Navigating Accent Variation: A Developmental Perspective

Elizabeth K. Johnson¹, Marieke van Heugten², & Helen Buckler³

¹ University of Toronto Mississauga, Canada (elizabeth.johnson@utoronto.ca, ORCID: 0000-0002-9941-9949)
² University at Buffalo, The State University of New York, USA (mariekev@buffalo.edu, ORCID: 0000-0002-4758-2314)
³ University of Nottingham, UK (helen.buckler@nottingham.ac.uk, ORCID: 0000-0002-9039-0406)

Correspondence should be addressed to Elizabeth K. Johnson, Department of Psychology, University of Toronto Mississauga, 3359 Mississauga Road, Mississauga, Ontario, Canada, L5L 1C6. E-mail: elizabeth.johnson@utoronto.ca. Phone: 905-569-4785. Fax: 905-569-4850.
Abstract

Adult processing of other-accented speech is fast, dependent on lexical access, and readily generalizes to new words. But infants’ processing of other-accented speech looks quite different. Laboratory studies suggest that infants below 12 months are incapable of cross-accent word recognition, and recognition of familiar words spoken in novel accents does not emerge until well after the first birthday. However, by the time infants reach their second birthday, their accent processing is qualitatively (though not quantitatively) similar to that of adults. By kindergarten, children are sensitive to the social value of accents, and like adults, use this information to make inferences about speakers. In this review, we discuss how children make this transition from infancy to later childhood. We conclude that making sense of accent variation is fundamental to human communication, and no model of acquisition can be complete without considering how children develop the ability to handle accent variation.

Keywords: accent accommodation, language acquisition, speech development, linguistic diversity, perceptual adaptation, linguistic diversity
1.0 Introduction

Accents are ubiquitous. Despite the lay notion that some speakers have accents and others do not, we all speak in an accent – and our accent can reveal our geographic or cultural origins, or tag us as an in-group or out-group member in any particular setting (Fuertes et al., 2012). Indeed, accent identification and accent adaptation are such an integral part of human interactions that it is hard to imagine a world without accents, or accent adaptation. If we were to imagine a hypothetical person who does not make any social inferences based on an interlocuter’s accent, or who is unable to adapt to the pronunciation of words produced in different accents – we imagine someone who struggles to function in day-to-day communications. Luckily, in the adult world, this is not something we typically find. Adults vary in how successfully they navigate accent variation in their day-to-day life (Banks et al., 2015; McLaughlin et al., 2018), perhaps in part due to how much experience they have doing so (Laturnus, 2018), yet to ensure efficient communication abilities, they necessarily possess at least some flexibly in adapting to different pronunciations and accents – an ability that we would argue is a core design feature of human language competency.

But there is a plot twist to this story, and this plot twist (and the conundrum it presents to language researchers) is the focus of our review. Although our hypothetical speaker is not typically found in the adult world, the limited data we have on the development of accent processing abilities suggest that young children may in some sense (at least temporarily) be that struggling language user, initially lacking sufficient tools to make sense of speech produced in an unfamiliar accent, and unaware of many of the social implications encoded in the way different interlocuters speak. This presents a bit of a paradox for developmental researchers. How can children – who struggle with other-accented speech – otherwise show such great facility in acquiring language (e.g.,
Johnson, 2016) and using the linguistic knowledge they have to navigate their social environment (e.g., Kinzler, 2021)?

There is currently no consensus regarding the answer to this question. Indeed, reasonable answers range from the possibility that we have overestimated children’s struggles with other-accented speech to the possibility that the skills and knowledge used when adapting to accents are not so much a prerequisite for language acquisition, but rather a byproduct of the process. In this article, we review what we do and do not know about the development of children’s ability to adapt to accent-related variation in the realization of spoken words, emphasizing crucial gaps in our knowledge. We outline different theoretical approaches to explaining developmental changes in children’s perception of other-accented speech, and weigh the evidence for each viewpoint. Where relevant, we will also discuss the practical implications of how children cope with different accents, and support our arguments by drawing on literatures outside of linguistics – including social and developmental psychology, early childhood literacy, and the speech sciences. We will end by discussing key questions that remain unanswered (or even as of yet unexplored) in the current literature, and challenge researchers to tackle these questions in the years to come.

2.0 A Chronological View of Development

In the past 50 years, we have learned a great deal about the remarkable efficiency with which the typical infant masters their native language or languages (see Johnson and White, 2020; Swingley, 2009; Werker and Curtin, 2005 for an overview). Crucially for the purposes of this review, however, most of this work has involved presenting infants with stimuli spoken in the regionally dominant variety of the child’s native language. Thus, much less is known about children’s ability to process languages spoken by other-accented talkers. Here, we review children’s developing
ability to handle accent variation from birth to adolescence, and discuss these findings in relation to well-established milestones in early language development (see Figure 1).

