Title: Effects of language background on the development of sociolinguistic awareness: the perception of accent variation in monolingual and multilingual 5-7 year old children.

Short Title: Effects of language background on the development of sociolinguistic awareness

Submitted to: SPECIAL ISSUE - CONSTANCY AND VARIATION IN SPEECH

Authors:

Bronwen G. Evans[1] & Gisela Tomé Lourido[1,2]

[1]Dept of Speech Hearing and Phonetic Sciences, University College London, U.K.

[2]School of Languages, Cultures & Societies, University of Leeds, U.K.

Corresponding author

Dr Bronwen G. Evans

Dept of Speech, Hearing and Phonetic Sciences,

University College London,

Chandler House,

2, Wakefield Street,

London,

WC1N 1PF

bronwen.evans@ucl.ac.uk

Tel. 44 (0)20 7679 4089

ABSTRACT

As a result of complex international migration patterns, listeners in large urban centres such as London, UK, likely encounter large amounts of variation in spoken language. However, although dealing with variation is crucial to communication, relatively little is known about how the ability to do this develops. Still less is known about how this might be affected by language background. The current study investigates whether early experience with variation, specifically growing up bilingually in London, affects accent categorization. Sixty children (30 monolingual, 30 bilingual) aged 5-7 years old, were tested in their ability to comprehend and categorize talkers in 2 out of 3 accents; a home, unfamiliar regional and unfamiliar foreign-accented variety. All children demonstrated high, above chance performance in the comprehension task, but language background significantly affected children's ability to categorize talkers. Bilinguals were able to categorize talkers in all accent conditions, but although all children were able to understand the talkers, monolingual children were only able to categorize talkers in the home-foreign accent condition. Overall, the results are consistent with an approach in which gradient representations of accent variation emerge alongside an understanding of how variation is used meaningfully within a child's environment.

[195 words]

ACKNOWLEDGEMENTS: This paper presents data that was collected by Amy Madigan as part of her undergraduate dissertation at University College London in partial fulfilment of the BSc Speech Communication. The writing of this paper has benefitted from discussions with many people, in particular Paul Foulkes, Ella Jeffries & Josiane Riverin-Coutlée, and participants of the symposium on "Children's understanding and use of sociolinguistic variation in language acquisition" at the 14th International Congress for the Study of Child Language, organized by Laura Wagner & Cynthia Clopper, where aspects of this work were presented. We would also like to thank all the parents and children and their schools who generously gave of their time to take part in this research.

CONFLICT OF INTEREST STATEMENT: The authors have no conflicts of interest to declare.

INTRODUCTION

Contact between speakers of different languages and accent backgrounds is a long-established phenomenon, but recently, national and international migration patterns have become more complex (e.g., Vertovec, 2007). In turn, this has created increasingly diverse multidialectal and multilingual communities, where people regularly come into contact not just with speakers of different regional accents but also different languages (e.g., Cheshire et al., 2011). Children growing up in such diverse communities are thus likely to encounter speakers from a wide variety of backgrounds, native and non-native.

Urban centres like London are a prime example of this. For instance, in the recent 2011 U.K. census (Office for National Statistics, 2016), over a third of London's population reported that they had been born overseas (vs. 13% nationally) with over a fifth (22.1%) reporting that they spoke a main language other than English (vs. 8% nationally). This means that children growing up monolingually in the dominant language, English, are likely to be exposed not just to different local varieties such as Popular London and Multicultural London English (cf. Cheshire et al., 2011), but also to different national and international varieties of English, as well as so-called World Englishes (e.g., Indian English) and foreign-accented speech (i.e., speakers with English as a second language). Children growing up bilingually will likely be exposed to still more variation. They will be exposed to different accents of English (McCarthy et al., 2013; 2014), but also to additional language, accent and talker variation. These children may be raised as simultaneous bilinguals (i.e., exposed to both languages from birth, e.g., Sundara & Scuttellaro, 2011; Mayr et al., 2014) or as sequential bilinguals, who are initially exposed primarily to the family or community language, and then learn the host community language when they enter school or nursery education at around 3-4 years of age (Darcy & Krüger, 2012; Xu et al., 2013; McCarthy et al., 2014). This means that on starting nursery or school education, sequential and possibly also simultaneous bilinguals are likely to have exposure not just to different languages and native-speaker varieties of English (e.g., through their monolingual peers, teachers) but also to non-native, i.e., foreign-accented varieties (e.g., through grandparents; see e.g., McCarthy et al., 2013).

Yet research on how young children's language development is affected by their immediate language environment, and in particular, the role that variation plays in language development, is scarce. What little we do know is largely limited to studies of children growing up monolingually in communities where children interact with talkers who speak similar varieties of the same language (e.g., Wagner et al., 2014; Jones et al., 2017; Weatherhead et al., 2016; Kinzler & DeJesus, 2012). In these studies, young children aged 5-6yrs have been shown to be unable to distinguish an unfamiliar regional accent from their own accent, but to be able to tell apart their home accent from a foreign accent. One explanation for this finding is that, at least initially, children's representations are organized along a quantitatively defined

continuum, anchored by their 'home' accent. This is thought to form the core set of their experiences, with regional accents treated as intermediate between the Home and Second-Language accent (Wagner et al., 2014). The current study aims to extend this work to examine the effects of early experience, specifically whether or not a child is raised monolingually or bilingually within a highly diverse speech community (London, UK), on the ability to extract and use patterns of variation in the speech signal to categorize home, unfamiliar regional (Yorkshire) and unfamiliar foreign-accented (Singapore English) talkers. Specifically, we aim to investigate whether children who are exposed to more variation develop a more detailed, gradient representation of variation within their community

1.1 The role of experience in the processing of accent variation

Previous work on the role of variation in speech processing has focussed predominantly on its effect on comprehension. For adults, unfamiliar-accented speech has been shown to have an initial processing cost (see Cristia et al., 2012 for a review); listeners are slower to decide whether a word is real or not (Floccia et al., 2006), are more adversely affected by background noise (Clopper & Bradlow, 2008) and make more errors in deciding whether or not a sentence is true or false (Adank et al., 2009). However, with exposure they are able to rapidly adapt to these differences (Maye et al., 2008), perhaps after only a 1-2 sentence lengths (Clarke & Garrett, 2004). Adaptation effects have also been found in long-term contact situations; native speakers of British English who have lived in the USA are better able to process American English speakers' flapped productions of medial-/t/, [ɾ], e.g., in city, where they would typically have [t] (Scott & Cutler, 1984). Likewise, Adank et al. (2009) found that Glaswegians living in Glasgow who were regularly exposed to Standard Southern British English (SSBE) were able to make true/false judgements on sentences presented in SSBE with the same speed and accuracy as when they were presented in their native accent, Glaswegian, but that SSBE speakers living in London with little or no experience of Glaswegian were slower and less accurate with Glaswegian than SSBE.

However, the mechanism by which listeners are able to adapt remains unclear. Some have interpreted these results as suggesting that speech processing is highly flexible and that, with enough exposure, listeners are able to adjust their category representations with relative ease (e.g., Johnson, 2007). In contrast, others have argued that abstract, underlying categories have limited flexibility, and that instead, short-term adaptation effects arise as a result of top-down changes, e.g., becoming better able to map variants to one's own underlying categories (Iverson & Evans, 2009). For example, in a series of experiments, Shaw et al. (2018) investigated whether Australian listeners' categorization of their own vowels, along with those in 4 unfamiliar varieties of English, was facilitated by short-term multi-talker exposure. In contrast to previous studies, (e.g., Maye et al., 2010; Norris et al., 2003) the results

demonstrated that listeners' performance on the categorization task was unaffected by exposure. Based on these findings, the authors hypothesize that listeners' representations can tolerate a good range of natural variation (cf. Evans & Iverson, 2007), and that this ability to "hear through" variation in order to recover the phonological category may be what is fundamental to maintaining perceptual flexibility.

