The role of social information in infants' and toddlers' linguistic representations

by

Drew Weatherhead

A thesis

presented to the University of Waterloo

in fulfillment of the

thesis requirement for the degree of

Doctor of Philosophy

in

Psychology

Waterloo, Ontario, Canada, 2017

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External Examiner	Kang Lee Professor
Supervisor(s)	Katherine White Associate Professor
Internal Member	Ori Friedman Professor
Internal Member	Elizabeth Nilsen Associate Professor
Internal-external Member	Grit Liebscher Professor and Department Chair

Author's Declaration

This thesis consists of material all of which I authored. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In seven experiments I demonstrate that social group information plays a role in infants' and toddlers' speech processing. In Chapter 2, I provide evidence that 16-month-old infants index social information to their linguistic representations. Specifically, infants do not automatically map the linguistic features of their own racial group to a speaker of an unfamiliar race. Instead, infants wait for experience with the speaker before deciding how to interpret her pronunciations. Chapter 3 demonstrates that race is indexed to toddlers' linguistic representations as they are being formed, even after only very little experience with a single group member. This information then influences their future interactions with speakers of the same race. Furthermore, I demonstrate that toddlers link fairly specific race information to linguistic representations. Finally, Chapter 4 shows that toddlers also use abstract cues such as previous affiliative behaviour to determine which social group an individual belongs to, and to predict how she will pronounce words. Overall, these studies demonstrate that social information is linked to infants' and toddlers' linguistic representations, and is used when interpreting speakers' utterances.

Acknowledgments

First and foremost I would like to thank my supervisor Katherine White. When I came to grad school I really did not believe that I could do it, and without you as my supervisor I honestly do not think I would have been able to finish. You were an amazing mentor and friend. The worst part about leaving grad school is leaving you. I would also like to thank the other faculty members in Developmental who contributed to my growth over the past five years. A special shout out to Ori who introduced me to research in social cognition, Stephanie who was always there for me when I needed advice, and to Liz who was able to step in at the last moment when I was in a bind.

I would also like to thank all my friends. Particularly my #squad: Darcy, Melissa, Jeff, Gord, Dan, Tim, and Michael. Shay and Madison for keeping me sane in Developmental, and for the many hours spent ranting, gossiping, and politicking. A special mention goes to my partner Tom for putting up with my rollercoaster of emotions throughout the last five years. Having someone to keep me grounded and put things in perspective has been very important. Then at the same time your ability to overreact on my behalf was also very helpful at times (I'm still worried you might try to sue SSHRC).

Then of course my family who have been incredibly supportive of me my entire life. My mom and dad, Nancy and Andy, my brother Alex, and my sister Aly. My parents worked so hard to ensure that I would have every opportunity to pursue my dreams. Without you both I would not have been able to get this far. I guess I really was a baby axfruit all along.

Table of Contents

List of Figures

List of Tables

1. Introduction

1.1 General Introduction

Language is highly variable, and this variation is strongly tied to the social world (e.g., Henrich & Henrich, 2007; Labov, 2006). A language community is a social group in which members share the same linguistic norms and expectations about how language will be used. Within a language community, despite some individual differences, there is a coherent linguistic structure (Labov, 2006); group members tend to use the same words and the pronunciation of these words tends to be consistent. At the same time, across language communities, differences in linguistic properties can be observed: people may use different pronunciations or other words entirely. For example, across speakers of English in the US, there are regional differences in pronunciations of certain vowels in different contexts (Labov, 1966; Labov, 1972; Labov 2006).

There are also dialectal differences within geographic regions that reflect social differences such as socioeconomic status. For example, Labov (1966) demonstrated that New Yorkers were socially ranked by their differential use of /r/ in postvocalic position (as in "car" or "four"). He found that individuals working in the highest ranked department stores had /r/ pronunciations rated highest in status, those in the middle-ranked stores had intermediate /r/ ratings, and those in the lowest ranked stores had the lowest ratings. Thus, within the context of New York City, /r/ is a strong social differentiator.

Throughout history we can see the consequences of linguistic division. Language communities have imposed language-related restrictions and policies on their members. For example, in the present day, The Charter of the French Language in Quebec

stipulates French text must appear larger than English text on any commercial signage, and English companies (even those such as Wal-Mart that originate in the United States) must translate their names to French or provide a French slogan on their storefronts. In more extreme cases, linguistic division has also led to the execution and genocide of speakers of a particular language (Phillipson & Skutnabb-Kangas, 1994; Shell, 2001; Sparks, 1996).

Adults are very sensitive to linguistic differences, and make rich social inferences on the basis of them, as in the study of r-pronunciation described above (e.g., Giles & Billings, 2004; Gluszek & Dovidio, 2010; Labov, 2006). Furthermore, social information about a speaker influences their speech perception (Drager, 2006; Ladefoged & Broadbent, 1957; Niedzielski, 1999). Children are also sensitive to linguistic variation, and make social inferences based on linguistic differences (Kinzler & DeJesus, 2013; Weatherhead, White, & Friedman, 2016); however, to date, there has been very little developmental work investigating how social group information may affect speech processing (both speech perception and speech production) in infancy and toddlerhood. In fact, developmental work on this topic has focused exclusively on children preschool-aged and older (e.g., Roberts, 1997; Roberts & Labov, 1995; Smith, Durham, & Fortune, 2007). Understanding how and when in development social information might be integrated into speech perception is critical for building a complete and accurate model of speech perception.

The overarching theme of my dissertation will be the effect of social group markers, in particular race and affiliation, on toddlers' linguistic expectations and word processing. In the following sections, I give a brief overview of existing literature on

early social processing, early language processing, language as a social marker, and how social information may be integrated into word processing.

1.2 Early Social Processing

Human beings are very sensitive to social information, and this sensitivity begins early in life. For example, children have many social preferences for, and draw inferences about, social groups based on a number of group-level properties such as gender, race, and age (Aboud, 1988; Baron & Banaji, 2006; Gelman, Collman, & Maccoby, 1986; Hirschfeld, 1996; Katz & Kofkin, 1997; Kircher & Furby, 1971; Maccoby & Jacklin, 1987; Martin, Fabes, Evans, & Wyman, 1999). Children even make inferences about members of arbitrarily defined groups (Sherif, Harvey, White, Hood, & Sherif, 1961), such as those based on t-shirt colours (Bigler, Jones, & Lobliner, 1997).

Many of these social preferences emerge in infancy. Very young infants prefer to look at faces of a more familiar gender (i.e., that of their primary caregiver: Quinn, Yahr, Kuhn, Slater, & Pascalis, 2002) and race (Bar-Haim, Ziv, Lamy, & Hodes, 2006; Kelly, Quinn, Slater, Lee, Gibson, Smith, Ge, & Pascalis, 2005). They also prefer to attend to faces of other infants over faces of adults (Bahrick, Netto, & Hernandez-Reif, 1998; McCall & Kennedy, 1980; Sanefuji, Ohgami, & Hashiya, 2006). In general, infants prefer others with whom they share similarities (or with whom they have more experience). There are many cues that infants can use to determine whether individuals are members of the same social group. This dissertation will focus on two major social indicators: (1) Race, and (2) Affiliative behaviour. In the next sections, I will provide an overview of work that has investigated infants' processing of race and affiliative behaviour.

1.2.1 Race

For the purposes of this dissertation, race is defined as an individual's skin colour and facial characteristics. Infants are very sensitive to race. By 3 months, infants prefer to attend to same-race faces over other-race faces (Bar-Haim et al., 2006; Kelly et al., 2005). With age, infants become less capable of discriminating or recognizing other-race faces; by 9 months, infants categorize faces by race and only recognize individual same-race faces (Anzures, Quinn, Pascalis, Slater, & Lee, 2010; Kelly, Liu, Lee, Quinn, Pascalis, et al., 2009; Kelly, Quinn, Slater, Lee, Ge, et al., 2007), and their scanning patterns for same-race and other-race faces differ (Wheeler, Anzures, Quinn, Pascalis, Omrin, & Lee, 2011). This behaviour appears to be driven by familiarity. For example, Black infants residing in Africa attend longer to Black faces than White faces, but Black infants residing in Israel look equally long at Black and White faces (Bar-Haim et. al, 2006; Kelly et al., 2005).

Interestingly, infants' ability to discriminate between people of different races does not to lead to a social preference for individuals of their own race. When given the option to accept a toy from a same-race or other-race individual, 10-month-old infants take toys equally from the two speakers. Similarly, 2.5-year-old toddlers offer toys equally to same-race and other-race individuals. Only at 5 years old do children demonstrate an explicit social preference for same-race speakers (Kinzler & Spelke, 2011). Once developed, these race-based social preferences persist into adulthood. In fact, adults not only prefer same-race individuals, they also use race as a social group indicator, and to make social judgements about an individual. For example, race has

been shown to influence American adults' judgements of whether an individual is American or foreign (e.g., Devos & Banaji, 2005). Even 4- to- 6-year-old children use race as a marker of whether someone is from their own country or a country far away, though it is not considered as strongly as linguistic cues (Weatherhead, Friedman, & White, 2017).

Although race does not appear to play a strong role in infants' social preferences until later on, there is some evidence that infants use race as a linguistic marker. For example, 6-month-olds match other-race faces with non-native languages (Uttley, de Boisferon, Dupierrix, Lee, Quinn, Slater & Pascalis, 2013). Thus, infants appear to have some understanding of same-race and other-race speakers as separate groups, and different beliefs about the languages these groups speak – in particular, they infer that same-race individuals speak in a familiar way, and other-race individuals speak in a novel way.

1.2.2 Affiliative Behaviour

Adults reason about others' affiliative relationships with relative ease (e.g., Seyfarth & Cheney, 2012; 2013; Cosmides, Tooby, & Kurzban, 2003), and base their own affiliative behaviour on similarities between themselves and other individuals. For example, adults are more likely to affiliate with individuals who have the same beliefs as them (e.g., a shared religion or political belief) than individuals with differing beliefs (e.g., Byrne & Nelson, 1965).