2.1 Early beginnings: 0 to 12 months

Speech development proceeds at a remarkable pace in the first 12 months of life. One of the first linguistic competencies demonstrated by infants is their preference to listen to their mother’s language over another language (e.g., Moon et al., 1993). This ability is presumably driven by exposure in the womb, and has been touted as evidence for the precocious nature of infants’ language abilities. But follow-up studies have revealed that newborns only distinguish between their mother’s language and an unfamiliar language if the two are rhythmically dissimilar (e.g., Mehler et al., 1988). Indeed, infants’ ability to tell apart their own native language(s) from rhythmically-similar other languages does not emerge until around 4 to 5 months of age (e.g., Bosch and Sebastián-Gallés, 1997; Nazzi et al., 2000). Thus, given that most (but not all) varieties of any language tend to belong to the same rhythmic class, it is not surprising that infants cannot typically tell apart different varieties of their native language until 5 months (e.g., English-learning American 5-month-olds distinguish British and American English; Nazzi et al., 2000; though see Kitamura et al., 2013, for evidence that this sensitivity may sometimes be in place by 3 months). However, even at 5 months, infants only appear to distinguish two varieties of their native language if one is highly familiar to the infant. For instance, 5-month-old infants learning a South-West variant of British English can tell apart their home accent from an unfamiliar Welsh accent, but they fail to discriminate Scottish and Welsh English, two unfamiliar variants of their native language (Butler et al., 2011; see Wagner, Cooper, & Pate, 2014, for similar findings with older children). Taken together, these studies suggest that in the first few months of life, infants tend to
treat different varieties of the native language as equivalent, and that sensitivity to differences does not develop until after they have accumulated a few months of relevant language experience – although it seems reasonable to predict that this sensitivity might emerge earlier in cases where the home accent is compared to a rhythmically distinct variety of the language (e.g., stress-timed North-American English vs. syllable-timed Singapore English).

Over time, infants’ experience with their native language stimulates their language growth. In the latter half of the first year, we begin seeing much more sophisticated attunement to the native language. Infants no longer show a universal sensitivity to all sound contrasts used in the world’s languages (Best et al., 1988; Kuhl et al., 2006; Werker and Tees, 1984), and they move beyond simply identifying or preferring familiar languages and language varieties. For instance, by 6 to 7.5 months, children begin segmenting words from speech (Bortfeld et al., 2005; Johnson et al., 2014; Jusczyk et al., 1999), and begin to demonstrate some comprehension of highly frequent words (Bergelson and Swingley, 2012; Tincoff and Jusczyk, 1999). By 7.5 to 9 months, infants begin using knowledge of the native language sound structure, including prosody and phonotactics, to find words in fluent speech (e.g., Jusczyk et al., 1999; Mattys and Jusczyk, 2001; Polka and Sundara, 2012). The highly frequent words infants first discover fuel further lexical growth by facilitating the segmentation of additional wordforms from speech (e.g., Bortfeld et al., 2005; Mersad and Nazzi, 2012; Shi et al., 2006). By 12 months, some have estimated that infants’ receptive lexicons contain as many as 70 words (Frank et al., 2017).

Throughout this early infancy period, as children are rapidly tuning in to the sound structure of their native language, they seem to handle talker variability fairly well (Bergelson and Swingley, 2018; Johnson et al., 2014; Kuhl, 1979; van Heugten and Johnson, 2012; though see Houston & Jusczyk, 2000). Indeed, current evidence suggests that the number of speakers in an infant’s
environment has no impact on their early speech development (Bergmann and Cristia, 2018). But in line with claims that talker and accent adaptation are distinct processes (Kriengwatana et al., 2016), accent variation seems to initially present infants with a much bigger challenge. For example, although young infants can sometimes successfully recognize words familiarized in an unfamiliar accent when they are once again presented in that same unfamiliar accent (Nazzi et al., 2014), 9- to 10-month-old infants familiarized with a word form in one accent fail to recognize it in a second accent (Schmale et al., 2010; Schmale and Seidl, 2009). This has been interpreted as evidence that children’s early lexical representations are over-specified and not equipped to handle accent variation, leading many of the most prominent models of infant speech perception to argue for an episodic account of early word recognition (Jusczyk, 1993; Werker and Curtin, 2005).

As infants’ speech perception abilities rapidly develop, and they become more sensitive to the fine acoustic-phonetic details identifying their own language variety, so too does their understanding that language can act as a social marker (e.g., Kinzler, 2021). By 6 months of age, infants (in at least some communities) associate foreign languages with other-race individuals (Uttley et al., 2013). Around the same time, they tend to direct their attention to linguistic in-group members, for example preferring to listen to a tune introduced by a speaker of their native language than an unfamiliar language (Soley and Sebastián-Gallés, 2015). By 11 to 12 months, infants preferentially try foods introduced by speakers of their native language as opposed to a foreign language (Shutts et al., 2009) and also preferentially imitate speakers of their own language (Buttelmann et al., 2013; de Klerk et al., 2019). Thus, there is ample evidence that the language spoken by an interlocuter informs infants’ attentional preferences - but do children also make social inferences based on the accent spoken by an interlocuter? To date, not much work has been conducted on how infants use accent information to make social inferences. Notably, however,
one study demonstrates that six-month-olds look longer to people whom they have previously heard speaking their native language with a familiar accent than to people they have heard speaking in an unfamiliar accent (Kinzler et al., 2007). Although such attentional biases do not necessarily imply social preference (Haith, 1998), they may indicate that infants have identified accent as a cue to group membership, an initial stage in the development of accent-based social categorization (Bigler and Liben, 2007; Imuta and Spence, 2020). Thus, in the latter 6 months of the first year of life, although infants struggle to cope with accent variability in speech perception, they may have already begun using accent information to guide their interactions. That said, they are still a long way off from adults, who are heavily reliant on subtle accent information to draw social inferences about others in their environment (Kozlowski, 2015).