Evidence from studies investigating children's perception of accented speech provides further support for this view. Nathan et al. (1998) tested monolingual English children, aged 4-7yrs old, born and raised in London in their ability to recognize words produced in their own (London) and a Glaswegian accent, as measured in a paradigm in which they had to repeat and then define a word. As expected, all children found it harder to understand words in the unfamiliar Glaswegian accent, but there were developmental differences in how children responded; older children were more accurate in defining the word and also gave significantly more phonological responses (i.e., repeated the word in their own accent) than did younger children, who instead tended to respond using the Glaswegian pronunciation. For the majority (61%) of these phonetic responses, children were also unable to define the word, indicating that they had not achieved lexical access. One possibility is that this shift from phonetic to phonological processing enabled older children to better map the incoming signal to their own, underlying stored representations. That is, with more experience of variation, older children were able to retrieve constant phonological and lexical information, i.e., achieve phonological constancy (Best et al., 2009; Nathan et al., 1998).

1.2 Accent variation, categorization & the role of experience

Language acquisition must, however, also involve distinguishing the linguistic from the social, i.e., understanding how a given variable is used meaningfully within a speech community. For example, adults make inferences about people based on the speech signal (e.g., Bayard et al., 2001; Preston, 1989) and these judgements can change their basic categorization of speech stimuli (Strand, 1999). Indeed, listeners' stereotypes about gender, as activated by the faces and voices of speakers, have been shown to alter perception of a phonetic continuum such that a female face or voice causes the perceived boundary between /s/-/ʃ/ to shift up in frequency, whilst a male face or voice causes the boundary to shift down in frequency (Strand, 1999). Moreover, faces rated by independent judges as highly masculine or feminine, have been shown to lead to greater changes than those rated as less masculine and feminine (Strand, 1999). Similar effects have been found for regional accent. Niedzielski (1999) showed that evoking a specific region by writing the name of a place on a response sheet, caused listeners who were familiar with the accent spoken there to categorize vowels according to the way in which they are produced in that regional accent (cf., Hay & Drager, 2010). Together, these findings indicate that socially constructed beliefs play a

role in basic phonological categorization, and that social information affects how listeners perceive and categorize speech sounds (see Campbell-Kibler, 2016 for an excellent discussion of how social and linguistic information might be integrated in models of cognitive processing).

Such a role for socially-constrained information in speech processing is consistent with an exemplar-based model of speech perception (e.g., Goldinger, 1997; Johnson, 1997; Hawkins & Smith, 2001). This approach argues that linguistic knowledge is based on detailed memory representations which simultaneously encode non-linguistic (e.g., who was speaking, in what situation) and linguistic (e.g., segmental features, voice quality, pitch) information (Docherty & Foulkes, 2013; Munson, 2010). Any abstract representations are thought to arise from common factors amongst the many memories a speaker-listener has (Hawkins, 2003), such that knowledge of phonetic and social variation derives from the clustering of similar forms, with knowledge of socially-structured variation emerging where linguistic and social differences coincide (Docherty & Foulkes, 2013). Such a model supports a probabilistic approach to phonetic learning and categorization (Foulkes & Docherty, 2006); as a speaker-listener builds up a library of traces, he/she develops sensitivity to the statistical properties of the usage of different variables, i.e., different patterns of usage. It is from this awareness that abstract, phonological categories emerge (Pierrehumbert, 2003; Pierrehumbert & Gross, 2003). Thus, a listener who has experience with different British accents may be able to recognize northern-accented speech by mapping it onto similar stored exemplars produced by northern talkers, and recognize southern-accented speech by mapping it onto similar stored exemplars produced by southern talkers. Specifically, being able to make hypotheses about where a speaker is from, may enable a listener to predict which phonetic variants are likely to be used and consequently, enable them to tune-in to the talker more easily. For example, hearing tokens of BATH produced with the northern English [æ] may 'prime' a listener not only to expect more tokens of BATH words produced with this vowel, but also to expect to hear other northern English variants, such as STRUT and FOOT words produced with the same vowel, [ʊ], even if they haven't yet heard tokens of this variable. Likewise, by adjusting their expectations of how /s/-/ʃ/ will be produced according to whether or not listeners think they are listening to a male or female talker (Strand, 1999), listeners may be better able to map the incoming signal to their own, underlying abstract representations of these categories.

In terms of child language development, we might hypothesize that as children build up a bank of detailed representations through experience with speech that is rich in language-specific and potentially talker-specific phonetic detail, there is a developmental progression from specific to more abstract knowledge of sound structure (Munson, 2010). Acquisition might therefore in part be driven by frequency of exposure, with those differences that are both phonetically large and experienced more frequently (e.g., male-female differences based on pitch range) emerging earlier than those which are more arbitrary (e.g.,

ethnicity, social class) and which might be experienced relatively rarely within the speech community (see Foulkes & Docherty, 2006).

Recent studies investigating children's ability to detect accent-related differences between talkers, lend experimental support to this view. For example, children aged 5-6yrs old are able to group talkers according to their accent, but only when the phonological differences between the accents are maximized (Girard et al., 2008; Wagner et al., 2014). That is, they are able to successfully categorize native vs. foreign-accented talkers, but are unable to reliably tell apart talkers of their own accent and another unfamiliar regional accent where the acoustic-phonetic differences between the talkers are relatively small. They also understand that a speaker's accent can index where someone lives (Weatherhead et al., 2016) and can make meaningful social associations with a talker based on his/her accent. For example, they can link familiar (e.g., a woman dressed in a business suit) and unfamiliar (e.g., a woman dressed in a kimono) cultural items with either a Home vs. Foreign or a Regional vs. Foreign accent contrast, but again, not a Home vs. Regional contrast (Wagner et al., 2014). One explanation for these findings is that at least initially, children's representations are organized along a quantitatively defined continuum, anchored by their 'home' accent, which forms the core set of their experiences, with regional accents treated as intermediate between the Home and Foreign accent (Wagner et al., 2014). With more experience, older children are able to reliably distinguish their own from another unfamiliar regional variety, (Floccia et al., 2009), and make judgements in line with general social stereotypes. For example, children aged 9-10yrs in Illinois & Tennessee USA, judged Northern-accented American English speakers to be more intelligent than those with Southern-accented American English (Kinzler & DeJesus, 2012).

However, almost all of this work has focussed on monolingual children growing up in relatively homogenous communities. Yet one prediction of an exemplar-based approach is that with greater exposure to variation, and a more detailed bank of representations, more gradient, fine-grained representations will emerge. One possibility then, is that children growing up in a multidialectal or multilingual environment like inner-city London, might be better able to extract and use variation in the speech signal, e.g., to group talkers, than those growing up in more homogenous environments. Few studies have investigated this question and the results are mixed. For example, Jeffries (2016) found that monolingual English children aged 3-4yrs, growing up in Yorkshire but with a parent from outside Yorkshire (i.e., who spoke with a different accent from that of the community) were better able to group talkers according to accent when the target variables contained different rather than the same phonemes (see also Floccia et al., 2009). Interestingly, children with parents from outside Yorkshire were also better able to recognize a familiar talker (in this case, a familiar teacher) using phonetic accent features. In contrast, Beck (2016) found that early exposure to multiple regional varieties, again through a parent with a different accent from that used in the

local community, did not lead to better discrimination between different regional accents. One possibility is that this is because the task – an ABX discrimination task in which children heard and compared isolated words – was too easy and thus did not enable more fine-grained differences between the different listener groups to emerge. Indeed, differences in awareness did emerge, although perhaps not in the direction expected; children with parents from outside the community (Outsiders) performed more poorly at identifying whether or not an accent was local or non-local than their peers with parents born and raised in the community (Insiders), and being an Outsider negatively influenced the ability to recognize regional accent as a kind of linguistic variation between speakers. One possibility is that Outsiders could hear the variation in the input, but given their exposure to different varieties within the same place, did not equate speaking with a regional accent as being from a specific place (cf. Weatherhead et al., 2016). This lends further support to the notion that an understanding of social variation is derived from experience, and that how and when this emerges will vary between individuals as a result of early experience with language.