Infants also appear to be able to reason about two individuals' social relationships or behaviour (e.g., Kuhlmeier, Wynn, & Bloom, 2003; Mascaro & Csibra, 2012; 2014;

Powell & Spelke, 2013; Spokes & Spelke, 2016; 2017). There is even some work suggesting that infants can reason about third-party affiliative relationships (i.e., whether two individuals are affiliated with the same social group), just as adults do. For example, 9-month-olds expect that individuals who share the same food preferences will later have a positive interaction, while two individuals with opposing preferences will later have a negative interaction (Liberman, Kinzler, Woodward, 2014).

These studies focus on infants' expectations of affiliation based on previous compatible or incompatible information (that is, whether their preferences or behaviour are the same or different). Overall, while there is some work on infants' and toddlers' processing of affiliative behaviour, there has been comparatively less work done on this topic than on their processing of race.

1.3 Language as a social marker

Infants and toddlers are clearly tuned into the social world. However, whether they recognize that this social information is tied to linguistic variation remains to be seen. As we saw earlier in the work of Labov, linguistic differences are quite prevalent in the real world, and as adult listeners, we have developed strong associations between social information and linguistic properties. The existence of these associations has two implications: 1) Listeners use linguistic information as an indicator of social dimensions of a speaker, and 2) Social dimensions of a speaker shape listeners' expectations about their speech. In this section I focus on both these implications in turn.

1.3.1 Social Inferences based on Language and Accent

Accent is defined as the distinctive pronunciation style in a given language. Although listeners typically attribute accents only to speakers who talk differently than them, in reality everyone has an accent. Speakers simply either have the same accent or a different accent than the listener. For the purposes of clarity I will refer to native accents as either native-accented or unaccented speech, and non-native accents as either foreign-accented or accented speech, throughout this dissertation.

1.3.1.1 Adults

Adults make many inferences about a speaker based on their accent, such as their nationality, place of origin, ethnic group, social status, sexual orientation, etc. (Labov, 2006). Additionally, adults attribute different personality traits, such as friendliness, warmth, or reliability, as well as other traits like intelligence or physical attractiveness to speakers of different dialects and accents (Bayard, Weatherall, Gallois, & Pittam, 2001; Campbell-Kibler, 2006; Dornic, Nystedt, Laaksonen, & Arberg, 1989; Paltridge & Giles, 1984). Adults will even infer different social properties for speakers producing the same content depending on their supposed age, speaking rate, and accent (Giles, Henwood, Coupland, Harriman & Coupland, 1992).

Perhaps unsurprisingly, accent plays a large role in adults' social preferences. For example, accent appears to weigh more heavily in adults' social categorizations than other salient cues such as ethnicity (Rakic, Steffens, & Mummendey, 2011). Adults rate native-accented speakers more favourably than foreign-accented speakers, even when both speaker types are equally intelligible (Bresnahan, Ohashi, Nebashi, Liu, & Shearman, 2002). For example, White and Hispanic American adolescents in California both rate American-accented speech more favourably than Spanish-accented speech.

However, those Hispanics with higher exposure to Spanish show less of a bias than those with lower exposure (Dailey, Giles, & Jansma, 2005). These types of preferences are true for non-foreign accents as well. For example, Canadian listeners rate Jewish-accented speakers as shorter, less good-looking and less likely to be in a leadership position than Canadian-Accented speakers, even if the listeners themselves are Jewish-Canadian (Anisfeld, Bogo, & Lambert, 1962). Similarly, American-English and Received-Pronunciation (British) dialects are favoured heavily over Australian or New Zealand English dialects (e.g., Huygens & Vaughn, 1984; Stewart, Ryan, & Giles, 1985; Bayard et al. 2001).

Some work has even suggested that the more accented the speech is, the less favoured it will be (Ryan, Carranza, & Moffie, 1977). However, other work has shown that this is not always the case. For example, Cargile & Giles (1997) found that while the "Japaneseness" of an American accent did influence participants' ratings of the speaker, the strength of accent did not, nor did it have an effect on listeners' arousal levels.

Thus, for adult listeners, it appears that knowing how a person talks is to assume what type of person they are. Many of these inferences are not accurate, like the inference that accented people are less friendly or intelligent, and are likely born out of a preference for people who are similar. However, some of these associations are grounded in the real world: for example, a person who speaks with a Canadian Accent is likely to be from Canada. The latter types of associations are likely due to the great deal of experience adults have with both linguistic variation and social variation. Although infants and children have far less experience, they show some of the same tendencies.

1.3.1.2 Infants and Children

Even very young infants have some expectation that speakers of the same language are members of the same social group. For example, 9-month-olds are more surprised when individuals who speak the same language have a negative interaction than when they have a positive interaction (Powell & Spelke, 2013). However, this is only found when the speakers use the infants' native language; when the two speakers use the same foreign language, infants do not expect a positive nor negative interaction. When two individuals speak different languages (with one speaker speaking their native language and the other speaking a foreign language), infants are more surprised when they later have a positive interaction than when they have a negative interaction. These results suggest that infants expect speakers of their native language to have affiliative relationships with each other, but not with speakers of a foreign language.

Young children recognize that not all people speak the same language; however, previous work has suggested that it is only at 6 years old that children infer that individuals from different cultures speak different languages (Kuczaj, 1982; Kuczaj & Harbaugh, 1982). Similarly, 6-year-olds assert that language differences are caused by nationality differences, and that a shared language is caused by shared nationality (Jahoda, 1961; Piaget & Weil, 1951; also see Hirschfeld & Gelman, 1997, for similar discussion). When explicitly told that an individual speaks a certain language, 6-year-olds reliably use this information to predict the individual's national group (Penny, Barrett, & Lyons, 2001). Thus, by age six, children infer that speakers of foreign languages are from different places or cultures, and speakers of the same language are from the same place or culture. Similarly, 6-year-olds infer that speakers of a foreign language were

born in, and currently live in, far away places (Weatherhead et al., 2017). Finally, 5-year-olds infer that speakers of a foreign language are from a different racial group, wear unfamiliar garb and live in novel looking houses (Hirschfeld & Gelman, 1997).

In terms of accented speech, there is some evidence that even young children are aware of the relation between linguistic variation and geographic background. Preschool-aged children recognize that two speakers with the same foreign accent live in similar places, while speakers with different foreign accents live in different places (Weatherhead, White, & Friedman, 2016). Five- and 6-year-old children categorize speakers based on their regional dialect (Wagner, Clopper, & Pate, 2014). Four-, 5-, and 6-year-olds infer that speakers with foreign accents were born in far away places (Weatherhead et al., 2017). Likewise, in a forced choice task, 5- and 6-year-olds use accent to infer who is American, or "lives around here" (Kinzler & DeJesus, 2013). When presented with two speakers, one who shares the same native accent as the child and the other who has a novel French accent, 5-to-6-year-old children are more likely to infer that the speaker who shares their native accent is American.

As with adults, language and accent weigh heavily in infants' and children's social preferences. For example, infants as young as 10-months-old prefer to interact with, and receive a toy from, a speaker of their native language rather than a speaker of a foreign language. Furthermore, infants selectively eat foods that are endorsed by speakers of their native language (Shutts, Kinzler, McKee, & Spelke, 2009). These preferences extend and strengthen into childhood: 5-year-old children would rather be friends with a speaker of their native accent over a foreign language or foreign accent speaker (Kinzler, Shutts, DeJesus, & Spelke, 2009). These preferences have been replicated cross-

culturally and with both monolingual and bilingual children (e.g., Okumura, Kanakogi, Takeuchi, & Itakura, 2014; Souza, Byers-Heinlein, & Poulin-Dubois, 2013).

At age 5, children prefer both own-race and native-accented individuals, when each category is tested in isolation (Aboud, 1988; Kinzler et al., 2009). When the two social cues are pitted against each other, White children in the United States prefer to be friends with native-accented, other-race individuals over foreign-accented, same-race individuals (Kinzler et al., 2009). Thus, like adults, language is more heavily weighted in children's social categorizations than other salient cues such as race. However, there are some cases in which other cues play a stronger role. For example, 5- and 6-year-old children prefer a nice foreign-accented speaker to a mean native-accented speaker (Kinzler & DeJesus, 2013). Moreover, native Xhosa-speaking children living in South Africa, and attending school in English, express social preferences for speakers of English over speakers of Xhosa, even when tested by a Xhosa-speaking experimenter (Kinzler, Shutts, & Spelke, 2012). The experimenters suggest that this preference is due to English being a more prestigious language than Xhosa. Thus, children may weigh socially desirable traits like friendliness or prestige more heavily than whether the accent/language is native.

Overall, linguistic properties of a speaker appear to play a strong role in infants' and children's social preferences. Additionally, children appear to use accent and language in their social categorizations, and make inferences about speakers accordingly.

1.3.2 Social Influences on Speech Processing

This dissertation will explore whether social information about a speaker affects how infants and toddlers interpret or recognize words. To date there has been no work looking at the influence of social information on infants' and toddlers' speech processing. However, there has been work investigating the effects of social information on adults' speech processing. For adult listeners, the same speech sequence can be interpreted differently, and be better or worse understood, based on social properties of the speaker such as gender, age, social class, and nationality (e.g., Ladefoged & Broadbent, 1957). For example, females have a higher acoustic boundary between /s/ and /ʃ/ in their productions. That is, on a continuum from /s/-/ʃ/, the boundary between /s/ and /ʃ/ (essentially where a /s/ becomes a /ʃ/) is closer to /s/ for a female speaker than a male speaker. As a result, people are more likely to perceive a sibilant on a /s/-/ʃ/ continuum as /ʃ/ when shown a photograph of a woman (Strand & Johnson, 1996). Also, the gender of a visually presented face affects the perceptual identification of phonemes on a continuum between [ʊ] and [ʌ] ("hood" and "hud"). This is found for both stereotypical and non-stereotypical faces (e.g., a feminine female face vs. a masculine female face), with the more stereotypical having an even larger effect (Johnson, Strand, & D'Imperio, 1999). Likewise, the perceived age of a voice has similar influences on vowel perception (Drager, 2006).

