(3), 280–300.

Tamminga, M. (2018). Modulation of the following segment effect on English coronal stop deletion by syntactic boundaries. *Glossa: a journal of general linguistics*, *3*(1), 1–27.

Tanner, J., Sonderegger, M., & Stuart-Smith, J. (2019). Vowel duration and the voicing effect across dialects of English. *Toronto Working Papers in Linguistics*, *41*(1), 1–13.

Thomas, E. R. (2002). Instrumental phonetics. In J. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The handbook of language variation and change* (1st ed., pp. 168–200). Oxford; Malden, MA: Blackwell.

Thomas, E. R. (2011). *Sociophonetics: an introduction*. Basingstoke, Hampshire; New York: Palgrave Macmillan.

Thomason, S. G., & Kaufman, T. (1988). *Language Contact, Creolization and Genetic Linguistics*. Berkeley: University of California Press.

Tomasello, M., & Stahl, D. (2004). Sampling children's spontaneous speech: how much is enough? *Journal of Child Language*, *31*(1), 101–121.

Torgersen, E., Gabrielatos, C., & Hoffmann, S. (2017). A corpus-based analysis of the pragmatic marker *you get me*. In E. Friginal (Ed.), *Studies in corpus-based sociolinguistics* (1st ed.). Oxford: Routledge.

Torgersen, E., Gabrielatos, C., Hoffmann, S., & Fox, S. (2011). A corpus-based study of pragmatic markers in London English. *Corpus Linguistics and Linguistic Theory*, *7*(1), 93–118.

Torgersen, E., & Szakay, A. (2012). An investigation of speech rhythm in London English. *Lingua*, *122*(7), 822–840.

Traugott, E. C. (1989). On the rise of epistemic meanings in English: An example of subjectification in semantic change. *Language*, *65*(1), 31–55.

Trudgill, P. (1986). *Dialects in contact*. Oxford: Blackwell.

Trudgill, P. (2004). *New-Dialect Formation: The Inevitability of Colonial Englishes*. Edinburgh: Edinburgh University Press.

Urban Dictionary. (2020). Retrieved from `https://www.urbandictionary.com/` (Last accessed 13th November 2020)

Van Hofwegen, J., & Wolfram, W. (2010). Coming of age in African American English: A longitudinal study. *Journal of Sociolinguistics*, *14*(4), 427–455.

van Son, R., & Pols, L. (1992). Formant movements of Dutch vowels in a text, read at normal and fast rate. *The Journal of the Acoustical Society of America*, *92*(1), 121–127.

Vasishth, S., & Nicenboim, B. (2016). Statistical methods for linguistic research: Foundational ideas – part I. *Language and Linguistics Compass*, *10*(8), 349–369.

Vasishth, S., Nicenboim, B., Beckman, M. E., Li, F., & Kong, E. J. (2018). Bayesian data analysis in the phonetic sciences: A tutorial introduction. *Journal of Phonetics*, *71*, 147–161.

Vaughn, C., & Kendall, T. (2018). Listener sensitivity to probabilistic conditioning of sociolinguistic variables: The case of (ING). *Journal of Memory and Language*, *103*, 58–73.

Verschik, A. (2007). Jewish Russian and the field of ethnolect study. *Language in Society*, *36*(2), 213–232.

Verschik, A. (2010). Ethnolect debate: evidence from Jewish Lithuanian. *International Journal of Multilingualism*, *7*(4), 285–305.

Vorperian, H. K., & Kent, R. D. (2007). Vowel acoustic space development in children: A synthesis of acoustic and anatomic data. *Journal of Speech, Language, and Hearing Research*, *50*(6), 1510–1545.

Waters, C. (2016). Practical strategies for elucidating discourse-pragmatic variation. In H. Pichler (Ed.), *Discourse-pragmatic variation and change in English: new methods and insights* (pp. 41–56). Cambridge: Cambridge University Press.

Watt, D., & Fabricius, A. (2002). Evaluation of a technique for improving the mapping of multiple speakers' vowel spaces in the F1–F2 plane. *Leeds Working Papers in Linguistics and Phonetics*, *9*, 159–173.

Weinreich, U., Labov, W., & Herzog, M. I. (1968). Empirical foundations for a theory of language change. In W. P. Lehmann & Y. Malkiel (Eds.), *Directions for historical linguistics: a symposium* (pp. 95–188). Austin, Texas: Texas University Press.

Wells, J. C. (1982a). *Accents of English* (Vol. 2: The British Isles). Cambridge: Cambridge University Press.

Wells, J. C. (1982b). *Accents of English* (Vol. 1: an Introduction). Cambridge: Cambridge University Press.

Wenger, E. (1998). *Communities of practice: learning, meaning, and identity*. Cambridge: Cambridge University Press.

Wiese, H. (2009). Grammatical innovation in multiethnic urban Europe: New linguistic practices among adolescents. *Lingua*, *119*(5), 782–806.

Wiese, H. (2013). What can new urban dialects tell us about internal language dynamics? The power of language diversity. *Linguistische Berichte*(19), 207–245.

Wiese, H. (2020). Contact in the city. In R. Hickey (Ed.), *The Wiley handbook of language contact* (2nd ed.). Oxford: Wiley-Blackwell.

Williams, A., & Kerswill, P. (1999). Dialect levelling: continuity vs. change in Milton Keynes, Reading and Hull. In P. Foulkes & G. Docherty (Eds.), *Urban voices: accent studies in the British Isles* (pp. 141–162). London: Arnold.

Williams, D., & Escudero, P. (2014). A cross-dialectal acoustic comparison of vowels in Northern and Southern British English. *The Journal of the acoustical society of America*, *136*(5), 2751–2761.

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

Wolfson, N. (1976). Speech events and natural speech: some implications for sociolinguistic methodology. *Language in Society*, *5*(2), 189–209.

Wong, A. W. (2012). The lowering of raised-THOUGHT and the low-back distinction in New York City: Evidence from Chinese Americans. *University of Pennsylvania Working Papers in Linguistics*, *18*(2), 157–166.

Wormald, J. (2015). Dynamic variation in 'Panjabi-English': analysis of F1 and F2 trajectories for FACE and GOAT. In *Proceedings of the 18th International Congress of Phonetic Sciences.* Glasgow, UK.

Wormald, J. (2016). *Regional variation in Panjabi-English* (Unpublished doctoral dissertation). University of York, York.

Youssef, V. (1991). Variation as a feature of language acquisition in the Trinidad context. *Language Variation and Change*, *3*(1), 75–101.

Zellou, G., & Tamminga, M. (2014). Nasal coarticulation changes over time in Philadelphia English. *Journal of Phonetics*, *47*, 18–35.