In this study we examine whether children raised monolingually or bilingually in a diverse community in London, differ in their ability to categorize talkers in three conditions; (1) home (London) vs. unfamiliar regional accent (Yorkshire), (2) home vs. unfamiliar foreign accent (Singapore English), (3) unfamiliar regional vs. unfamiliar foreign accent. Based on previous work we hypothesize that though our monolinguals may outperform those in previous studies as a result of their increased exposure to accent variation (e.g., Wagner et al., 2014), they will likely find this task challenging. In contrast, we hypothesize that bilinguals will be better able to do this task, not just because they are exposed to more variation (regional and foreign-accented speech, as well as different languages) but also because this variation does meaningful work in their community, e.g., in signalling family relationships, and that this will mean that they are better able to categorize talkers even when the differences between them are relatively small (e.g., in the home vs. unfamiliar regional accent condition).

METHOD

1. Participants

Sixty children aged 5;4 – 7;11 yrs (mean 6;7 yrs) were recruited from North London primary schools in the neighbouring London boroughs of Brent and Barnet. None of the children who participated had any reported speech, hearing or language impairments, and prior to testing, all children obtained age-appropriate scores on the British Picture Vocabulary Scale 3 (BPVS 3; Dunn & Dunn, 2009). Ethical approval for the study was granted by the UCL Ethics committee and all children completed the study on an opt-in basis[1].

Before testing, parents completed a language background questionnaire which provided basic details of their child's language background and exposure to other foreign languages or accents. Thirty were

native monolingual English speakers (13 females, 17 males; 5;06 – 7;10 yrs, mean = 6;05yrs) and 30 were either bi- or multilingual (18 females, 12 males, 5;04-7;11yrs, mean 6;08yrs). Of these, only 5 were multilingual, i.e., they had acquired more than one additional language and so for ease, we will refer to this group as bilingual. Bilingual children were either simultaneous bilinguals (N=13), who had acquired both English and any additional languages simultaneously from birth, or sequential bilinguals (N=17) who had initially acquired their home language and had first received consistent exposure to English when attending nursery or pre-school. Bilingual children spoke a range of different languages (Gujarati, Urdu, Hindi, Polish, Romanian, Spanish, Portuguese), but none were bilingual in any of the other official languages spoken in Singapore, (i.e., Mandarin, Malay, Tamil), or in any other Chinese language (e.g., Cantonese). The majority of children had been born and raised within the area of North London in which they were tested. A small number of bilingual children (N=5) had either been born or spent time living outside of the U.K., but had moved or returned to the U.K. to attend nursery at age 3yrs where they had lived ever since.

Information from the questionnaires was also used to inform the assignment of children to the different experimental conditions. For example, if a child fit the criteria for selection but his/her parents reported that he/she had had regular exposure to a variety of Northern English, the child was placed in the Home-Foreign condition. Regular exposure was defined as being through a close or near-relative (e.g., a parent, grandparent), or a caregiver with whom the child regularly interacted (e.g., a childminder). Only one child in the bilingual condition was reported to have regular exposure to Yorkshire English and this child thus completed the Home-Foreign condition. This ensured that as far as possible, all children across all conditions, were equally unfamiliar with both the regional and foreign accent.

2. Stimuli

Twelve speakers (2 male, 2 female per accent) recorded the stimuli for both the comprehension and categorization tasks. They were encouraged to read in an engaging style, as if reading a story to a child and all speakers were judged by a trained phonetician to have a representative accent. All speakers were recorded in a sound attenuated booth at Chandler House, UCL, at a sampling rate of 44.1 kHz, 16-bit resolution, using a Rode NT-1A condenser microphone connected to a PC via an RME Fireface UC processor. Recordings were made in stereo and later converted to mono. Sentence stimuli were manually extracted and saved to individual .wav files with boundaries placed as close as possible to the onset and offset of each sentence. Stimuli were then band-pass filtered at 60Hz-20kHz, equalized for intensity at 70 dB and a 100ms buffer added to the beginning and end of each file. All processing was completed in Praat (Boersma & Weenink, 2017).

Four monolingual native speakers from North London recorded the Home accent (i.e., London) stimuli. The Home accent was Popular London (Wells, 1982), which is characterized by the use of glottal stops word-medially, TH-fronting, vocalized L, labialized R and differences in the realization of some vowels as compared with SSBE, e.g., the PRICE and MOUTH vowels. We chose this accent rather than SSBE because this has been shown to be highly intelligible, even for speakers of other regional accents (Pinet et al., 2015), and because although the children would have been familiar with SSBE, Popular London is more representative of the accent used by them and those in their local community. Speakers were aged 24-49yrs (mean = 31yrs) and had been born and raised in North London in the London boroughs of Camden and Islington. None had spent any significant amount of time living outside North London and were resident in the area at the time of recording. Although these speakers were from a different part of North London to the children who participated in the study, the accent is very similar to that used in the children's local communities, both of which border Camden.

The regional accent was Yorkshire English and was recorded by 4 monolingual native speakers from Yorkshire. Speakers were aged 22-50yrs (mean = 30yrs). Of these, 3 speakers (1 male, 2 female) had been born and raised in West Yorkshire whilst the other speaker had been born and raised in North Yorkshire. All speakers were recorded in London. Three speakers had been living in London for 3 years, having moved there from Yorkshire to study at university, whilst the other had lived in London and the surrounding area for 15 years. Like other northern accents, Yorkshire English differs from southern English accents in production of BATH, which is produced with a short vowel, [æ], and STRUT, which is produced with either the FOOT vowel, [ʊ] (i.e., no STRUT-FOOT split) or a raised vowel similar to schwa, [ə]. Our speakers used both of these variants of the STRUT vowel, as is common in speakers of northern Englishes who have lived in the south of England (e.g., Evans & Iverson, 2007). Likewise, although our speakers also used some other features of Yorkshire English, e.g., a lowered /ɔː/ vowel, sometimes phonetically close to [ɒː] (Wells, 1982) they produced other vowels with more standard rather than regionally-specific variants. For example, in these recordings they all produced FACE with a diphthong, [eɪ], rather than a monophthong, [eː] and GOAT, as [əʊ] (raised onset) or [əʊ], rather than as a monophthong [oː]. Thus, our speakers maintained key variants typical of their home community, which they used alongside some features of what Wells (1982) refers to as northern Near-RP features.

The foreign accent was Singaporean English and was recorded by 4 native speakers of Singaporean English aged 22-26yrs (mean = 23yrs). All speakers had been born and raised in Singapore and had been living in London for 4yrs at the time of the recordings, where they had moved for university. All speakers had acquired Mandarin initially, but had been exposed to English from an early age and had completed their education in Singapore bilingually in English and Mandarin. Key features include what has

traditionally been referred to as syllable-timed rhythm (Trudgill & Hannah, 2002), avoidance of schwa in unstressed syllables (Wells, 1982), neutralization of vowel length distinctions (Deterding, 2005), TH-stopping (Bao, 1998) and consonant cluster reduction (Lim, 2004). Our speakers displayed all of these features and also differed from native English speakers in their use of connected speech features, such as linking and intrusive /r/, which are absent from SSE (Trudgill & Hannah, 2002), and realization of word final consonants, which in SSE are typically voiceless, leading to minimal pairs like “niece” and “knees” in SSBE becoming homophones in SSE (Lim, 2004).

Stimuli were selected to include key phonetic differences between the different accents but were not intended to be highly confusable. Differences between London and SSE affected many different phonetic features, but all sentences were also designed to include key differences between London and Yorkshire English, i.e., a BATH or STRUT vowel (see Appendix 1 for a full set of target and distractors). The sentence stimuli for the comprehension task consisted of six, 1-2 sentence phrases, four of which were themed to include characters from The Gruffalo (Donaldson & Scheffler, 1999). The stimuli for the categorization task consisted of twelve 1-2 sentence passages. Of these, 6 were taken from the popular children’s book Mrs Plug the Plumber (Ahlberg & Wright, 1980) and 6 were original. Vocabulary was selected to be age-appropriate; average acquisition of each word was either at or below age 5yrs as detailed in the Age of Acquisition database (Kuperman et al., 2012).

3. Procedure

All testing took place in a quiet room in the child’s school and was carried out by the second author, who had grown up in North London and spoke with a North London accent. During the experiment, the child and experimenter were seated next to each other at a table. Stimuli were presented at a comfortable listening level via a laptop PC over headphones, worn by both the experimenter and child, using the Experiment MFC interface in Praat (Boersma & Weenink, 2017). The child gave his/her responses using picture response cards and these were simultaneously logged by the experimenter via the Experiment MFC interface (Boersma & Weenink, 2017). The child was seated so he/she was unable to see the computer screen. Presentation of stimuli was pseudo-randomized across participants and conditions. Children completed the Comprehension and then the Categorization task in a single testing session lasting 10-15mins.