One of the most profound social factors affecting speech perception is nationality. Niedzielski (1999) found that Detroiters shifted in their perception of the /aʊ/ diphthong when the word "Canadian" or "Detroiter" appeared at the top of their answer sheet. Hay et al. (2006) report on a similar effect in New Zealand, where New Zealanders shifted in their perception of /i/ depending on the presence of the word "Australian" or "New

Zealander." These effects are seen even when listeners are primed with these properties very subtly. For example, Hay & Drager (2010) exposed New Zealander participants to either a stuffed kangaroo (associated with Australia) or a stuffed kiwi bird (associated with New Zealand) prior to completing a vowel perception task. They found that participants' vowel perception shifted as a function of the exposure toy, such that those who saw the kangaroo were more likely to classify vowels as Australian-like than participants who saw stuffed kiwis.

Another important social factor affecting speech perception is race. Like Kangaroos and Australians, there can be strong ties between physical appearance and geographic origin. American listeners understand unaccented English better when it is paired with a picture of a Caucasian face than with a Chinese face (Kang & Rubin, 2009; Rubin, 1992; consistent with Babel & Russel, 2015). Likewise, Mandarin-accented English is better understood when it is paired with a Chinese face than with a Caucasian face (McGowan, 2015). Together, these studies suggest that adult listeners form associations between properties of social groups and accent, which in turn lead to expectations that affect speech processing. These effects can be seen in adults' productions as well. For example, African American individuals (including Oprah Winfrey) shift their speech style depending on the race of their conversation partner (Hay, Janned, & Mendoza-Denton, 1999; Rickford & McNair-Knox, 1994). These results are consistent with the idea that listeners' speech styles are designed primarily for their audience, and a speech style can be predicted based on the associations between a given group and its linguistic features (Bell, 1984; 2001).

To date there has not been research investigating the effect of social information on infants' and toddlers' word processing. But before it can be determined how social information might be linked to infants' and toddlers' word representations, a more basic understanding of how infants tune in to their native language, and more importantly, their native accent, is necessary. In the following section I review some of this work.

1.4 Early Language Processing

1.4.1 Tuning into their Native Language

Infants are tuned into speech from birth. In fact, even very young infants have a preference for listening to speech over non-speech signals (e.g., Columbo & Bundy, 1981; Vouloumanos & Werker, 2004), with some evidence suggesting that this preference is present as early as birth (Vouloumanos & Werker, 2007; though this preference may be more broadly defined, see Vouloumanos, Hauser, Werker, & Martin, 2010). Thus from a very early age, infants recognize that language is a special signal of sorts, and have an interest in language stimuli.

Infants' language preferences appear to be determined by their language environment. This is true even of their prenatal language environment. Newborn infants prefer to listen to a language rhythmically similar to the language heard in utero as opposed to one that is not (Byers-Heinlein, Burns, & Werker, 2010; Moon, Cooper, & Fifer, 1993). Furthermore, 2-day-olds prefer to attend to their own language over a foreign language (Moon et al, 1993). As their experience with their native language grows, their preference for their native language becomes even more specific. Four- to 5-month-old infants attend to their native language over a non-native language of the same

rhythmic class (Bosch & Sebastian-Galle, 1997), and 6-month-olds prefer to listen to their native language spoken naturally rather than unnaturally (backward speech) (Kinzler, Dupoux, & Spelke, 2007). At 9-months-old, infants prefer to listen to words that follow the same stress patterns and phonotactic rules as their native language over structures found in other languages, whereas 6-month-olds do not show this preference (Jusczyk, Cutler, & Redanz, 1993; Jusczyk, Friederici, Wessels, Svenkerud, & Jusczyk, 1993).

As infants become more attuned to the specific properties of their native language, their sensitivity to properties of other languages decreases. The most dramatic instance of this can be seen with sound discrimination: 6-to-8-month-olds are able to discriminate minimally different phonetic contrasts that exist both in their native language, and in non-native languages. But by 10-to-12-months, infants have difficulty discriminating some sound contrasts that do not exist in their native language (Kuhl, Williams, Lacerda, Stevens, & Lindblom, 1992; Werker & Tees, 1984; Saffran, Werker, & Werner, 2006).

Thus, infants have a special interest in their native language, and they begin to ignore information that is not relevant to their native language. Gradually over the course of their first two years, infants begin to learn words and amass a surprising receptive vocabulary. However, as I have discussed, there is a lot of variability in how words are pronounced within a language. How do infants contend with this variability?

1.4.2 Infants' and Toddlers' Processing of Accented Words

Infants are quite good at recognizing the sounds of words. For example, after being familiarized to new words in the lab, 7.5-month-olds will prefer to listen to passages containing those words over passages that do not contain those words (Jusczyk & Aslin, 1995). However, this effect is not found when infants are familiarized with mispronunciations of the words in the passage, either at the word onset (e.g., "bog"; Jusczyk & Aslin, 1995) or word offset (e.g., "dob"; Tincoff & Jusczyk, 1996). Similarly, young infants fail to recognize newly familiarized words when they are presented in a new accent at test (Schmale & Seidl, 2009; Schmale, Cristia, Seidl, & Johnson, 2010). Thus, infants' word representations seem quite specific.

Other studies have examined infants' recognition of known words; that is, words known prior to coming into the lab (unlike the above-mentioned studies, which familiarized infants with new words in the lab). In these studies, 11-month-old infants show a preference for listening to known words (e.g., high-frequency words like "bottle") over low-frequency or non-words, which are unfamiliar to them (e.g., "boogle"). However, they do not show this preference if the known words are accented or mispronounced by even a single-feature, at least in stressed syllables (Halle & de Boysson-Bardies, 1996; Swingley, 2005). When the words are produced in an unfamiliar accent, infants do not show a preference for the known words even at later ages (Best, Tyler, Gooding, Orlando, & Quann, 2009), unless they are given sufficient exposure to the accent first (van Heugten & Johnson, 2014).

In studies that have tested word comprehension (rather than recognition of word forms), young toddlers similarly struggle to recognize accented words. For example, when presented with two visual referents (e.g., a dog and a novel distractor) and an

accented word that corresponds to one of these two objects (e.g., "dog"), infants under 19-months-old fail to recognize that accented pronunciations map onto familiar referents (White & Aslin, 2011; Mulak, Best, Tyler, & Kitamura, 2013), suggesting that infants do not access the meaning of accented words. By 25-months-old, however, toddlers can understand accented familiar words under some conditions, regardless of prior exposure to the speaker or the accent (van Heugten, Krieger, & Johnson, 2015). However, children continue to struggle to recognize accented words throughout development (e.g., Bent, 2014; Nathan, Wells, & Donlan, 1998; Newton & Ridgway, 2016).

1.5 Incorporating Social Information in Speech Perception

As mentioned previously, adults have vast experience with both social and linguistic variation, and the co-occurrences between them. Adults' social expectations influence how they perceive a speaker's speech, suggesting that social information is tied to their linguistic representations. However, while adults have many experiences to draw from during speech perception, infants and toddlers do not. This provides a unique opportunity to explore the formation of the link between social and linguistic information. For example, how do infants and toddlers interpret the speech of a type of speaker for which they have no previous experience? If social information is not linked to their linguistic representations, then they should process words the same way, regardless of social information about the speaker. When encountering a speaker from an unfamiliar social group, they should process that speaker's speech in the same way as they would for a speaker from their own social group. Thus, they would be extending their experience with their immediate social group to an unfamiliar group (in the case of

the following experiments, Caucasian speakers with a South Western Ontario Dialect of Canadian English). However, if social information is indexed to infants' and toddlers' linguistic representations, then experiences with their own linguistic group should not be extended to a speaker from an unfamiliar group.

Virtually no existing work has empirically tested the nature of sociolinguistic processing in infants and toddlers. However, there have been some theoretical observations. Foulkes and Docherty (2006) suggest that children adapt to their social world throughout the course of development, re-weighting and re-defining social categories as they become relevant. Specifically, they suggest that children may retain all the salient details associated with an individual, even if children are unaware that these details are linguistically relevant, and over time the associations between linguistic categories and these social characteristics emerge. At certain stages of development only some social categories may be considered linguistically relevant. For example, children may first tune into race as being a linguistically relevant social marker, while ignoring more abstract information like behaviour or preferences. Foulkes and Docherty also suggest that the correlations between more arbitrary, less salient, social information and linguistic information may not occur until adulthood. Finally, they suggest that it is only through direct experience with a social group, and once the amount of experience with this group reaches a specific threshold, that the construction of a sociolinguistic category occurs. At present, it is unknown whether young children establish adult-like sociolinguistic categories both in terms of how they are structured and the type of information they contain.

To even begin to answer this question, we first need to know how much experience is necessary to form a sociolinguistic category, and what about the experience is stored. With respect to the first question, Foulkes and Docherty (2006) suggest that because sociolinguistic categories are formed through experience with variation, they may take an extended amount of time to develop. It could be that vast experience with a social group is necessary before group-level associations are formed. Thus, infants and toddlers would be working on a more "speaker by speaker" basis, learning about each new speaker as they encounter them until some threshold is reached. To address this question, I ask how experience with a new type of speaker affects future interactions with similar (or dissimilar) speakers. I ask whether toddlers are able to learn about a particular type of speaker and apply this information to a new speaker of the same type right away.

In regards to the second question, Foulkes and Docherty (2006) do not speculate about what children may be specifically encoding about new speakers. Infants and toddlers could begin by tracking features of the speaker more generally (e.g., speakers "like me" vs. speakers different than me) in which case they may over-generalize the category. Or the features they are tracking could be hyper specific (e.g., speaker is ethnically Indian, tall, long hair, female, etc.) in which case they may under-generalize the category. I ask how specific the social information that infants and toddlers track is.

One way to address these questions is to use race as the dimension varying across speakers. Even very young infants are highly sensitive to race and are able to discriminate between races (e.g., Bar-Haim et al., 2006; Kelly et al., 2005). Importantly, race serves as an important linguistic marker for adult listeners. As such race is the social indicator used in the following two chapters. In particular, I ask whether the social

information linked to infants' and toddlers' linguistic representations is specific to each type of speaker (i.e., properties of Caucasian speakers vs. Indian speakers vs. Chinese speakers).