2.2 Dramatic changes in accent accommodation: 12 to 24 months

In the second year of life, as children transition from infancy to toddlerhood, their language abilities dramatically improve. They develop sensitivity to fine-grained mispronunciations in familiar words (e.g., Mahr et al., 2015; Paquette-Smith et al., 2016; White & Morgan, 2008), broaden their repertoire of word learning strategies to facilitate their fast mapping (e.g., Graf Estes et al., 2007; Paquette-Smith and Johnson, 2016; Spiegel and Halberda, 2011), begin using grammatical knowledge to constrain the processing of spoken language (e.g., Cauvet et al., 2014; van Heugten and Christophe, 2015; van Heugten and Shi, 2009), show sensitivity to semantic relatedness between items in their lexicon (e.g., Arias-Trejo and Plunkett, 2009; Johnson et al., 2011; Mani et al., 2013; Willits et al., 2013), and experience an explosion in their receptive vocabulary (e.g., Bloom, 1973). Children’s production abilities also take off during this period, with children in the United States typically producing 90 words by the time they reach 18 months.
of age, and 300 words by the time they reach 24 months of age (Frank et al., 2017). During this time, we also see substantial growth in children’s ability to use social information to make inference about intended meaning and to make judgements on the reliability of different types of information (Howard et al., 2015; Stenberg, 2009).

Perhaps unsurprisingly, these dramatic developmental changes in language processing are accompanied by a striking improvement in children’s ability to handle accent variation. In the first six months of the second year of life, these improvements are still slow, but by the latter half of this year, progression is more rapid. At 12 to 13 months, cross-accent recognition of newly segmented word forms is first observed (i.e., word forms learned in a familiar variety of the native language can be recognized when spoken in a different variety; Schmale et al., 2010; Schmale & Seidl, 2009). This is a great advance from the word form recognition abilities we see in younger infants; however, spontaneous recognition of familiar word forms produced in unfamiliar accents emerges much later, and is not reliably observed until about 19 months of age. That is, although children at 12.5 and 15 months of age listen longer to known word forms (e.g., ball, diaper) than to low-frequency or nonsense word forms (deuce, koddy) in their own variety of English, we do not see comparable performance when the same word forms are produced in an unfamiliar accent until about 19 months of age (Best et al., 2009; van Heugten et al., 2018; van Heugten and Johnson, 2014). Eye tracking studies, where infants are presented with two images side-by-side on a screen and where an auditory prompt directs them to look at one of the two images, reveal a similar pattern of results (Mulak et al., 2013; van der Feest and Johnson, 2016; van Heugten et al., 2015). Without any previous exposure to an accent, infants typically fail to recognize familiar words spoken in unfamiliar accents, even with a visual prompt for named targets clearly presented. But as their vocabulary size and linguistic skills increase around their second birthday, toddlers
suddenly begin to spontaneously accommodate to unfamiliar accents so well that they rapidly comprehend familiar words in unfamiliar accents even in the absence of prior exposure (Mulak and Best, 2013; van Heugten et al., 2015; van Heugten and Johnson, 2014). Some have argued that this dramatic shift in children’s accent processing abilities is linked to a qualitative shift in children’s phonological abilities (see section 3.1 below).

But is it really true that infants under the age of 19 months have no ability to contend with other-accented speech? The implications of this would be profound. How could children acquire language so quickly and efficiently if their word recognition abilities ground to a halt every time they encountered a novel accent? If this were the case, shouldn’t we expect to see dramatic differences in the language processing abilities of children growing up in linguistically diverse environments (where other-accented speech would be frequently encountered) compared to linguistically homogeneous environments (where other-accented speech would be less frequently encountered)? Upon closer inspection, it becomes clear that all of the studies discussed above examine children’s ability to contend with accent variation in a relatively unnatural lab task, where unfamiliar, disembodied voices produce words in isolation or short sentences with little communicative context. This does not fully represent everyday listening situations for children, who may know (and may have previously interacted with) the other-accented interlocutor. Maybe children – like adults – would perform better with other-accented speech if given a chance to adapt to it (e.g., Bradlow & Bent, 2008; Clarke & Garrett, 2004; Maye et al., 2008; Trude & Brown-Schmidt, 2012).

To address these issues, an accent adaptation paradigm has been used where infants are given a brief (typically well-controlled and highly engaging) exposure to an unfamiliar accent in the lab. By comparing word recognition performance in a specific accent with and without
exposure to that same accent, researchers can examine what children can learn from engaging with or listening to other-accented talkers. This approach to understanding children’s ability to handle accent variation is in many ways more ecologically valid than the approaches outlined above, and has revealed a great deal about how infants process other-accented speech (see discussion in section 3.0 below). Results show that by 15 to 24 months, just a few minutes of exposure can dramatically improve infants’ ability to recognize words in that accent (van der Feest and Johnson, 2016; van Heugten and Johnson, 2014; White and Aslin, 2011), and this adaptation is found even for words that had never been heard in that accent before. It appears, though, that the efficiency and generalizability of this process is dependent on various factors, such as whether the exposure phase was prerecorded or presented live (Paquette-Smith et al., 2021), how many accents were present during exposure (Potter and Saffran, 2017), how different the unfamiliar accent is from the infant’s own accent (Cooper et al., in revision; Newman et al., 2018), whether social information is available (Singh et al., 2020; Weatherhead and White, 2018), and what the task demands are (van Heugten et al., 2015; van Heugten and Johnson, 2016). This speaks to a complex interaction between all sources of information available to the child during early accent accommodation.