Comprehension Task. The task was a 4-alternative forced choice task, in which children selected the picture which they thought best matched the sentence they had heard. Previous work has shown that adults are able to quickly tune-in to an unfamiliar accent and so the purpose of this task was both to familiarize children with the novel accent as well as to test that they were able to understand it, i.e., that any differences in performance in the categorization task were not because they couldn’t understand the talker.

Children completed 6 trials in the non-Home dialect, i.e., either the Regional (Yorkshire) or Foreign (SSE) accent, produced by 2 talkers (1 male, 1 female). Before beginning the task, children were given a small booklet which contained separate response sets for each of the 6 trials. Each set consisted of 4 picture cards (1 target, 3 distractors) which were attached by Velcro, with only one set visible during a given trial (Fig. 1). The position of the pictures in each set was randomized across participants to reduce any potential effects of side bias on picture selection. The experimenter explained to the child that he/she would hear someone speaking and that it was his/her job to listen carefully, select the picture that the speaker had asked for or described (i.e., remove it from the booklet) and give the picture card to the experimenter. For example, on hearing "Can you see the ball in the grass?" (Trial 4), children saw pictures of a doll, a teddy, a bowl and a ball in the grass and had to choose the picture with the ball (Fig. 1). To keep the child 'on task', the first, third and fifth trials were preceded by appropriate attention-grabbing audio files. For example, Trial 5 ("His favourite food is a glass of mouse milkshake"), was preceded by the sound of liquid being sucked through a straw. All clips were downloaded from the Internet and converted to mono .wav files in Praat (Boersma & Weenink, 2017). A trial was considered complete when a picture had been selected and given to the experimenter. Children were praised uniformly throughout.

[INSERT FIG. 1 about here]

Categorization Task. The task was divided into 2 parts; training in group membership (4 trials) and a test phase in which children completed 6 categorization trials, interspersed with 2 reminder training trials which followed the 2nd categorization trial.

Children were first introduced to two soft toys, the Mouse and Gruffalo's Child. Both are well-known characters from the children's picture book *The Gruffalo's Child* (Donaldson & Scheffler, 2001; also an animated short film) and were familiar to all the children. Each toy represented one of the two accents (i.e., Home and Foreign, Home and Regional, Regional and Foreign) with the assignment of character to accent counterbalanced across participants: both soft toys were on view during the whole of the categorization task. Children were told that the characters' families had got lost in the deep, dark wood (the setting for the story) and that their job was to help each of them find the correct family (i.e., Mouse or Gruffalo's Child). However, because the wood was so dark they couldn't see them, and so they would have to listen carefully to their voices instead. The children were then introduced to how the Mouse and Gruffalo's Child families talked. At the beginning of each training trial, the experimenter held up the appropriate soft toy, and said "This is what the Mouse/Gruffalo's Child family sounds like!". A sound file was then played whilst the experimenter held up and jiggled the appropriate soft toy. Children were trained in this way with 2 sentences (1 male, 1 female) from each accent (i.e., 4 different talkers) before continuing to the test phase, in which they completed a 2-alternative forced choice task where they responded whether the talker they heard belonged to the

Gruffalo's Child's or Mouse's family. Before beginning the test phase, children were reminded that they would hear a sentence from a character lost in the deep, dark, wood, and that their job was to say which family the speaker belonged to by selecting the appropriate picture (Mouse or Gruffalo's Child) from the response booklet. There were 6 test trials (3 per accent). For each accent (i.e., 3 trials), the speakers were 1 talker who was familiar from the training and 2 novel talkers. The familiar talker was also used in the reminder trials. As in the comprehension task, a test trial was considered complete when a picture had been selected and given to the experimenter. Children were praised uniformly throughout.

RESULTS

1. Comprehension

Table 1 shows performance (proportion correct) for each accent in the comprehension task (regional, foreign) for monolingual and bilingual children. All children performed very well at this task and there did not appear to be any large differences in overall performance between monolingual and bilingual children, though monolinguals had a higher mean comprehension score than bilinguals.

[INSERT TABLE 1 HERE]

These observations were confirmed in a series of analyses. First, a series of one-sample t-tests confirmed that all children performed above chance in all conditions; Regional accent, monolinguals - $t(14) = 25.58$, $p < 0.001$, bilinguals - $t(14) = 21.75$, $p < 0.001$: Foreign accent, monolinguals - $t(14) = 40.12$, $p < 0.001$, bilinguals - $t(14) = 25.66$, $p < 0.001$. Secondly, a univariate ANOVA with accent (regional, foreign) and language background (monolingual, bilingual) coded as independent variables, and proportion correct as the dependent variable, demonstrated that there were no main effects or interaction of accent or language background, $p > 0.05$, confirming that monolinguals and bilinguals performed similarly with the two accents.

Previous work has shown that children's identification of sentences presented in quiet, in a foreign accent is significantly poorer than in a home accent (Bent & Atagi, 2015), and so it might seem somewhat surprising that our children performed so well in this task. However, our task was designed to be relatively easy to enable children to tune into the accents and to build confidence in doing the more difficult categorization task. It is thus possible that children did experience some difficulties in processing the unfamiliar accents, but that the use of pictures may have enabled them to make use of context in selecting the correct answer (cf. Bent, 2014).

2. Categorization Task

a. Overall performance

Fig. 2 displays the mean proportion correct for each accent (home, regional, foreign) in each condition (Home-Regional, Home-Foreign, Regional-Foreign) split by language background (monolingual, bilingual). Responses were coded as correct when children selected the soft toy that matched that accent. That is, if children were trained that the Mouse's family spoke with the home (Popular London) accent and the Gruffalo spoke with the foreign (SSE) accent, answering that a speaker who speaks with a London accent belonged to the Mouse's family was coded as a correct response, whilst answering that he/she belongs to the Gruffalo's family (foreign accent) was coded as an incorrect response. Inspection of the data suggests that bilinguals do better than monolinguals overall, performing at near ceiling level in all conditions. In contrast, monolinguals appear to find the task more difficult; although they do better in the Home-Foreign condition than in the Home-Regional and Regional-Foreign conditions, they perform more poorly than bilinguals overall.

To assess children's overall performance, an overall categorization score (i.e., the mean categorization score across both dialects in the pair) was calculated for each child in each group (monolingual, bilingual) for each condition (Home-Regional, Home-Foreign, Regional-Foreign) and compared to chance. Bilinguals performed above chance in all conditions (Bonferroni-corrected for multiple comparisons); Home-Regional $t(9) = 10.94$, $p < 0.001$, Home-Foreign $t(9) = 10.73$, $p < 0.001$, Regional-Foreign $t(9) = 9.04$ $p < 0.001$. In contrast, monolinguals performed at above chance in the Home-Foreign condition [$t(9) = 4.01$ $p < 0.01$], but not in the Home-Regional and Regional-Foreign conditions, $p > 0.05$.

A mixed effect logistic regression model was built with the binomial response (correct/incorrect) as the dependent variable, condition (Home-Foreign, Foreign-Regional, Home-Foreign) and language background (monolingual, bilingual) as fixed factors and 'participant' and 'speaker' as crossed random effects. Treatment coding was used for both factors, with 'Home-Foreign' as the reference level for condition, and 'monolingual' as the reference level for language background. For monolingual listeners, the model revealed a significant contrast between the Home-Foreign ($M_{prop} = 0.683$) and Regional-Foreign ($M_{prop} = 0.467$) conditions and a marginal significant contrast between the Home-Foreign ($M_{prop} = 0.683$) and Home-Regional ($M_{prop} = 0.483$) conditions (Table 2). There was a significant effect of language background, which indicated that bilinguals performed better than monolinguals (Fig 1). The interaction between condition and language background was not significant. To investigate whether bilinguals differed from monolinguals in each condition, pairwise post-hoc tests were calculated using the lsmeans package (Lenth, 2016) in R (R Core Team, 2018), adjusting for multiple comparisons using the Tukey method. The tests indicated that bilinguals outperformed monolinguals in all conditions (Home-Foreign: $\beta = -1.262$, SE = 0.490, $z = -2.575$, $p = 0.01$; Foreign-Regional: $\beta = -1.883$, SE = 0.447, $z = -4.208$, $p < 0.001$; Home-Regional: $\beta = -1.958$, SE = 0.463, $z = -4.225$, $p < 0.001$).