However, while race is certainly a salient social category marker, it is not always a marker of linguistic group membership. For example, a Russian speaker and a Canadian speaker may be virtually identical in terms of appearance, but speak very different languages. Adults use "invisible cues" such as nationality as predictors of linguistic properties (e.g., Hay et al, 2006). Thus, the type of social information indexed to sociolinguistic representations must include not only salient concrete features (e.g., age, race, gender), but also more abstract features (e.g., nationality, socioeconomic status, intelligence). Foulkes & Docherty (2006) suggest that for social categories that are not visually transparent (like race and gender), social information may take a great deal of time to be indexed to linguistic properties. I hypothesize that toddlers interpret the speech of speakers from the same social group in the same way, even if the cue to group membership is more abstract. To investigate this question, I ask whether toddlers use a speaker's previous affiliative behaviour as a cue to group membership. If toddlers use behavioural features of a speaker, in addition to physical characteristics, it would suggest that a social cue does not have to be a visible feature of the speaker in order to be associated with linguistic properties.

1.6 Present Studies

The present studies explore the role of social information on infants' and toddlers' word processing.

The first chapter investigates the role of race in infants' word processing. Three experiments investigate the effect of speaker race on 16-month-olds' recognition of familiar words in the context of a familiar accent and an unfamiliar accent. This chapter addresses how infants interpret the speech of new types of speakers for which they have no previous experience.

In chapter two, three experiments determine if, when toddlers are familiarized with the linguistic properties of an individual of a different race, they interpret productions from new speakers of that race the same way. This chapter provides insight into what information toddlers store about a new speaker (i.e., is race indexed?), and whether they generalize this information to new speakers of the same type. Additionally, the specificity of these race-based generalizations is addressed. That is, are toddlers learning about a specific type of person or are they learning something more general about people different from them?

In the final chapter, I look at the role of affiliative relationships in toddlers' word processing. In particular, I look at whether toddlers generalize linguistic properties across social group members based on previous affiliative behaviour. This chapter addresses whether abstract social information is indexed to toddlers' word representations.

2. Infants' interpretations of an other-race speaker's pronunciations

This chapter focuses on infants' initial processing of speech from a new speaker who belongs to a social group for which they have little previous experience. Specifically, I ask if a speaker's race impacts infants' recognition of familiar and unfamiliar pronunciations of familiar words. As previously mentioned, young language learners initially struggle to recognize familiar words when they are produced with an accent, at least in the absence of a learning period (Best, Tyler, Gooding, Orlando, & Quann, 2009; van Heugten & Johnson, 2014; van Heugten & Johnson, 2015; White & Aslin, 2011). For example, without prior exposure to the accent, 15-month-olds do not look preferentially at a target object when its label is produced with an accent (Mulak, Best, Tyler, Kitamura, & Irwin, 2013). One interpretation is that infants' difficulty with unfamiliar pronunciations is due to their lack of exposure to variety. That is, infants have only heard a small number of speakers produce words with a narrow range of variation. If social information is not initially indexed to infants' linguistic representations, then infants should initially process new speakers, of any kind, in a way that is consistent with their previous experience with how words are pronounced. In this case, when an unfamiliar accent is encountered, existing linguistic representation are unable to account for this variability, leading to processing difficulties.

However, infants could have linked these familiar pronunciations with a familiar social group. Just as infants have only heard words produced in a narrow range of pronunciations, most have only heard words produced by a narrow range of (in many cases, same-race) people. Thus, infants may not have much experience with other-race speakers and how they pronounce words. If social information is indexed to infants'

speech representations, then when encountering a speaker from a novel social group, recognition of familiar pronunciations may fail, as they have no prior experience with speakers from this group. Furthermore, infants may more readily accept unfamiliar pronunciations of words from an other-race speaker than a same-race speaker because they are not constrained by their previous experience.

Experiment 1 tested whether infants' processing of familiar and unfamiliar pronunciations of familiar words differs for other-race and same-race speakers. Experiment 2 tested whether infants were learning systematic pronunciation differences, that is, whether they are willing to accept non-systematic mispronunciations from an other-race speaker, simply because they look unfamiliar. Experiment 3 replicated the findings of Experiments 1 and 2 with a different other-race speaker.

2.1 Experiment 1

Infants heard familiar words produced in their native accent (e.g. "dog") and in an accent involving a vowel shift (e.g. "dag"), in the context of either a same-race speaker or an other-race speaker. If infants' familiar words representations are linked to their own social group, they should expect to hear familiar, or natively accented, pronunciations from the same-race speaker, and not unfamiliar pronunciations. However, when they encounter the other-race speaker, who belongs to a social group with which they are not familiar, infants may respond in one of two different ways. If social information is not indexed to infants' linguistic categories, then they should treat the other-race speaker as a same-race speaker and accept only the natively accented pronunciations. However, if

social information does play a role, then they should treat the other-race speaker differently than the same-race speaker.

2.1.1 Participants

Forty 16-month-old infants were tested (23 females; mean age: 16 months 0 days; age range: 15;16-16;16). Nine additional participants were tested, but not included due to non-completion (3), failure to attend to both objects during the baseline period for at least half of each trial type (3), or an overall difference score exceeding 2.5 standard deviations from the mean for either the unaccented or accented word trials (3).

Infants were randomly assigned to one of two conditions: Same-race speaker or Other-race speaker. Participants in both conditions were monolingual English-learners and Caucasian. Overall, participants had very minimal exposure to speakers who spoke a foreign language, had a non-local English accent, or were of a different race (average exposure per week was 2.6%, 7.2%, and 7.3%, respectively, as indicated by parental reports; by condition: Same Race Condition - 3.1%, 7.2%, and 7.2%, respectively; Other Race Condition - 2.1%, 6.5%, and 7.5%, respectively).

2.1.2 Stimuli

Audio Stimuli

The audio stimuli were modelled on White & Aslin (2011). The test words were six words highly familiar to 16-month-olds (Dale & Fenson, 1996), all containing the same vowel, /a/: “ball”, “block”, “bottle”, “car”, “dog”, and “sock”. All of these words are comprehended by 67%-95% of children by 15-months of age according to the MacArthur

Communicative Development Inventories (Dale & Fenson, 1996)[1]. A female native speaker of English produced each word four times, twice in natively accented pronunciations (i.e., the native accent of the child), hereafter referred to as unaccented pronunciations, and twice with an unfamiliar accent, hereafter referred to as accented pronunciations, in which the /a/ vowel was shifted to /ae/ (i.e., "bottle" to "battle", "sock" to "sack", etc.). Each version was produced in each of two sentence contexts, "Do you see the X" or "Find the X". All sentences were naturally produced in an infant-directed-manner. The same audio recordings were used for both conditions. Stimuli were recorded in a sound-treated booth at a sampling rate of 44100 Hz and equated for amplitude in Praat (Boersma & Weenink, 2009).

Percent known at 15 months	Unaccented Pronunciation (a)	Accented Pronunciation (æ)	Random Mispronunciation
95.3	Ball	Bæll	Bull
67.2	Block	Blæck	Blick
84.4	Bottle	Bættle	Boottle
82.8	Car	Cær	Cor
87.5	Dog	Dæg	Dag
78.1	Sock	Sæck	Seck

Table 1. Words used in Experiments 1-3. The accented pronunciations used in Experiment 1 and 3 contained a systematic vowel shift consistent across words. The

[1] Parental reports in this study indicate that for each test word, 78%-97% of children had "seen the object before and understand the word very well". There were no differences across conditions (Same Race: average = 86%, range = 76%-96%, Other Race: average = 91%, range = 80%-99%; in both conditions the least known word was "Bottle" (76% and 80%) and the most known word was "Ball" (96% & 99%).

random mispronunciations from Experiment 2 and 3 were not consistent. Percent known refers to the percentage of parents who report that their child comprehends the word, according to the MacArthur Communicative Development Inventories (Dale & Fenson, 1996).

Visual Stimuli

Depending on the condition, participants either saw a still image of a same-race woman or an other-race speaker. The same-race woman was a 22-year-old Caucasian with pale skin and long brown hair. The other-race woman was a 23-year-old mixed-race female with Black, Caucasian and Native-Canadian heritage. Like the same-race speaker, she had long brown hair.

A B C

Figure 1. A: Same-race speaker (Experiment 1). B: Other-race speaker (Experiments 1 and 2). C: Other-race speaker (Experiment 3).

Six familiar-unfamiliar object pairs were created. Each object appeared in an outline of a box on either the right or left side of the screen (counterbalanced). The pairs were as follows: bottle – stapler, ball – abacus, block – hair dryer, car – turkey baster,

dog – hourglass, sock – can opener[2].

2.1.3 Procedure

The participant sat on his/her parent's lap approximately 1.5 ft. from a 36x21-inch plasma screen television in a sound-treated testing room. A camera under the television recorded the child's looking behaviour for the entirety of the session. The camera was linked to a monitor and recording device in the lab area adjacent to the testing room for the experimenter's viewing purposes and for later off-line coding. Stimuli were played at approximately 65dB and presented in Psyscope X (Cohen, MacWhinney, Flatt & Provost, 1993). Parents were instructed not to interact with their infants during the session and wore noise-cancelling headphones playing instrumental music to mask the audio being played to the infant.

Infants first viewed a silent 8-second introductory video of the speaker smiling and waving to ensure they recognized that the speaker was a real agent. Infants then completed a total of 24 test trials, in two consecutive blocks of 12 trials. In each block, each of the test words occurred twice, once accented and once unaccented (whether the unaccented or accented version occurred first was counterbalanced across words and participants). Each trial was 10 seconds in length. At the start of each trial, a static image of the speaker's face and shoulders appeared at the top center of the screen for two seconds, with two black outlined boxes appearing on either side of the screen. Following this, the speaker's face disappeared, and an object appeared in each of the two outlined boxes (see Figure 2). One object corresponded to the test word (i.e., the target object),

[2] Parental reports confirmed that the target images were familiar to the participants and that distractor objects were unfamiliar to the participants.

and the other object was a novel distractor (i.e., the distractor object).