2.3 Beyond 24 months: Continued development

Although children are fairly adept at recognizing familiar words spoken in unfamiliar accents by their second birthday, quantitative differences in accent processing between children and adults continue to exist for an extensive period of time (see Sidebar 2). Research with school-aged children, for instance, has revealed that their comprehension of speech in unfamiliar accents lags behind that of adults (Bent, 2014; Bent and Atagi, 2017, 2015; Bent and Holt, 2018; Creel et al., 2016; Nathan et al., 1998; Newton and Ridgway, 2016). In one study, 5- to 6-year-olds and adults
were presented with a sentence repetition task, which is arguably more cognitively demanding than the tasks typically used with infants and toddlers. Crucially, the sentences participants were asked to repeat were presented in both the participant’s native North American accent as well as an unfamiliar Japanese accent. Children performed worse than adults, with this difference driven primarily by the unfamiliar accent condition. That is, although both children and adults repeated fewer correct words when listening to the other-accented than to the own-accented speaker, the other-accented speaker affected children’s performance more than adults’ (Bent and Atagi, 2015). In fact, it has recently been shown that the developmental trajectory of learning to cope with unfamiliar accents continues through adolescence (Bent, 2018). Thus, although the fundamental capacity to contend with accent variation is present early in life, accents may nonetheless complicate language processing for a protracted period of time.

Throughout the school-age years, as children’s processing of other-accented speech gradually improves, they also exhibit a growing recognition of their own language variety. Metalinguistic tasks demonstrate that children as young as 5 years-old are able to distinguish their own accent from a perceptually distinct foreign accent (Floccia et al., 2009; Girard et al., 2008; Paquette-Smith et al., 2019; Wagner et al., 2014), but the ability to tell apart less distinct accents is more challenging, with gradual improvement observed throughout childhood and into adolescence (Jones et al., 2017; McCullough et al., 2019).

As children become better at distinguishing their own language variety from others, their tendency to make social inferences based on accent information also matures. For example, 3- to 6-year-olds use accent to infer a speaker’s geographical origins (Kinzler and DeJesus, 2013; Weatherhead et al., 2019, 2018, 2016). That is, they believe that two speakers who share the same accent live in the same place (Weatherhead et al., 2016) and that speakers with more pronounced
accent differences live in more distant locations (Weatherhead et al., 2019). Interestingly, these effects are driven only by a speaker’s accent, not their grammatical skill; Hwang and Markson, 2018; St Pierre et al., under review). Children also link cultural items (Wagner et al., 2014) and cultural norms (Weatherhead et al., 2016) with accents. Between 3 and 8 years of age, children increasingly rely on speaker accent to decide who to trust and learn from (Kinzler et al., 2011; Ronfard et al., under review). During this same time period, children develop a preference for same-accent peers (Creel, 2018; Paquette-Smith et al., 2019; St Pierre and Johnson, 2020), and rate foreign-accented adults as being worse teachers than same-accented adults (Paquette-Smith et al., under review). These tendencies grow stronger between 4 and 11 years of age (Liberman et al., 2018; Spence and Imuta, 2020; St Pierre and Johnson, 2020). Children’s sensitivity to other people’s attitudes towards accents also develops during this period, and in the process, they come to understand that their own variety may be stigmatized. This can have an impact on their literacy development and educational achievement (see Patton et al., 2010; Snell and Andrews, 2017 for review in US and UK contexts).

3.0 Taking a closer look at children’s perception of accented speech

As discussed above, children’s ability to handle other-accented speech improves dramatically during their second year of life. In fact, one could argue that by their second birthday, toddlers’ perceptual adaptation is fundamentally similar in quality to that of adults, for whom adaptation is fast (Clarke and Garrett, 2004), dependent on lexical access (Norris et al., 2003), and generalizes across words (Bradlow and Bent, 2008; Maye et al., 2008; McQueen et al., 2006). But how do children accomplish this? In this section we will take a twofold approach to addressing these questions. First, we will review and compare three accounts for explaining the dramatic shift seen
in children’s accent processing abilities in the second year of life (section 3.1) and then we will examine how infants and toddlers use prior exposure to an accented speaker to adapt to different accents (section 3.2).

3.1 Making the quantum leap from infancy to toddlerhood

At approximately 19 months of age, infants appear to make a quantum leap in their ability to cope with accent-related variation. But is what we observe as researchers truly a quantum leap, or is it just the result of tiny incremental changes in infants’ abilities that eventually result in what appears to be a quantum leap? If the former, what drives the quantum leap? And if the latter, what types of incremental changes could be happening prior to 19 months of age that could build up into what appears to be a quantum leap?

As it turns out, the way researchers answer this question in part depends on how they view adult speech perception, and in part on which theory of phonological development they find most plausible. Some researchers have argued for word representations being episodic in nature across the lifespan (e.g., Goldinger, 1996; Houston and Jusczyk, 2000). In a nutshell, the exemplar-based account posits that the mental lexicon consists of traces of previously experienced surface forms of words. These traces include indexical details such as the speaker’s voice and accent. Only if the incoming word is acoustically similar to one or more stored traces will the word be recognized. According to this view, cross-accent word recognition is initially a challenge for infants because their trace-based memories do not capture the acoustic variation of a different accent. But over time, as listeners experience a greater amount of variation in the pronunciation of words, their ability to cope with other-accented speakers improves (see Rost and McMurray, 2009 for such evidence with infants in the domain of speaker variability). That is, when more exemplars are
stored, less prototypical pronunciations are more likely to activate a trace associated with the intended word. This view thus stipulates that the improvement in the ability to cope with accents between infancy and toddlerhood can be attributed to relatively small incremental changes over time.