[INSERT TABLE 2 HERE]

As noted by Girard et al., (2008; see also Floccia et al., 2009), measures of correct responses do not exclude potential effects of response bias, that is, the possibility that children might over- or underestimate the number of unfamiliar/untrained accented sentences in the categorization task. To assess whether or not children were able to detect the unfamiliar/untrained accent using a bias-free measure, separate measures of sensitivity (*A'*) and bias (B''_D) were computed using signal detection analysis (see Girard et al., 2008; Floccia et al., 2009). Sensitivity (*A'*) was used rather than *d'* (Grier, 1971) because it does not rely on the strict assumptions of normality and equal variance of signal and noise distributions. Additionally, unlike *d',* this measure, which varies between 0 and 1, can be computed for extreme values. Thus, a score of 1.0 represents a 100% correct hit-rate (i.e., no false alarms), and a value of 0.5 indicates chance performance. In this case, a hit was defined as correct categorization of the home accent in the Home-Foreign and Home-Regional conditions, and the regional accent in the Regional-Foreign condition. A false alarm occurred when the regional (Home-Regional condition) or foreign accent (Home-Foreign and Regional-Foreign conditions) was mistakenly identified as the home or regional accent. Response bias, measured as B''_D, is independent from *A'*, and ranges from -1 to 1. Negative values indicate a liberal response bias and positive values indicate a conservative bias (Donaldson, 1992).

[Insert Fig 2 about here]

Table 3 displays average sensitivity and bias for monolingual and bilingual children in each condition (Home-Regional, Home-Foreign , Regional-Foreign). The measure of sensitivity gives a similar pattern of results to the overall categorization score; bilingual children are highly sensitive in all conditions, but monolinguals have only above chance levels of sensitivity in the Home-Foreign condition. Potential differences in performance according to language background were investigated in a univariate ANOVA with language background and condition coded as between-subject variables. There was a significant main effect of language background, $F(1,49) = 62.34$, $p < 0.001$, confirming that bilinguals showed greater sensitivity to accent differences than monolinguals. Additionally, there was a main effect of condition, $F(2,49) = 7.01$, $p < 0.001$, and interaction of language background and condition, $F(2,49) = 5.47$, $p < 0.01$. Again, this reflects the finding that monolinguals had greater sensitivity in the Home-Foreign condition than in the Home-Regional or Regional-Foreign conditions, with bilinguals having similarly high sensitivity across all 3 conditions.

[INSERT TABLE 3 HERE]

In terms of bias, monolingual and bilingual children showed similar tendencies in their responses in the different conditions (Table 3). All children had negative bias in the Home-Regional condition, indicating that they had a tendency to answer “Home” (i.e., to select the soft toy associated with the London accent),

whilst in the Regional-Foreign condition they had positive bias, indicating a tendency to answer “Foreign” (i.e., to select the soft toy associated with SSE). Given the success with which children were able to complete the task in the Home-Foreign condition, it is not surprising that neither monolingual nor bilingual children showed evidence of bias in this condition. These observations were confirmed in a univariate ANOVA on bias, with language background and condition coded as between-subject variables. There was no significant main effect of language background or interaction of language background and condition. However, there was a significant effect of condition, $F(2,44) = 3.27$, $p < 0.05$, confirming that responses, regardless of language background, showed differing effects of bias in the 3 conditions. Thus, in the Home-Regional condition, all children, regardless of language background, had a bias towards answering that the talker had the ‘Home’ accent, whilst in the Regional-Foreign condition, all children had a bias towards answering that the talker had the ‘Foreign’ accent.

4. GENERAL DISCUSSION

Previous research with monolingual children has shown that the ability to explicitly categorize talkers according to regional accent begins to emerge at around 5 years, and that initially they can only successfully categorize speakers when the acoustic-phonetic differences between the accents are maximized (Floccia et al., 2009; Girard et al., 2008; Wagner et al., 2014). This led researchers to hypothesize that children have gradient representations of accent variation, with their Home accent, the one that is experienced the most, forming the core set of experiences and other varieties identified relative to this (Wagner et al., 2014). Children are thus thought to perform poorly in categorizing talkers from their home vs. an unfamiliar regional accent community because even though these accents might be phonetically quite different, they are interpreted by the children as being similar enough that they cannot be well differentiated. In contrast, foreign accents are hypothesized to go well beyond what children will accept as similar, and so they are able to categorize these as separate from their own, native accent (Wagner et al., 2014).

Based on these findings, our initial hypothesis was that with greater experience of variation, children would develop a more detailed, gradient representation of variation within their community, such that our children, growing up in a community where they are frequently exposed to regional and foreign-accented speech, would perform well on an accent categorization task. Specifically, we hypothesized that our monolingual children would perform better than those in previous studies, who were growing up in more homogenous communities (Wagner et al., 2014; Girard et al., 2008), and that our bilingual children, who are exposed to the greatest amount of variability, would perform still more successfully than both of these groups. Our results supported our second prediction but not the first. Bilingual children performed best and were able to categorize talkers successfully in all 3 conditions (Home-Regional, Home-Foreign, Regional-

Foreign). In contrast, although all children were able to understand the different accents, our monolingual children were only able to categorize talkers with above chance accuracy in the Home-Foreign condition, replicating findings from these previous studies but with a different population.

Why did our monolingual children perform more poorly on the categorization task than our bilinguals? One interpretation is that these results simply lend support for the gradient representation hypothesis outlined above. One could imagine that at this stage of development, children may be more sensitive to the broader and less predictable differences that typically occur in foreign-accented speech (Girard et al., 2008) and that for our monolingual children, with potentially less exposure to foreign-accented varieties than our bilingual children, these differences were particularly noticeable. However, the foreign accent in our study was SSE, an established World English variety, with systematic patterns of variation which were well exemplified by our talkers, who had acquired this variety in childhood (cf. Wagner et a., 2014). All our speakers were highly fluent and it is therefore highly unlikely that they differed from speakers of our other varieties in their consistency of phonological patterning.

Another possibility is that success at this task depends on experience with different accents. For example, in exemplar-theoretic models, adults have been hypothesized to be able to identify talkers according to their regional accent by mapping to stored memories of similar talkers (e.g., Johnson, 2007). The regional and foreign accents used in this study were selected to be unfamiliar to the children, and it is possible that our monolingual children were unable to successfully categorize talkers according to accent as a result of their inexperience with these particular accents. That is not to say that they heard the regional and home accent as equivalent – our monolinguals performed poorly in the regional vs. foreign accent condition but not the home vs. foreign accent condition – but rather that the regional accent was within the 'noise' tolerance for the home accent and that their sparse representations for the regional and foreign accent meant that they were unable to separate them reliably (Wagner et al., 2014). One obvious problem with this explanation though, is that our bilinguals were selected to be equally unfamiliar with the regional and foreign accents used here; although they likely had more experience of variation overall, they did not have experience with these particular accents. Likewise, although we did not conduct any tests of cognitive abilities, there were no significant differences in vocabulary in monolingual and bilingual children, with all achieving age-appropriate scores on the BPVS. It therefore seems unlikely that our monolingual children were unable to do this task either because it is too difficult for children at this age (cf. Jones et al., 2017) or because they had not experience with those accents.

Perhaps a better explanation is that successful categorization of talkers according to regional accent involves not only development of the ability to track acoustic-phonetic differences between talkers, consistent with the gradient representation hypothesis, but also an understanding of how patterns of

variation are used meaningfully within a community. In turn, this enables children to associate patterns of variation with a given attribute. In support of this, previous research has shown that bilinguals have enhanced performance on phonological tasks in comparison to their monolingual counterparts (Yelland et al., 1993; Campbell & Sais, 1995; Bialystok et al., 2003), enhanced attentional processing (Kovacs & Mehler, 2012) and greater linguistic flexibility (Bialystok, 1986). Bilinguals arguably also have a greater need to use linguistic information in order to navigate relationships within their community (see McCarthy et al., 2014). Increased exposure to variation in a community where such variation serves to differentiate talkers as belonging to different sections of that community may thus promote sensitivity to talker variation, such that children growing up bilingually in a diverse, multilingual community develop the ability to extract, store and use talker variation in speech processing in a more fine-grained way, earlier in development.