These two objects stayed on the screen for eight seconds, the first three seconds of which was a silent baseline period, followed by an audio recording of the test word in the naming phrase (either "Do you see the X" or "Find the X"). Each block was pseudo-randomized such that the target object was never on the same side for more than three trials in a row, the same sentence context did not occur more than two trials in a row, no more than three accented or unaccented words occurred in a row, and the same word did not occur fewer than four trials apart.

2.1.4 Coding of looking times.

Looking time was coded off-line using customized software (Brown University), frame-by-frame (1 frame = 33 msec). Looking proportions to the objects were determined for the baseline period and for the test period, which began 300ms after the onset of the test word to account for the time necessary to program an eye movement (e.g., Swingley & Aslin, 2002). Both the baseline and test period were 3 seconds in length.

For all experiments in this dissertation looking behaviour was hand coded by either myself or research assistants. Each frame was assigned an L, R, or O depending on if the participant was fixated on the left side of the screen, the right side of the screen, or another area of the screen/not at the screen. Random videos were selected for reliability testing in which a second coder would independently code the video to ensure that videos were being coded consistently and correctly.

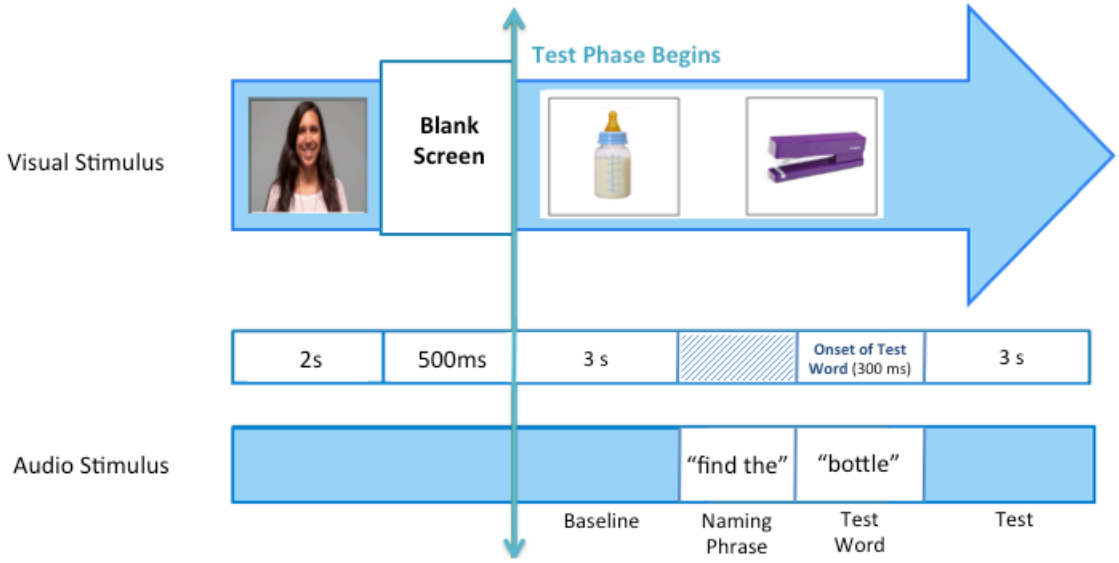

Figure 2. An example of a test trial in Experiment 1. This example depicts an unaccented trial for the Other-race Speaker. Note that the test phase began 300ms after the onset of the test word ("bottle" in this example).

2.1.5 Results and Discussion

For both the baseline and test periods, the proportion of time infants spent looking at each of the objects was computed (out of the total time looking at either object during that 3-second period). Trials in which infants did not look at both objects during the baseline period (or at either object during the test period) were not included in the analyses. There was no difference in the percentage of discarded trials across conditions, $t(38) = -.28$, $p = .783$. In addition, the proportion of time that infants spent looking at the familiar object during baseline was equivalent between conditions, $t(38) = 0.48$, $p = .144$ (.53 in the same-race condition and .50 in the other-race condition).[3]

[3] Additionally, an independent t-test revealed that infants in the two conditions paid an

To assess infants' recognition of the words, a difference score was calculated for each trial using the looking proportions for each period (proportion target $object_{test}$-proportion target $object_{baseline}$). This measure indicates the change in looking towards the target object after labeling. Note that a difference score of zero (no change following labeling) indicates a failure to recognize the word.

To determine the effect of accent on infants' word processing across the whole experiment, a mixed measures ANOVA with the within-subject factor of word type (Unaccented vs. Accented) and a between-subject factor of condition (Same-race vs. Other-race) was run. No main effect of word type was found, $F(1, 38) = 2.72$, $p = .107$, but there was a significant main effect of condition, $F(1, 38) = 6.78$, $p = .013$, and a significant word type X condition interaction, $F(1, 38) = 12.96$, $p = .001$. Thus, infants interpreted the same words differently depending on which speaker they saw.

For infants in the same-race speaker condition, paired sample t-tests comparing unaccented pronunciations to the accented pronunciations revealed that infants interpreted the unaccented and accented pronunciations differently, $t(19) = 4.13$, $p < 0.001$. One-sample t-tests against chance (zero change) revealed that for the unaccented pronunciations, infants' looking increased significantly to the target object, $t(19) = 4.97$, $p < .001$. For the accented words, infants did not increase their looking to either object, $t(19) = -1.48$, $p = .155$. Thus, for the same-race speaker, infants recognized only the unaccented pronunciations.

---

equivalent amount of attention to the speaker overall during the 2-second speaker presentation prior to the objects' appearance, $t(38) = .92$, $p = .362$.

For infants in the other-race speaker condition, however, paired sample t-tests comparing unaccented pronunciations to the accented pronunciations revealed that infants did not interpret these two types of words differently, $t(19) = -1.26$, $p = .222$. One-sample t-tests against chance revealed that infants' looking increased significantly toward the target object for both the unaccented words, $t(19) = 2.28$, $p = .035$, and the accented words, $t(19) = 4.94$, $p < .001$. Thus, for the other-race speaker, infants accepted both the accented and unaccented pronunciations.

Because previous studies (White & Aslin, 2011) have found that toddlers may learn about a speaker's accent during the test phase, planned analyses were conducted with test block as a factor. The first block is more representative of infants' initial interpretation of the words, whereas the second block indicates what they learned after some exposure to the speaker.

To determine if infants' responses changed over time, a mixed measures ANOVA with the within-subject factors block and word type, and the between-subject factor condition was run. It revealed a main effect of block, $F(1, 38) = 9.22$, $p = .004$, a main effect of word type, $F(1, 38) = 26.23$, $p = .01$, and, crucially, the significant condition X word type interaction $F(1, 38) = 17.21$, $p < .001$. No other effects were significant, ps > .083. The lack of a 3-way block X word type X condition interaction indicates that the infants' differential treatment of the pronunciations in the two speaker conditions was present in both blocks of testing. Consistent with this, the critical word type X condition interaction found in the overall analysis was found for each block separately: for block 1, there was a significant condition X word type interaction, $F(1, 38) = 8.09$, $p = .007$, but no main effect of condition or word type, ps > .071. For the second block, there was a

significant condition X word type interaction, $F(1, 38) = 14.42$, $p = .001$, and a main effect of condition, $F(1, 38) = 4.37$, $p = .043$, but no effect of word type, $p = .520$. Thus, for both blocks individually, infants interpreted the same words differently depending on which speaker they saw (see Figure 3).

Each speaker condition was then considered separately. A repeated measures ANOVA with the within-subject factors block and word type found that for the same-race speaker, there was a main effect of word type, $F(1, 19) = 21.83$, $p < .001$, no effect of block, $F(1, 19) = 3.43$, $p = .08$, and no block X word type interaction, $F(1, 19) = 1.17$, $p = .293$. In both blocks of trials, infants in the same-race speaker condition interpreted the unaccented and accented words differently, recognizing only the unaccented versions (block 1: $t(19) = 3.46$, $p = .003$; block 2: $t(19) = 3.11$, $p = .006$). Their looking significantly increased to the target object for unaccented words (block 1: $t(19) = 3.33$, $p = .004$; block 2: $t(19) = 3.69$, $p = .002$), but did not for the accented pronunciations (in block 1, there was a significant increase in looking to the *distractor* object, $t(19) = -2.69$, $p = .014$; in block 2, there was no change from baseline, $t(19) = .08$, $p = .936$). In other words, infants in this condition recognized only the unaccented pronunciations.

For the other-race speaker, a repeated measures ANOVA with the within-subject factors block and word type revealed a main effect of block, $F(1, 19) = 5.80$, $p = .026$, but no main effect of word type, $F(1, 19) = 2.35$, $p = .142$, and no block X word type interaction, $F(1, 19) = 0.56$, $p = .464$. Analyses by block revealed that in block 1, infants showed no difference between the unaccented and accented words, $t(19) = -.67$, $p = .510$, failing to recognize either type of word (unaccented words, $t(19) = .66$, $p = .519$; accented words, $t(19) = 1.73$, $p = .099$). By block 2, infants did treat the words

differently, $t(19) = -2.25$, $p = .036$. Infants' looking increased significantly toward the target object for both the unaccented words, $t(19) = 2.40$, $p = .027$ and accented words, $t(19) = 6.62$, $p < 0.001$.