Other researchers have instead argued that abstract phoneme representations are necessary to explain accent adaptation in both childhood and adulthood (e.g., Cutler, 2008; McQueen et al., 2012; van Heugten and Johnson, 2014). Abstractionist accounts for the sudden shift we see in toddlers’ behavior around 18 to 19 months come in two closely related, but somewhat different, flavors. We will henceforth refer to these two accounts as the phonological reorganization and the continuous abstract account. According to the phonological reorganization account, the changes we see in toddlers’ ability to spontaneously recognize other-accented words at around 18 to 19 months are attributable to a sudden shift in children’s representational capabilities. Proponents of this view tend to emphasize that this shift coincides with the naming explosion (e.g., Mulak and Best, 2013). The change in children’s accent processing abilities could be instigated by children’s greater ability to detect the invariant phonological structure of words (i.e., the development of phonological constancy; Best et al., 2009) or because initially holistic representations of lexical items transform into more abstract segmental representations (Metsala and Walley, 1998).

But not all researchers who emphasize the importance of abstract phoneme representations for accent processing agree that children undergo a dramatic change in their representational capabilities at 18 to 19 months (e.g., van Heugten and Johnson, 2014). The continuous abstract account, in contrast to the phonological reorganization account, assumes a stronger continuity between young infants’ early (pre)lexical representations and toddlers’ more mature lexical representations. Proponents of this view attribute toddlers’ improved handling of accented speech
to things such as increased attunement to the native language phonology and greater experience mapping the speech signal to underlying representations (cf. van Heugten and Johnson, 2014). According to this view, the striking changes observed between infancy and toddlerhood are not due to a stage-like shift in representational capacities. Rather, children’s abilities to handle accent variation improves gradually over time; the appearance of a stage-like is attributed largely to the limited sensitivity of our testing procedures. Thus, continuity (or lack thereof) in the representations used by infants and toddlers to recognize other-accented words are a major factor distinguishing different accounts for early accent processing (see Table 1). In this regard, debates over how toddlers improve in their ability to process accented speech echo classic debates on the stage-like versus continuous development of cognitive abilities (e.g., Courage and Howe, 2002).

To summarize, we have discussed three different accounts for why children show a sudden shift in their ability to handle accent variation in toddlerhood. The exemplar-based account posits that the lexicon consists of traces of specific tokens. Support for this view comes from the protracted period of development in children’s ability to handle accent variation. In contrast, the phonological reorganization and continuous abstract accounts both assume that abstract phonological representations are necessary to efficiently handle accent variation. Importantly, these latter two accounts differ in how these representations emerge – that is, they differ in whether they assume qualitative or quantitative changes in phonological processing abilities at around 18 to 19 months of age. The continuous abstract account makes a clear assumption about abstract phonological representations existing early in infancy, arguing that abstraction is a fundamental component of human speech processing, and necessary to explain infants’ early success at building a proto-lexicon and handling talker variability. Support for this view comes from studies demonstrating accent adaptation in infants under 18 to 19 months of age (van Heugten and
In contrast, different variants of the phonological reorganization account may or may not assume abstract representations before 18 to 19 months, but when abstract representations are assumed, they are rudimentary (at best) and are qualitatively different from those that emerge later (see Table 1). Support for this view comes from the dramatic changes we see in toddler’s behavior at about 18 to 19 months.

3.2 Taking a closer look at how accent exposure induces accent adaptation

As mentioned above, toddlers are able to adapt to a speaker’s accent after as little as two minutes exposure (van Heugten et al., 2018; van Heugten and Johnson, 2014; White and Aslin, 2011). But how is this accomplished? Two distinct types of explanations have been proposed. One class of explanations, often referred to as the specific mapping hypothesis, proposes that any other-accented input can cause toddlers to change their perception by generating precise between-accent mappings, perhaps in a lexically-guided fashion (van Heugten et al., 2018; van Heugten and Johnson, 2014; White and Aslin, 2011). By updating the specific signal-to-phoneme mappings in line with the observed evidence, specific, directional shifts are created. For example, if a child hears a speaker produce the words kiss and bib with a long, raised vowel, [kiːs] and [biːb], instead of [kɪs] and [bɪb], they will deduce that this speaker pronounces the sound [i] as [iː]. They will then adjust their expectations and will not be surprised when words they have never heard the speaker produce follow this same pattern (e.g., fish produced as [fiːʃ]). As the child’s expectations adjust, their processing of the other-accented speech should show rapid and notable improvement. This view generally assumes an abstract pre-lexical level of processing, so that a listener’s lexical representations themselves are unaffected by short-term exposure to a different-accented talker. Instead, the mapping between the pre-lexical and lexical level are adjusted so that the accented
surface forms can be mapped onto the underlying representations. Note that since this strategy makes use of the prelexical phoneme level, it is really only viable with abstractionist accounts of language processing.