This explanation is in part consistent with Labov's (1964) proposal that there are multiple stages in the development of the processing of indexical variation. In this model, children are initially thought to focus on learning the basic grammar of their language from their parents/caregivers, with the ability to understand the social significance of dialect variation only emerging in early adolescence (cf. Jones et al., 2017). Whilst our results support the idea that children's representations are likely initially influenced primarily by their core set of experiences with their home dialect(s), they also suggest that social context plays an important role in the development of sociolinguistic awareness. The fact that our bilinguals, exposed to greater amounts of variation than our monolinguals, were highly successful at this task indicates that although all children likely have the capacity to perceive differences between talkers early in development, the ability to do this might present differently and at a different stage in development, depending on the child's linguistic environment. Crucially, our evidence suggests that for bilinguals growing up in a diverse environment, an understanding of the social significance of dialect variation may begin to emerge earlier than previously thought. Of course, not all bilingual children grow up in this kind of environment. Others grow up in communities where they are exposed to a single, dominant variety of the host country language in their community, and though they speak their other language at home with a caregiver and other close family members, may only have limited experience of it outside of this setting (cf. Grosjean, 1998). This differs from the experience of the bilingual children growing up in dense immigrant communities like the one tested here, where children have increased exposed not only to different native-speaker varieties of English (e.g., through their monolingual peers, teachers) but also to foreign-accented varieties (McCarthy et al., 2013). In this case, we would predict that those bilinguals, exposed to less variation, would perform more like our monolinguals, with the ability to categorize talkers according to spoken accent emerging later in development.

What is being encoded? Previous research has suggested there is a developmental progression from specific and concrete knowledge (phonetically rich, potentially talker-specific) to more abstract

knowledge of sound structure (generalizations), with age and exposure playing an important role in the development of an understanding of the social significance of variation (cf. Munson, 2010, Beckman et al., 2007). Our results provide some tentative support for this view; despite being unfamiliar with the accents tested, our bilingual children were able to successfully categorize talkers, perhaps because they had developed representations in which particular speaker attributes were associated with more fully developed macrosociological categories. However, the exact nature of these representations requires further exploration, and whilst our results highlight the need for an important role for social context, we would not want to exclude a role for a phonetic-phonological explanation. For example, one could imagine that If listeners' representations develop to be tolerant of variation, at least to some degree, then as well as facilitating perceptual adaptation by making it easier to map phonetic variation to underlying phonological (i.e., abstract) categories (Shaw et al., 2018), they might also facilitate mapping of variation across individual talkers to macrosociological constructs. In developmental terms, this might mean that as children shift from phonetic to phonological processing (Nathan et al., 1998), they may be better able to spot and extract patterns of variation, and map these to social categories based on their developing understanding of their social world.

CONCLUSION

This study has shown that early experience with speech affects children's ability to use variation in an explicit categorization task. Despite the fact that children were exposed to a small number of talkers and heard only a small number of sentences, bilingual children were able to categorize talkers in all 3 accent pairings. In contrast, monolingual children were only able to categorize talkers in the Home-Foreign conditions, where the differences between the accents were maximized. Although these findings are consistent in part with an exemplar-based account, they could also be interpreted in the context of a gradient representation hypothesis, with the home accent forming the core set of experiences and other accents identified relative to this. They further suggest an important role for social context, such that the development of the ability to extract and use patterns of variation to categorize talkers may differ according to a child's linguistic environment. Knowledge of phonetic and social variation is thus likely emergent, with any abstract categories formed in part at least, from social and phonetic knowledge gained and filtered through experience. Thus, just as phonetic perception becomes specialized for one's native language(s) early in life (Kuhl et al., 2006; Werker & Tees, 1984), so too might sociolinguistic awareness, with perception filtered through the lens of an individual's understanding of their social world.

FOOTNOTES

1. Parents of all children in the target age group received a letter and information sheet informing them about the study and asking them if they would like their child to take part. Participation was on an opt-in basis: only those who returned the completed questionnaire and consent form to school and met the criteria took part.

REFERENCES

Adank P, Evans BG, Stuart-Smith J, Scott, SK (2009): Comprehension of Familiar and Unfamiliar Native Accents under Adverse Listening Conditions. Journal of Experimental Psychology: Human Perception and Performance 35 (2): 520-529.

Ahlberg A, Wright J (1980): Mrs Plug the Plumber. Puffin.

Bao Z (1998): The sounds of Singapore English. In English in New Cultural Contexts: Reflections from Singapore, (eds.) JA Foley, T Kandiah, Z Bao, AF Gupta, L Alsagoff, CL Ho, L Wee, IS Talib, and W Bokhorst-Heng, pp.127-151. Singapore: Oxford University Press.

Bayard D, Gallois C, Weatherall A, Pittam J (2001): Pax Americana? Accent attitudinal evaluations in New Zealand, Australia and America. Journal of Sociolinguistics 5(1): 22-49.

Beck E (2016): What It Means to Be an Outsider: How Exposure to Regional Variation Shapes Children's Awareness of Regional Accents in Their Native Language. In A. Babel (Ed.), Awareness and Control in Sociolinguistic Research (pp. 104-122). Cambridge: Cambridge University Press.

Beckman ME, Munson B, Edwards J (2007): Vocabulary growth and developmental expansion of types of phonological knowledge. In J Cole & JI Hualde (eds.), Laboratory Phonology 9, pp. 241-264. New York: Mouton de Gruyter.

Bent T (2014): Children's perception of foreign-accented words. Journal of Child Language: 1 - 22.

Bent T, Atagi E (2015): Children's perception of nonnative-accented sentences in noise and quiet. Journal of the Acoustical Society of America. 138(6), 3985-3993.

Best C, Tyler M, Gooding T, Orlando C, Quann C. (2009): Development of phonological constancy: toddlers' perception of native- and Jamaican-accented words. Psychological Science 20(5): 539 - 542.

Bialystok E (1986): Factors in the growth of linguistic awareness. Child Development 57: 498-510.

Bialystok E, Majumder S, Martin MM (2003): Developing phonological awareness: Is there a bilingual advantage? Applied Psycholinguistics 24: 27–44.

Boersma P, Weenink D (2017): Praat: Doing Phonetics by Computer.

Campbell R, Sais E (1995): Accelerated metalinguistic (phonological) awareness in bilingual children. British Journal of Developmental Psychology 13: 61–68.

Campbell-Kibler KC (2016): Towards a Cognitively Realistic Model of Meaningful Sociolinguistic Variation. In A Babel (Ed.), Awareness and Control in Sociolinguistic Research, pp. 123-151. Cambridge: Cambridge University Press.

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.

Clarke CM, Garrett MF (2004): Rapid adaptation to foreign-accented English. Journal of the Acoustical Society of America 116: 3647–3658.

Clopper CG, Bradlow A (2008): Perception of dialect variation in noise: intelligibility and classification. Language and Speech 51: 175–198.

Cristia A, Seidl A, Vaughn C, Schmale R, Bradlow A, Floccia C (2012): Linguistic processing of accented speech across the lifespan. Frontiers in Psychology 479: 1-6.

Darcy I, Krüger F (2012): Vowel production and perception in Turkish children acquiring L2 German. Journal of Phonetics 40: 568–581.

Deterding D (2005): Emergent patterns in the vowels of Singapore English. English World-Wide, 26(2): 179-197.

Docherty GJ Foulkes P (2013): An evaluation of usage-based approaches to the modelling of sociophonetic variability. Lingua

Donaldson W (1992): Measuring Recognition Memory. Journal of Experimental Psychology: General 121(3): 275-277.

Donaldson J, Scheffler A (2001): The Gruffalo's Child. Macmillan Children's Books.

Dunn DM, Dunn LM (2009): The British Picture Vocabulary Scale. National Foundation for Educational Research in England and Wales.; GL Assessment (Firm) London : GL Assessment 3rd ed.