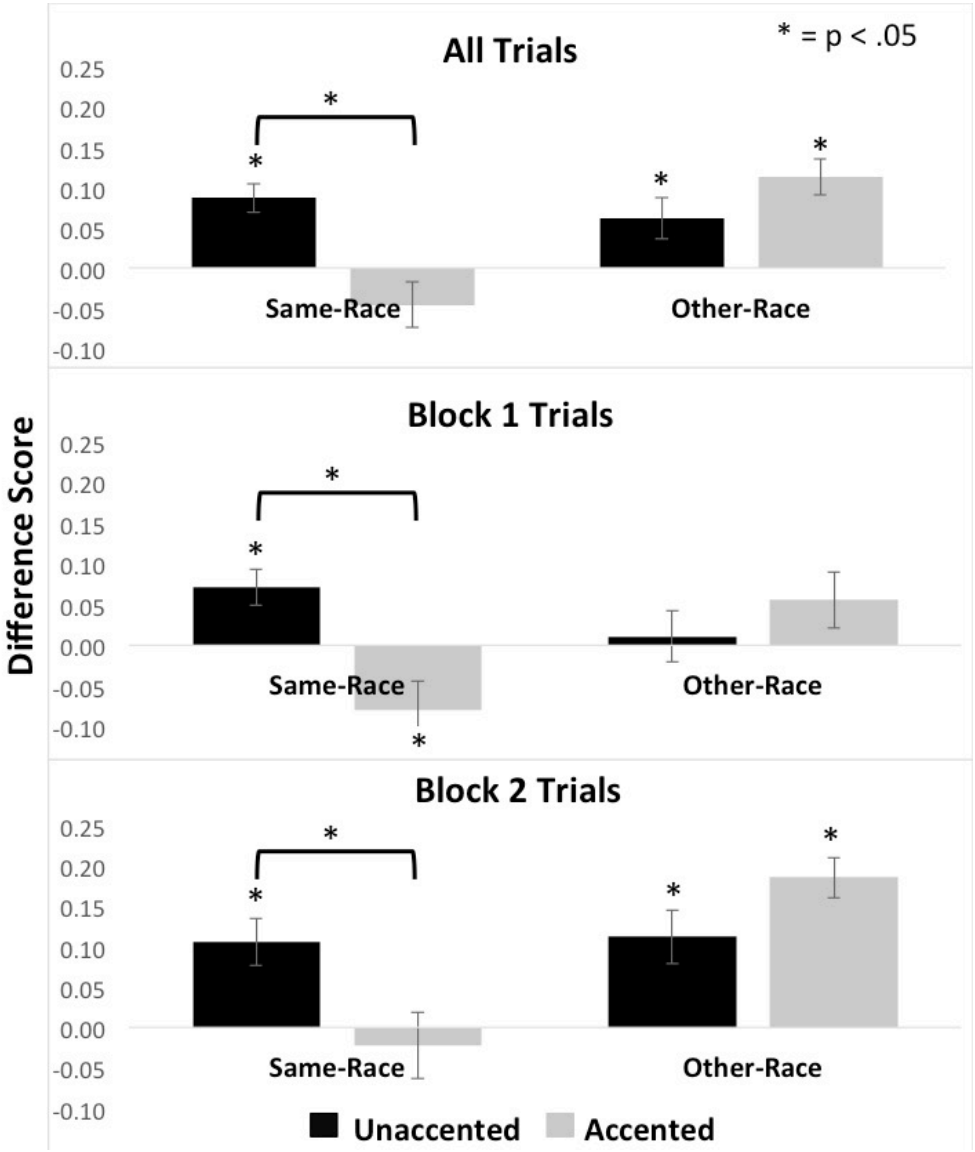

Figure 3. Difference scores and standard errors for Experiment 1. Speaker identity is on the X-axis. The Y-axis gives the difference between proportion looking at the target object in the naming phase and proportion looking at the target object in the baseline phase. A positive difference score indicates increased looking to the target

object, and a negative difference score indicates increased looking to the distractor object. The black bars correspond to the unaccented pronunciations and the grey bars correspond to the accented pronunciations.

2.1.6 Discussion

These findings demonstrate that infants interpreted the same words differently depending on whether the speaker was a same-race or other-race speaker. Overall, infants in the same-race condition accepted only familiar pronunciations, while infants in the other-race condition accepted both types of pronunciations. Moreover, the block analyses show the time course of this effect: when infants first encountered the other-race speaker, they were unsure how to interpret her speech (and failed to recognize both types of words). However, as they gained more experience with the speaker, they accepted her pronunciations, regardless of whether they were familiar or not.

2.2 Experiment 2

Infants clearly interpreted the speech of the same-race and other-race speakers differently in Experiment 1, suggesting that social information is indexed to infants' linguistic representations. Infants appear to have linked familiar pronunciations of words to members of their own social group, as they accepted the same-race speaker's use of familiar pronunciations and initially rejected unfamiliar pronunciations as labels for familiar objects. In contrast, for the other-race speaker, infants required experience before accepting either type of pronunciation, and eventually learned both.

One possibility for the latter finding is that infants were willing to accept any type

of unfamiliar pronunciations from an unfamiliar-looking speaker. An alternative possibility is that infants were attending to the speaker's pronunciations and learning the systematic difference between her accented pronunciations and the native pronunciations they are familiar with. Experiment 2 was designed to determine whether infants were indeed learning a systematic accent from the other-race speaker.

2.2.1 Participants

Twenty 16-month-old infants were tested (10 females; mean age: 15 months 27 days; age range: 15;10-16;15 days). Four additional participants were tested, but not included due to non-completion (3), and failure to attend to both objects during the baseline period for at least half of each trial type (1).

2.2.2 Stimuli

Audio Stimuli.

The same six highly familiar words from Experiment 1 were used. Recall, all six words contain the same vowel (/a/). Unlike the prior experiment, here there was no systematicity to the new pronunciations; a random vowel change was assigned to each word. For example, "bottle" was produced as "boottle", "sock" as "seck", "block" as "blick", etc. A female native speaker of English produced two versions of each word, one unaccented and one with a random mispronunciation. Once again, each version was produced naturally in infant-directed speech in the context of two sentences, "Do you see the X" or "Find the X".

Visual Stimuli.

Visual stimuli were identical to the other-race speaker condition of Experiment 1.

2.2.3 Procedure

This experiment follows the same general procedure as Section 2.1.3; however, only an other-race speaker was used.

2.2.4 Coding of looking times

See Section 2.1.4.

2.2.5 Results

As in Experiment 1, trials in which infants did not look at both objects during the baseline period (or at either object during the test period) were not included in the analyses. There was no difference in the percentage of discarded trials across Experiments 1 and 2, $F(2, 57) = 0.07$, $p = .937$. In addition, the proportion of time that infants spent looking at the familiar object during baseline in Experiment 2 (.54) was equivalent to the proportions found in Experiment 1, $F(2, 57) = 2.47$, $p = .094$.

To explore infants' recognition of the words, a difference score was again calculated for each trial (proportion looking target $object_{test}$-proportion looking target $object_{baseline}$). A paired sample t-test revealed a significant difference in how infants interpreted the unaccented and random pronunciations, $t(19) = 2.23$, $p = .038$. Infants' looking increased significantly toward the target object for the unaccented pronunciations, $t(19) = 3.85$, $p = .001$, but looking was at chance levels for the random pronunciations, $t(19) = 0.72$, $p = .478$.

As in Experiment 1, test block was included as a factor separately to explore changes over time. A repeated measures ANOVA with within-subjects factors block and word type revealed a main effect of word type, $F(1, 19) = 7.81$, $p = .012$, and a significant block X word type interaction, $F(1, 19) = 6.83$, $p = .017$. No main effect of block was found, $F(1, 19) = 0.02$, $p = .905$.

For the first block, a paired sample t-test revealed no difference between the unaccented and random pronunciations, $t(19) = 0.06$, $p = .954$. Infants did not show a significant change in looking to the target for either the unaccented pronunciations, $t(19) = 1.53$, $p = .143$, or the random pronunciations $t(19) = 1.77$, $p = .093$. In the second block, a paired sample t-test revealed a significant difference in how infants interpreted the unaccented and random pronunciations, $t(19) = 3.44$, $p = .003$. Infants' looking increased significantly toward the target object for the unaccented pronunciations, $t(19) = 3.56$, $p = .002$, but was at chance levels for the random pronunciations, $t(19) = -.93$, $p = .364$ (see Figure 4).

To determine whether infants' behaviour in this experiment was different from the behaviour of infants in the other-race speaker condition of Experiment 1, a mixed-measures ANOVA was conducted, with the within-subject factors block and word type, and the between-subject factor experiment (only the other-race condition was included for Experiment 1). This ANOVA revealed a significant word type X experiment interaction, $F(1, 38) = 8.44$, $p = .006$, and a significant block X word type X condition interaction, $F(1, 38) = 5.91$, $p = .02$. No other effects were significant, $ps > .066$. When each block was considered separately, for block 1, there were no statistical differences between experiments, $ps > .555$. However, for block 2, a main effect of experiment was

observed, $F(1, 38) = 5.33$, $p = .027$, and a significant experiment X word type interaction, $F(1, 38) = 16.84$, $p < .001$.

Therefore, as in the other-race speaker condition of Experiment 1, infants in Experiment 2 did not accept either type of pronunciation in the first block of testing. However, in contrast to Experiment 1, infants in Experiment 2 who heard random, rather than systematic, mispronunciations did not accept the pronunciations over time.