An alternative strategy of accent adaptation involves the *general expansion* hypothesis, which suggests that exposure to accented speech could lead children to rely less on the specific details of the word representations. For example, if a child hears the same speaker as before produce [kiːs] and [biːb] instead of [kɪs] and [bɪb], they will loosen their criteria for lexical access in such a way that they will readily perceive not only [iː] as an acceptable pronunciation of [ɪ], but will also accept [eː] or perhaps even [ɛ] as an instance of [ɪ], despite the lack of evidence that the other-accented talker produces [ɪ] in this fashion. This general expansion of acceptable pronunciations could be instigated either by expanding the mapping between the speaker’s pronunciation of the words and the underlying lexical representations and/or simply by lowering the activation threshold for word recognition. This would result in children being able to recognize words in unfamiliar accents, although one unfortunate by-product would be that minimally different word pairs would become more confusing (through increased lexical competition). While this approach can be realized as broadening on the prelexical phoneme, it can also be established at the level of lexical representations themselves. For this reason, the general expansion account is congruent with both abstractionist and exemplar-based models of language processing.

There is substantial evidence for both the specific mapping and general expansion hypotheses. On the one hand, evidence for the *specific mapping* hypothesis comes from an accent adaptation study with 19-month-olds. After brief exposure to an artificial accent, toddlers demonstrate accent-specific adaptation to novel productions of familiar words (White and Aslin, 2011). Furthermore, by 24 months, after two minutes of exposure to an unfamiliar variety of the
native language that maintains a phonological contrast not present in the child’s own variety, children begin demonstrating sensitivity to a voicing contrast they initially ignored (van der Feest and Johnson, 2016; see also van Heugten et al., 2018 for additional evidence of children’s detection of mispronunciations in unfamiliar accents). This requires access to a pre-lexical level, which exemplar-based models do not have. On the other hand, there is also good evidence for general expansion. For example, after brief exposure to some type of variability, whether in the form of speech or completely outside the speech realm, children appear to recognize words spoken in an unfamiliar accent never heard before (Schmale et al., 2015). Similarly, other work shows that prior exposure to multiple unfamiliar accents helps children cope with a fourth accent not previously heard (Potter & Saffran, 2017). These findings suggest that the presence of variability or unexpected events (either in the speech signal or elsewhere) can cause infants to loosen their expectations regarding the exact pronunciations of words, and that this can happen without generating a specific mapping between the child’s own accent and the talker’s accent.

To summarize, the developmental literature contains evidence for both the specific mapping and general expansion explanations for accent adaptation. Could both explanations be right? Although the specific mapping and general expansion theories are often presented as two categorically distinct explanations for early accent adaptation (cf. Schmale et al., 2015, 2012; van Heugten et al., 2018; van Heugten and Johnson, 2014; White and Aslin, 2011), in the adult literature some have concluded that listeners appear to employ both strategies (e.g., Cooper and Bradlow, 2018; Goldinger, 1996; Kleinschmidt and Jaeger, 2015; Pierrehumbert, 2016). Such a flexible, dual-systems account could work particularly well for young children, who lack the speech processing efficiency and vocabulary that adults have. We will henceforth refer to this possibility as the hybrid flexibility hypothesis. One can imagine that children could employ a
general expansion approach when initially confronted with an unfamiliar accent (or when a newly confronted accent is particularly distinct from the child’s own accent), and then replace it with the specific mapping approach when possible. Note that this flexible hybrid approach makes specific predictions about what factors should elicit a general expansion versus specific mapping approach to accents, and ongoing work in our laboratories is currently further testing these predictions. For example, in support of the category broadening hypothesis for distant accents, we have recently found that toddlers pre-exposed to a distant novel accent before a word recognition test subsequently recognized words more slowly in their own accent than children who did not experience this pre-exposure. This suggests that the children experiencing the distant accent pre-exposure broadened their criteria for what counts as acceptable pronunciation for familiar words, resulting in slower word recognition even in their own accent variety (e.g., Cooper et al., in revision). We are also currently exploring whether routine exposure to multiple accents in everyday life enables toddlers and preschoolers, like older school-aged children (Levy et al., 2019) and adults (Laturnus, 2018), to better adapt to speech produced in unfamiliar accents.

**4.0 Looking forward: accent processing in linguistically diverse settings**

Laboratory studies clearly demonstrate that infants initially struggle to handle accent variation, and although toddlers’ accent processing abilities may be qualitatively similar to adults’ accent processing abilities, children continue to experience difficulty processing other-accented speech well into late childhood (and perhaps even adolescence). But how does this play out in the real world? Exposure to multiple accents in early childhood is not rare, and is in fact the norm in many parts of the world, such as in large multicultural cities (e.g., Hong Kong, Toronto, and Los Angeles) and even more rural locations where pockets of linguistic diversity overlap or are tightly
packed together. Imagine a child growing up in Canada with one Australian parent and one American parent. Based on input at home, this child would need to learn that the similar-sounding forms of the word *ball* pronounced with an Australian accent and the word *bowl* pronounced with an American accent refer to different entities, whereas the more dissimilar pronunciations of the word *ball* in the two accents refer to the same referent. And what would this same child do when they went to daycare and heard their peers speaking in a Canadian English accent, and their teacher speaking with a French accent? If infants struggle with accent variation as much as laboratory studies suggest, then shouldn’t language development differ in children routinely exposed to one versus many accents? And how do social relationships and other social cues factor into their learning and processing abilities? It is surprising that despite the prevalence of multi-accent exposure in early childhood, and despite an increasingly awareness of the consequences of multilingual development amongst infant language researchers and speech therapists (e.g., Byers-Heinlein, 2018; Mcleod et al., 2017), very little work to date has examined how multi-accent exposure affects early language acquisition (Johnson, 2018). Examining the development of accent processing abilities in these children is not only crucial to testing the ecological validity of the laboratory studies reviewed above, it is also crucial to understanding the factors that shape language learning trajectories in the real world, and possibly even important to developing fair access to education (e.g., Morgan et al., 2015) and speech language assessments (e.g., Goldstein and Iglesias, 2001; Morgan et al., 2016).