Evans BG, Iverson P (2007): Plasticity in vowel perception and production: A study of accent change in young adults. J. Acoust. Soc. Am. 121 (6): 3814-3826.

Floccia C, Butler J, Girard F, Goslin J. (2009): Categorization of regional and foreign accent in 5- to 7-year-old British children. International Journal of Behavioral Development 33 (4): 366 -375.

Floccia C, Goslin J, Girard F, Konopczynski G (2006): Does a Regional Accent Perturb Speech Processing? Journal of Experimental Psychology: Human Perception and Performance 32 (5): 1276–1293.

Foulkes P, Docherty GJ (2006): The social life of phonetics and phonology. Journal of Phonetics 34: 409-438.

Girard F, Floccia C, Goslin J (2008): Perception and awareness of accents in young children. British Journal of Developmental Pyschology. 26, 409-433.

Goldinger S (1997): Words and voices: Perception and production in an episodic lexicon. In K. Johnson, & J. Mullennix (Eds.), Talker variability in speech processing (pp. 33–66). San Diego: Academic Press.

Grier JB (1971): Nonparametric indexes for sensitivity and bias: Computing formulas. Psychological Bulletin 75: 424-429.

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

Hawkins S (2003): Roles and representations of systematic fine phonetic detail in speech understanding. Journal of Phonetics, 31: 373–405.

Hawkins S, Smith R (2001): Polysp: A polysystemic, phonetically-rich approach to speech understanding. Rivista di Linguistica, 13: 99–188.

Hay J, Drager K (2010): Stuffed toys and speech perception. Linguistics 48(4): 865–892

Iverson P, Evans BG (2009): Learning English vowels with different first-language vowel systems II: Auditory training for native Spanish and German speakers. Journal of the Acoustical Society of America 126(2): 866-877.

Jeffries E (2016): Children's developing awareness of regional accents: A socioperceptual investigation of pre-school and primary school children in York. Unpublished PhD thesis, University of York, UK.

Johnson K (1997): Speech perception without speaker normalization: An exemplar model. In K. Johnson, & J. Mullennix (Eds.), Talker variability in speech processing (pp. 145–165). San Diego: Academic Press.

Jones Z, Yang Q, Wagner L, Clopper CG (2017): The development of dialect classification across the lifespan. Journal of Phonetics 60: 20-37.

Kinzler KD, DeJesus J (2012): Northern=smart and Southern=nice : the development of accent attitudes. U.S. Quarterly Journal of Experimental Psychology 66(6): 1146–58.

Kovács, ÁM, Mehler J (2009): Cognitive gains in 7-month-old bilingual infants. PNAS. 106. (16) 6556–6560.

Kuhl PK, Stevens E, Hayashi A, Deguchi T, Kiritani S, Iverson P (2006): Infants show a facilitation effect for native language phonetic perception between 6 and 12 months. Developmental Science 9(2): F13-21.

Kuperman V, Stadthagen-Gonzalez H, Brysbaert M (2012): Age of Acquisition ratings for 30,000 English words. Behavior Research Methods 44(4): 978-90.

Labov W (1964): Stages in the acquisition of standard English. In R. Shuy (Ed.), Social dialects and language learning, pp. 77–104. Champaign, IL: National Council of Teachers of English.

Lenth RV (2016): Least-Squares Means: The R Package lsmeans. Journal of Statistical Software 69(1): 1-33.

Lim L (2004): Singapore English: A Grammatical Description (Varieties of English Around the World G33). Amsterdam/ Philadelphia: John Benjamins.

Maye J, Aslin RN, Tanenhaus MK (2010): The Weckud Wetch of the Wast: Lexical Adaptation to a Novel Accent. Cognitive Science 32(3): 543-562.

Mayr R, Howells G, Lewis R (2015): Asymmetries in phonological development: The case of word-final cluster acquisition in Welsh-English bilingual children. Journal of Child Language 42: 146-179.

McCarthy KM, Evans BG, Mahon M (2013): Acquiring a second language in an immigrant community: The production of Sylheti and English stops and vowels by London-Bengali speakers. Journal of Phonetics 41: 344-358.

McCarthy KM, Mahon M, Rosen S, Evans BG (2014): Speech perception and production by sequential bilingual children: A longitudinal study of Voice Onset Time Acquisition. Child Development 85(5): 1965-1980.

Munson B. (2010): Levels of phonological abstraction and knowledge of socially motivated speech-sound variation: a review, a proposal, and a commentary on the Papers by Clopper, Pierrehumbert, and Tamati; Drager; Foulkes; Mack; and Smith, Hall, and Munson. Laboratory Phonology 1, 157-177.

Nathan E, Donlan C, Wells, WHG (1998): Children's comprehension of unfamiliar regional accents: A preliminary investigation. Journal of Child Language, 25, 343–365.

Niedzielski NA (1999): The effect of social information on the perception of sociolinguistic variables. Journal of Language and Social Psychology 18(1): 62–85.

Norris D, McQueen JM, Cutler A (2003): Perceptual learning in speech. Cognitive Psychology 47(2): 204-38.

Office for National Statistics; National Records of Scotland; Northern Ireland Statistics and Research Agency (2016): 2011 Census aggregate data. UK Data Service (Edition: June 2016).

Pinet M, Gan Y, Evans BG, Iverson P (2015): Intelligibility of British English accents in noise for second-language learners. In the Proceedings of the International Congress of Phonetic Sciences, Glasgow, UK.

Pierrehumbert J (2003): Phonetic diversity, statistical learning, and acquisition of phonology. Language and Speech 46: 115-154.

Pierrehumbert JB, Gross P (2003): Community phonology. Paper presented at the Chicago Linguistic Society, Chicago, April 10-12: /http://babel.ling.northwestern.edu/_jbp/cls03.pdf.

Preston DR (1989): Perceptual dialectology. Dordrecht: Foris.

R Core Team (2018): R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org/.

Rasinger SM (2007): Bengali-English in east London: A study in urban multilingualism. In G Davis & KA Bernhardt (eds.), Contemporary studies in descriptive linguistics 11. Oxford, UK: Peter Lang.

Scott DR, Cutler A (1984): Segmental phonology and the perception of syntactic structure. J. Verb. Learn.Verb.Behav. 23: 450–466.

Shaw J, Best CT, Docherty GJ, Evans BG, Foulkes P, Hay J, Mulak KE (2018): Resilience of English vowel perception across regional accent variation. Laboratory Phonology.

Strand E (1999): Uncovering the roles of gender stereotypes in speech perception. Journal of Language and Social Psychology 18(1): 86–99.

Sundara M, Scuttellaro A (2011): Rhythmic difference between languages affects the development of speech perception in bilingual infants. Journal of Phonetics 39: 505–513.

Trudgill P, Hannah J (2002): International English: a guide to varieties of standard English. 4th ed. London: Arnold. p136-138.

Wagner L, Clopper CG, Pate JK (2013): Children's perception of dialect variation. Journal of child language 41(5): 1062-84

Weatherhead D, White KS, Friedman O (2015): Where are you from? Preschoolers infer background from accent. Journal of Experimental Child Psychology 143:171-178.

Wells JC (1982): Accents of English. Cambridge: Cambridge University Press.

Werker JF, Tees RC (1984): Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. Infant Behavior & Development 7: 49-63.

Vertovec S (2007): Super-diversity and its implications. Ethnic and Racial Studies. 30, (6),1024-54.

Xu Rattansone N, Demuth K (2013): The acquisition of coda consonants by Mandarin early child L2 learners of English. Bilingualism: Language & Cognition 17: 646–659.

Yelland GW, Pollard J, Mercuri A (1993): The metalinguistic benefits of limited contact with a second language. Applied Psycholinguistics 14: 423–444.