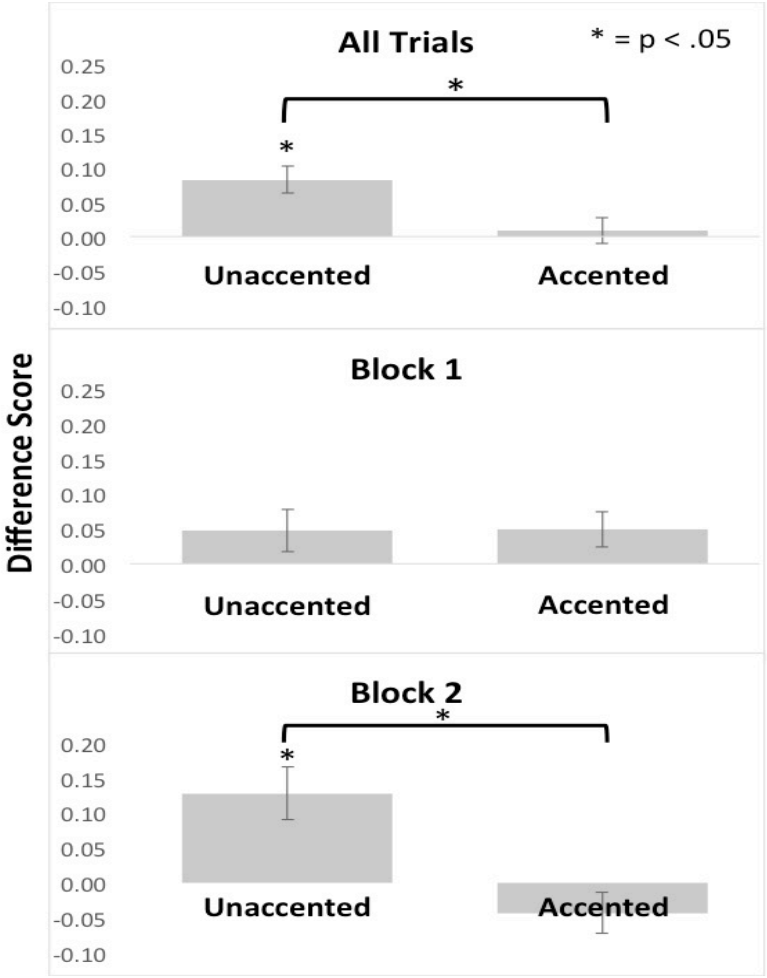

Figure 4. Difference scores and standard errors for Experiment 2. Test pronunciation is on the X-axis. The Y-axis gives the difference between proportion looking at the target object in the naming phase and proportion looking at the target object in the baseline phase.

2.2.6 Discussion

Infants in Experiment 2 did not recognize either type of pronunciation in the first block of testing. However, in contrast to Experiment 1, infants in Experiment 2 who heard random, rather than systematic, pronunciations did not accept the pronunciations over time. This finding demonstrates that for an other-race speaker, infants were not willing to accept any similar-sounding variant of a word, but rather, only pronunciations that systematically differed. The difference in infants' looking for the other-race speaker in Experiment 1 vs. Experiment 2 suggests that infants do not simply link any type of pronunciation to speakers; instead the accent must be systematically different. If infants were willing to accept any novel-sounding speech simply because the speaker looked novel, the same pattern of results would have been found across experiments.

2.3 Experiment 3

Experiment 3 was run to ensure that the findings held when a different other race speaker was used. Thus, Experiment 3 replicated the other-race speaker conditions of Experiment 1 and Experiment 2, in a between-subjects design using a different other-race speaker whose race was even more visually salient. This replication is important for two reasons: 1) it validates the generalizability of the previous two experiments, and 2) it allows for direct statistical comparison of the results for the other race speaker when she has a systematic vs. random accent.

2.3.1 Participants

Forty 16-month-old infants were tested (20 females; mean age: 16 months 1 days; age range: 15;18-16;17 days). Two additional participants were tested, but not included due to non-completion (1), and failure to attend to both objects during the baseline period for at least half of each trial type in each block of trials (1). As in Experiments 1 and 2, participants had minimal exposure to speakers who spoke a foreign language, had an accent, or were of a different race, and the amount of exposure was similar across conditions (Systematic Accent Condition: 3.7%, 6.9%, and 5.7%, respectively; Random Pronunciations Condition: 4.4%, 8.3%, and 9.1%, respectively).

2.3.2 Stimuli

Audio Stimuli. The same six highly familiar words from the previous two studies were used. Half the infants heard a systematic vowel shift (identical to that of Experiment 1), while the other half heard random pronunciations (identical to that of Experiment 2). A female native speaker of English produced all test stimuli. Once again, each version was produced naturally in infant-directed speech in the context of two sentences, "Do you see the X" or "Find the X".

Visual Stimuli. The test trials remained identical to the previous two experiments with the substitution of a new other-race speaker (see Figure 1).

2.3.3 Procedure

See sections 2.1.3 and 2.2.3.

2.3.3 Coding of looking times

See sections 2.1.4 and 2.2.4.

2.3.4 Results

As in previous experiments, trials in which infants did not look at both objects during the baseline period (or at either object during the test period) were not included in the analyses. A mixed measures ANOVA with the within-subject factor of word type (Unaccented vs. Accented) and a between-subject factor of condition (Systematic Accent vs. Random Pronunciations) found a main effect of word type, $F(1, 30) = 4.24$, $p = .046$, and a significant word type by condition interaction, $F(1, 30) = 7.46$, $p = .010$. No main effect of condition was found, $p = .187$.

To determine the effect of accent on infants' word processing across the whole experiment, paired sample t-tests compared unaccented pronunciations to the unfamiliar pronunciations (Figure 5). In the Systematic Accent condition, there was no difference across word types, $t(15) = 0.40$, $p = .698$; however, in the Random Pronunciations condition there was a significant difference between the unaccented and randomly pronounced words, $t(15) = 4.80$, $p < 0.001$. One-sample t-tests against chance (zero change) showed that in the Systematic Accent condition, infants' looking increased significantly to the familiar object for both word types (unaccented: $t(15) = 2.29$, $p = .034$; accented: $t(15) = 2.90$, $p = .011$). In the Random Pronunciations condition, infants increased their looking to the familiar object for the unaccented words, $t(15) = 5.34$, $p < .001$, but not the randomly pronounced words, $t(15) = 0.83$, $p = .420$.

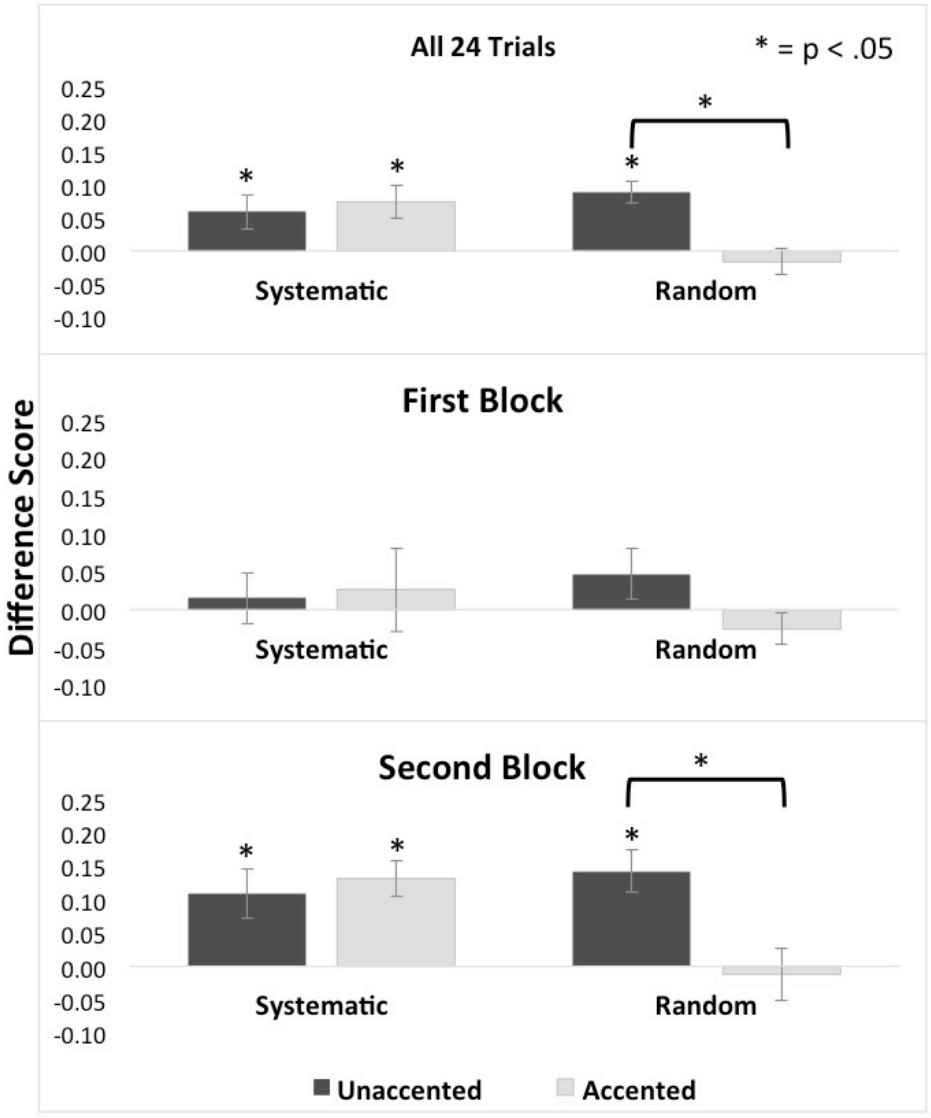

Figure 5. Difference scores and standard errors for Experiment 3. Test pronunciation is on the X-axis. The Y-axis gives the difference between proportion looking at the target object in the naming phase and proportion looking at the target object in the baseline phase.

As in the previous experiments, test block was included as a factor to explore changes over time for each condition. For the Systematic Accent Condition, a repeated measures ANOVA with within-subjects factors block and word type revealed a main

effect of block, $F(1, 15) = 5.91$, $p = .022$. No other effects were significant, ps > .666. For the first block, a paired sample t-test revealed no difference between the unaccented and accented mispronunciations, $t(15) = 0.15$, $p = .886$. Infants did not show a significant change in looking to the target for either the unaccented pronunciations, $t(15) = 0.43$, $p = .674$, or the accented pronunciations, $t(15) = 0.45$, $p = .657$. In the second block, a paired sample t-test revealed no difference between the unaccented and accented mispronunciations, $t(15) = 0.61$, $p = .551$. Infants' looking increased significantly toward the target object for the unaccented pronunciations, $t(19) = 2.98$, $p = .009$, and the accented pronunciations, $t(15) = 4.80$, $p < .001$ (see Figure 5).

For the Random Accent Condition, a repeated measures ANOVA with within-subjects factors block and word type revealed a main effect of word type, $F(1, 15) = 22.93$, $p < .001$. No other effects were significant, ps > .120. For the first block, a paired sample t-test revealed a marginal difference between the unaccented and random mispronunciations, $t(15) = 1.96$, $p = .069$. Infants did not show a significant change in looking to the target for either the unaccented pronunciations, $t(15) = 1.37$, $p = .191$, or the random pronunciations, $t(19) = 1.24$, $p = .233$. In the second block, a paired sample t-test revealed a significant difference in how infants interpreted the unaccented and random pronunciations, $t(15) = 2.75$, $p = .015$. Infants' looking increased significantly toward the target object for the unaccented pronunciations, $t(15) = 34.53$, $p < .001$, but was at chance levels for the random pronunciations, $t(15) = 0.32$, $p = .3756$ (see Figure 5).