The first experimental study to examine language development in toddlers exposed to one versus two varieties of the native language compared word recognition in British 20-month-olds. Mono-accent children were exposed to the locally dominant rhotic variety of British English from both of their parents. Multi-accent children had at least one parent who was a speaker of a non-
rhotic variety (Floccia et al., 2012). All children were tested on their recognition of words containing the rhotic/non-rhotic contrast across the two accents, and the results were quite surprising. Both groups of toddlers recognized only those words produced in the socially dominant rhotic variety. The authors concluded that the multi-accent children were selectively acquiring the dominant variety spoken in the community. In a follow-up study, 20-month-olds were tested on correctly pronounced and mispronounced versions of familiar words spoken in the locally dominant community accent, and the multi-accent children (unlike the mono-accent children, who only recognized correctly pronounced words) recognized all words, regardless of whether they were correctly pronounced or not (Durrant et al., 2015). Thus, once again, the multi-accent children performed quite differently than the mono-accent children. The authors concluded that multi-accent children have poorly specified lexical representations due to the variable nature of their language input. A more optimistic way to frame this finding, though, is that the multi-accent toddlers were simply behaving in a manner adaptive to their environment. In a Bayesian framework of speech processing (e.g., Kleinschmidt & Jaeger, 2015), this could perhaps be conceptualized as having weaker priors due to less exposure to the socially dominant variety of the native language, which in turn would give infants increased flexibility for recognizing words in a linguistically diverse environment. Or one could just think of multi-accent children as not assuming that everyone will necessarily speak in the dominant community accent. But of course, this more optimistic framing does not negate the fact that optimizing the recognition of more variable pronunciations of words in this way would likely come at the cost of slowing down the recognition of words in the locally dominant variety.

More recent studies examining the effect of multi-accent exposure have suggested that the relationship between accent exposure and language development is more complicated than the
initial studies with multi-accent British English learning infants suggested. In one set of mispronunciation detection studies, Dutch-learning 24-month-olds routinely exposed to two varieties of their native language in their day-to-day life appeared to outperform their age-matched peers who were routinely exposed to only one variety of their native language (van der Feest and Johnson, 2016). While the multi-accent group adjusted their speech processing expectations to suit the accent of the talker at hand, the mono-accent group only showed sensitivity to mispronunciations in an unfamiliar variety of Dutch after a two-minute exposure to a talker speaking the unfamiliar variety. A follow-up study working with this same population suggested that the multi-accent Dutch learners were not in general slower to recognize words than the mono-accent Dutch learners (van der Feest, Rose, & Johnson, under review). However, in another study testing monolingual English-learning 24-month-olds on their recognition of words produced in the locally dominant variety of Canadian English, children exposed to a mix of foreign-accented and local-accented English were indeed slower to recognize words than children exposed routinely to only the locally dominant variety of English (Buckler et al., 2017). Thus, in some studies, multi-accent children seem to outperform mono-accent children in their speech processing flexibility, but in others (specifically those comparing performance on the socially dominant variety of the native language) mono-accented children appear to recognize words more efficiently than multi-accent children. Interestingly, whatever differences we might observe in multi- versus mono-accent toddlers’ word recognition performance in the lab, an ongoing large-scale vocabulary study has thus far failed to find any link between vocabulary size and children’s exposure to accent variation (Fung et al., 2019).

So how can we reconcile lab studies on children’s processing of other-accented speech with real world studies examining how language unfolds in children growing up with exposure to
multiple varieties of the native language? And how might we need to adjust our models of early speech and language development to address children’s success (and struggles) when exposed to language variability early in life? Clearly, we still have a lot to learn on this topic. One key piece of the puzzle will come from testing younger infants, since this is the age group that typically displays the most difficulty in processing other-accented speech. To date, only one study has examined how multi-accent exposure affects early word recognition in infants, and this study suggests that multi-accent exposure has a dramatic effect on early word form recognition (van Heugten and Johnson, 2017). Whereas mono-accent children readily recognized word forms produced in the locally dominant variety of English, the multi-accent children did not, potentially because multi-accent infants come to the lab with different assumptions about how speakers sound. For example, given that they are used to hearing people speak in different accents, infants growing up in multi-accent environments may not make assumptions about the accent of the speaker, and may hence need additional time to figure out what signal-to-representation mapping scheme to use. Multi-accent infants may also lag behind mono-accent infants when it comes to attuning to native language sound structure, given that some L2 varieties of an infant’s native language may initially be mistaken for another language (Paquette-Smith and Johnson, 2015). A second piece of the puzzle will be provided by looking at long-term consequences of early multi-accent exposure. Some studies have suggested that any effect of multi-accent exposure on early word recognition dissipates by about 32 months of age (Buckler et al., 2017), but other studies suggest the effects may have a lifelong impact on speech processing (Antonio, Grohmann, Kambanaros, & Katsos, 2016; Clopper and Pisoni, 2004; Chen, Rattansone, Cox, & Demuth, 2017). A final piece of the puzzle will come from considering how exposure to accent variability interacts with children’s use of accent information as a social cue. Early work in this area suggests that monolingual children
more positively evaluate age-matched peers (Paquette-Smith et al., 2019) and adults (Paquette-Smith et al., under review; Ronfard et al., under review) who speak with the dominant community accent, even if they are exposed to significant accent variation in their day-to-day home and social life (see also DeJesus et al., 2017, for data with bilingual children).