FIGURES

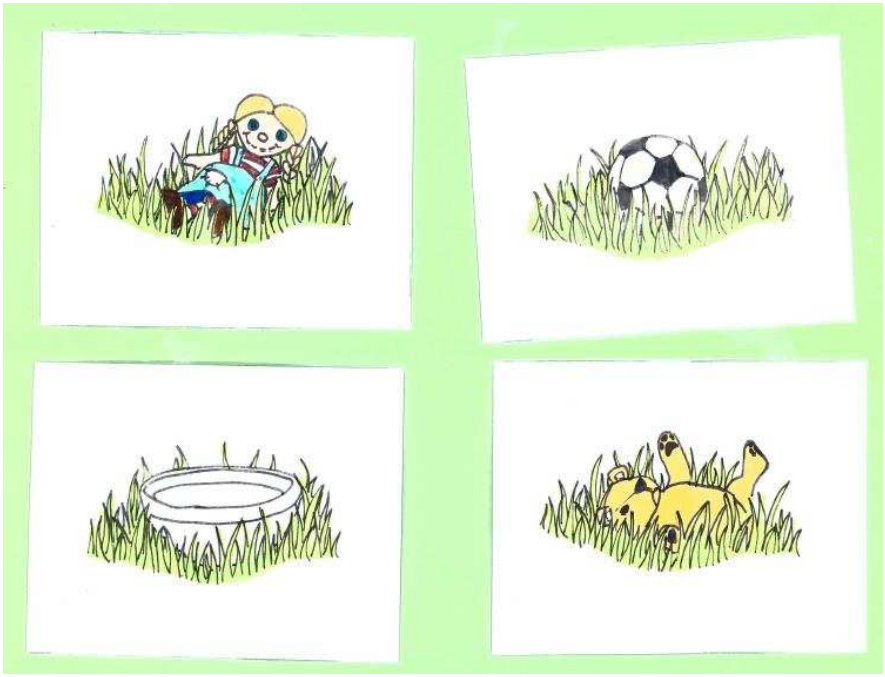

Figure 1: Pictures used in Trial 4 of the comprehension task. The target picture shows a ball in the grass, with the other 3 pictures distractors.

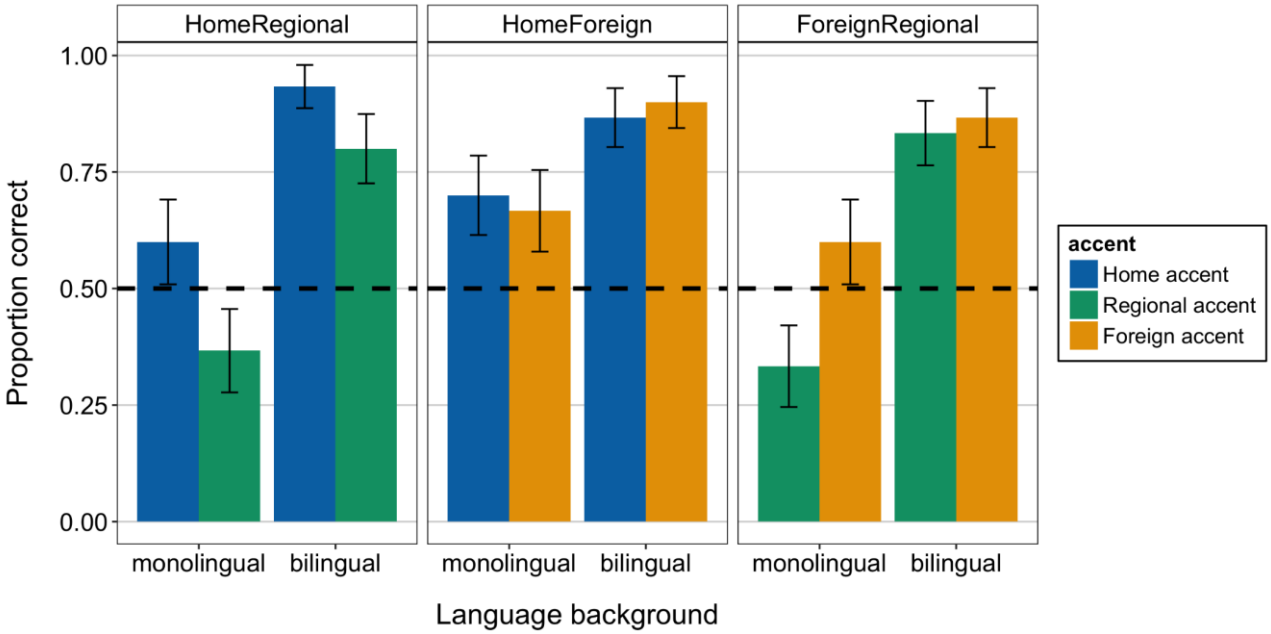

Figure 2: Mean categorization score (proportion correct) for monolingual and bilingual children in each condition (Home-Regional, Home-Foreign, Regional-Foreign) split by accent. Note that each bar represents the average score across 3 trials.

TABLES

Table 1: Mean proportion correct (standard deviation in brackets) for the regional and foreign accents in the comprehension task, split by language background (monolingual, bilingual).

	Regional	**Foreign**
Monolingual	0.91 (0.09)	0.91 (0.14)
Bilinguals	0.88 (0.13)	0.86 (0.15)

Table 2: Summary of the results of the regression model for accent categorisation. F-R = Foreign-Regional; H-R = Home-Regional; b = bilingual. Baselines for predictor variables: Home-Foreign for condition; monolingual for language background. Numbers represent Estimates (β), Standard Errors (SE), Wald statistics (z-values) and p-values

	β	SE	z	p
Intercept	0.733	0.297	2.463	0.014
Condition (F-R)	-0.840	0.410	-2.047	0.041
Condition (H-R)	-0.782	0.403	-1.942	0.052
Language (b)	1.262	0.490	2.575	0.010
Condition (F-R): Language (b)	0.621	0.663	0.937	0.349
Condition (H-R): Language (b)	0.696	0.673	1.033	0.301

Table 3: Mean measures of sensitivity (A') and bias (B''_D) for monolingual and bilingual children calculated for each condition (HomeRegional, HomeForeign, RegionalForeign)

	HomeRegional		**HomeForeign**		**RegionalForeign**	
	A'	B''_D	*A'*	B''_D	*A'*	B''_D
Monolingual	0.46	-0.32	0.76	-0.08	0.42	0.36
Bilingual	0.92	-0.42	0.93	-0.08	0.91	0.2

Appendix 1

Stimuli used in the comprehension and categorization tasks, with response options given where relevant.

i. Comprehension task

TRIAL	STIMULUS SENTENCE	RESPONSE OPTIONS (PICTURES)
1	One of his favourite foods is a buttered mouse roll	mouse in a roll, fox in a roll, owl in a roll, snake in a roll
2	The Gruffalo needs a bath, show me the soap	bath, soap, rubber duck and sponge
3	One of his favourite foods is a jammy doughnut	bubble gum, doughnut, fairy cake, chocolate bar
4	Can you see the ball in the grass?	ball in grass, patch of grass, teddy in grass and flower growing in the grass
5	His favourite drink is a glass of mouse milkshake.	glass of milkshake, glass of orange juice, glass of fizzy pop and carton of juice
6	Show me the giraffe with the purple scarf	giraffe with red scarf, green scarf, purple scarf, blue scarf

ii. Categorization task

TRAINING TRIAL	SENTENCE
1	Oh no! Mummy dropped her money in the puddle! And her gloves are soaking wet!
2	The children ran through the grass and jumped in the puddle and they got their coats very muddy.
3	One night there was a storm. The ship hit a rock and a hole appeared in its side.
4	The next night there was a big celebration party. Mrs and Mr Plug danced on deck under the stars.
REMINDER TRIALS	
1	The lady gave her a cup of tea. When she had finished the lady gave her some money, Mrs Plug set off home.
2	Mrs Plug got on with the job mending the hole in the boat. Mr Plug helped her. They both got soaking wet.

TEST TRIAL	STIMULI
1	Mrs Plug went home. She told Mr Plug about the robber and the rich man. she showed him the reward.
2	Mrs Plug had a useful bag. It had a saw, a hammer, a spanner, a purse, a comb, a lipstick, a plunger, a blow torch and a few other things in it.
3	One night there was some trouble in a lady's bathroom. A plumber was need. The neighbours said send for Mrs Plug!
4	The puppy ate all of the boy's lunch without leaving a crumb. Now his tummy was very, very full.
5	The monkey stuck out his tongue and licked the glass and all of the children laughed.
6	The boy put on his woolliest gloves and rolled a big snowball. He threw the snowball and it hit the window.