To summarize, there is a lot we do not know about how multi-accent exposure affects language development in young infants and toddlers, or how being tested in a different variety in the lab than children hear at home may affect performance in studies. But the little work that does exist suggests that these issues are imperative. On a theoretical level, multi-accent environments should be incorporated into our models of accent adaptation, and language acquisition more broadly. And on a more practical level, data on children’s development in multi-accent settings could inform debates regarding the best way to instruct multi-accent children, and ensure children’s accent background does not interfere with the delivery of appropriate speech and language services.

5.0 Wrapping things up

Children’s ability to cope with accent variation gradually improves between infancy and toddlerhood, with children’s processing of accented speech appearing to be qualitatively (though not quantitatively) adult-like by the time they reach their second birthday. Although children are exposed to a substantial amount of accent variation in their everyday lives, most research on early language development has focused on the processing of speech in the regionally dominant variety of the child’s native language. Moreover, psycholinguistics have, for the most part, not focused much on potentially important links between children’s perceptual sensitivity to accent variation, their ability to comprehend other-accented talkers, and the bi-directional relationship between
these two factors and the social inferences children draw about members of their community. By failing to fully consider the impact of these factors in developmental studies – treating children’s ability to handle accent variation as something we can study separately from language acquisition as a whole, and largely ignoring language variety differences in the input children receive – we have risked developing models of child speech and language development that are at best unrepresentative of the acquisition challenges experienced by many children in the world, and at worst a distortion of the overall acquisition process itself. We therefore suggest that future work in the field of children’s speech perception systematically take into account accent diversity. This will ultimately result in more inclusive models of early language processing.
References


Johnson, E.K., McQueen, J.M., Huettig, F., 2011. Toddlers’ language-mediated visual search: They need not have the words for It. Q. J. Exp. Psychol. 64, 1672–1682. https://doi.org/10.1080/17470218.2011.594165


https://doi.org/10.1016/j.jecp.2020.104933


https://doi.org/10.1080/0305764X.2016.1159660


https://doi.org/10.1016/j.jecp.2020.104807


https://doi.org/10.1080/15250000902994115


https://doi.org/10.1080/01690965.2011.597153


https://doi.org/10.1080/10489223.2015.1047096


Figure 1. Outline of children’s development of native language skills (blue), of their ability to cope with accents (green), as well as their ability to generate social inferences about speakers based on their language or accent (orange). Note that prior to 16 months, we know very little about how accent information affects infants’ social perception of the world.
Tables

Table 1  
Feature breakdown of the three accounts that explain the shift the ability to cope with accents between infancy and toddlerhood

<table>
<thead>
<tr>
<th>Account</th>
<th>Requires abstract phonological representations in infancy?</th>
<th>Assumes continuity in lexical representations?</th>
<th>Requires abstract phonological representations in toddlerhood?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exemplar-based Account</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td>Phonological Reorganization Account</td>
<td>Possibly, but only rudimentary</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Continuous Abstract Account</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 2  
Feature breakdown of the three types of explanations for toddlers’ exposure-based adaptation to unfamiliar accents

<table>
<thead>
<tr>
<th>Explanation</th>
<th>Adaptation accent specific?</th>
<th>Abstract phoneme representations required?</th>
<th>Can be influenced by non-linguistic factors?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Specific Mapping</td>
<td>Yes</td>
<td>Yes</td>
<td>Not specified</td>
</tr>
<tr>
<td>General Expansion</td>
<td>No</td>
<td>Not necessarily</td>
<td>Yes</td>
</tr>
<tr>
<td>Hybrid Flexibility</td>
<td>Sometimes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Running title: ACCENTS AND LANGUAGE ACQUISITION

1. In this review, our discussions of children’s accent processing have focused almost entirely on information conveyed in the speech signal. In everyday life, however, communication does not happen in a vacuum. Adult accent processing, for example, is affected by visual information, including the perceived race of accented interlocuters (Babel and Russell, 2015; Rubin, 1992). There is growing evidence that young children’s language processing is similarly sensitive to extra-linguistic contextual cues, such as a talker’s race (Singh et al., 2020; Uttley et al., 2013; Weatherhead and White, 2018). Considering these factors in our models of language processing may provide a fuller understanding of how children navigate their linguistic environment.

2. The developmental speech perception literature tends to emphasize how difficult children find it to process unfamiliar accents (e.g., Bent, 2018, 2014). But a much more positive outlook predominates in the speech production literature. Indeed, children are typically characterized as adopting new accents far more rapidly and successfully than adults (Chambers, 1992; Smith et al., 2007; Tagliamonte and Molfenter, 2007). To date, very little work has examined the intersection of accent perception and accent production in children (though see Paquette-Smith et al., in revision, for related work). Given claims that speech perception and production abilities are intimately related, even in infancy (e.g., Best et al., 2016; Bruderer et al., 2015), the lack of crosstalk between these literatures is surprising. Examining how the perception and production of accents relate could help researchers gain a new perspective on how children process other-accented speech – and acquire language more generally.