2.3.4 Discussion

In both conditions, infants did not accept either type of pronunciation in the first block of testing. However, infants who heard systematic, rather than random, pronunciations did accept the unfamiliar pronunciations over time. These results directly replicate the results with the other-race speaker in the previous two experiments, using a different other-race speaker.

2.4 General Discussion

In three experiments, infants' perception of same-race and other-race speakers' word pronunciations were explored, as well as what they learned about those speakers' pronunciations over time. Overall, infants interpreted both unaccented and accented words differently depending on the speaker's race. In Experiment 1, infants who viewed a same-race speaker accepted only unaccented versions of familiar words, whereas infants who viewed an other-race speaker accepted both unaccented and accented versions of the words. Experiment 2 further demonstrated that for an other-race speaker, infants did not simply accept any similar-sounding variant of a word, but rather, only recognized words produced with a systematic accent. Experiment 3 replicates the findings with the other-race speaker observed in Experiment 1 and 2. These results provide the first evidence that social properties of speakers, such as race, influence infants' speech processing.

The finding that infants in the same-race condition accepted only unaccented pronunciations of words suggests that infants have some form of representation in which social information (in this case Caucasian race) is associated with linguistic information (in this case their native accent). For the other-race speaker, in contrast, infants have had

almost no experience to draw on, and thus have no pre-existing link between this race and specific linguistic information. Infants were initially unsure about how to interpret words from other-race speakers. Infants did not accept the unaccented or accented words in block 1 in any experiment. The fact that they did not even accept the unaccented pronunciations is particularly interesting, given that infants reliably map such words to target objects when there is no information about the speaker's appearance. After some evidence that the speaker talked in a consistent manner, they eventually recognized both unaccented and accented pronunciations. Thus, *some* experience with a social group is necessary before sociolinguistic representations can be formed (Foulkes & Docherty, 2006); however, these results suggest that far less experience is necessary than previously hypothesized.

This is the first study to demonstrate that race information affects infants' speech processing. Additionally, it provides critical insights into how social information might be initially linked to linguistic representations. Infants in this study had very little exposure to other races and accents and, therefore, had mostly heard these familiar words produced in a particular way by same-race speakers. As a result, they appear to have linked those pronunciations with same-race speakers. In contrast, encountering an other-race speaker appeared to trigger a different process. With little experience to draw on, infants initially failed to recognize either unaccented or accented pronunciations, suggesting that they were waiting for information about the speaker's pronunciations.

One important thing to note is the difference in infants' looking for the other-race speaker in Experiment 1 vs. Experiment 2, and in Experiment 3. Infants were not simply willing to accept any novel-sounding speech because the speaker looked novel. Had this

been the case, the same pattern of results would have been found across experiments. This particular finding contradicts some proposals of early accent accommodation, in which infants are thought to generally relax their word processing to accept a degree of deviance after exposure to variation, whether linguistic or social (see evidence from Schmale, Cristia, & Seidl, 2012; Schmale, Seidl, & Cristia, 2015). Infants in the current study did not simply relax their categories for the other-race speaker, as demonstrated by the recognition failure for the other-race speaker in block 1. Furthermore, in the second block of trials, infants only recognized deviant pronunciations if they were in the form of a systematic accent.

In conclusion, Chapter 1 demonstrates that social information is indexed to 16-month-old infants' linguistic representations. When encountering a speaker from a familiar group, they accept familiar pronunciations from this speaker and initially reject unfamiliar pronunciations (though after some time they may begin to accept these unfamiliar pronunciations). However, when encountering a speaker from a novel group, they must have some experience with the speaker before deciding how to interpret their pronunciations.

3. Specificity of toddlers' sociolinguistic categories

Chapter 2 demonstrates that infants do not automatically generalize familiar pronunciations to members of a novel social group. Furthermore, infants accepted an alternative label for a referent when the speaker was an other-race speaker, but not if they were a same-race speaker. In this situation, infants had no prior experience with the novel social group. Chapter 3 will go one step further, and ask whether this social information is used to interpret utterances from future speakers of the same kind. Toddlers will be familiarized with the linguistic properties of an other-race individual, and later be tested on their interpretation of productions from new speakers of that race or another race. Adults have very specific associations between race and accent (i.e., American Accented English goes with Caucasian faces and Mandarin Accented English goes with Chinese faces), but they have also had extensive experience. In this study, prior to exposure, toddlers will not have had any experience with the specific linguistic variation (i.e., the artificial accent), and no (or very little) experience with the social group of the Novel Speaker (i.e., South Asian and East Asian speakers). If toddlers' linguistic representations are linked to social information, as they are first being formed, then toddlers should generalize the linguistic properties from one other-race speaker to a speaker of the same race.

Additionally, this chapter will examine the specificity with which social information is indexed to linguistic properties. Do toddlers simply have a "like me" vs. "other" organization of speakers? If this were the case, if toddlers learn that an ethnically Indian individual talks with a specific accent, they should generalize this accent to a new speaker who is ethnically Korean. The other alternative is that toddlers have a more

nuanced organization of social information, in which associations between race and accent are more specific. In this case, an Indian speaker and Korean Speaker would belong to different social categories, and therefore experience with an Indian speaker should have no influence on toddlers' perception of a Korean speaker.

3.1 Experiment 4

This experiment is a replication of Weatherhead & White (2016) using 24-to-26-month-old participants. In the original study, 11-month-old infants were able to track two Caucasian speakers' accents simultaneously, and use this information to guide their future interactions with each speaker. Additionally, infants learned something specific about the differences between their accents (i.e., that one speaker had higher front vowels than the other). When tested on words that did not have the systematically shifted vowel, infants recognized that the speakers would produce words in the same way. A replication of this study with 2-year-olds was necessary to first establish the extent to which 2-year-olds succeed in this task, as it is the basis for the key manipulations throughout Chapters 3 and 4.

3.1.1 Participants

Forty 24-to-26-month-olds were tested (17 females and 23 males; mean age: 748 days; age range: 724-795 days). Four additional participants were tested, but not included due to lack of attention during test trials (2), failure to attend to both objects during the baseline period of test trials (1), or difference scores exceeding 2.5 standard deviations from the mean of either speaker (1).

3.1.2 Stimuli

Audio Stimuli

The stimuli consisted of four pairs of CVCV nonsense words (see Table 2) as used in Weatherhead & White (2016), produced by two female native speakers of English. The pronunciations of the words differed only in the first vowel (a front vowel), while the remainder of the word was consistent across speakers. Three of the word pairs (*m[I/i]to*, *d[E/I]lu*, and *b[I/i]mo*) were shown during exposure without referents (exposure pairs). The Training Speaker used the word "tEpu" during exposure to label an object (object presentation event). The last word, "tIpu" was heard only at test. Stimuli were recorded in a sound-treated booth at a sampling rate of 44100 Hz and equated for amplitude in Praat (Boersma & Weenink, 2009). The audio stimuli for the exposure phase were inserted into the videos described below.

Word Type	Trained Speaker	Extension Speaker
Exposure Pair 1	mIto	mito
Exposure Pair 2	dɛlu	dIlu
Exposure Pair 3	bImo	bimo
Object Presentation Pair	tɛpu	
Test Word	tIpu	

Table 2. Audio stimuli used during exposure and test in Experiments 4-7.

Audiovisual Stimuli (Exposure Phase)

Both talkers, 24-year-old Caucasian females, were recorded against the same backdrop and wore different colored t-shirts (white and black). Both talkers recorded three exposure videos, in which a single exposure word was repeated three times in toddler-directed speech with approximately one second between each utterance. Each

talker also recorded an object presentation event. In the Training Speaker's object presentation event, she held and waved the target object while labelling it "tEpu" three times (this object is hereafter referred to as the trained object). In contrast the Extension Speaker was only seen holding and waving the trained object, providing no label. Toddlers were either trained with an unfamiliar blue object or an unfamiliar yellow object (counterbalanced).

3.1.3 Procedure

The participant sat on his/her parent's lap approximately 1.5 ft. from a 36x21-inch plasma screen television in a sound-treated testing room. A camera under the television recorded the child's looking behaviour during the entire session. The camera linked to a monitor and recording device in the lab area adjacent to the testing room for the experimenter's viewing purposes and later off-line coding. Stimuli were presented in Psyscope X (Cohen, MacWhinney, Flatt & Provost, 1993) at approximately 65dB. Parents wore noise-cancelling headphones playing instrumental music.

The exposure phase began with the object presentation events from both talkers, to indicate to the toddlers that they were in a word-learning situation. Next, the three pairs of yoked exposure videos (e.g., *mIto-mito*) were presented in random order (see Table 2). These pairs highlighted the front-vowel difference between the talkers. Finally, the object presentation event pair was presented again twice (see Weatherhead & White, 2016). In total, toddlers saw the Training Speaker label the trained object 9 times. An attention getter occurred between the video pairs, with the next pair beginning when the experimenter judged that the participant was focused on the screen.

The test phase began immediately after the exposure. There were two test trials, one per talker. Each trial was 10 seconds in length. At the start of each trial, the talker's face appeared alone for 2 seconds, followed by a display with the trained object and a novel untrained object. The objects remained on the screen for 8 additional seconds, the first 3 seconds of which was a silent baseline period, followed by an audio recording of the pictured talker saying the test word ("tIpu"). The talker in the first test trial and the side on which the trained object appeared were counterbalanced across participants (this side assignment remained constant for both test trials).

If toddlers were able to learn the Training Speaker's label for the trained object, "tEpu", during the exposure phase, then the novel label "tIpu" should be mapped to the untrained object for this talker. This would be in line with previous work demonstrating that toddlers typically show a disambiguation response, mapping novel labels to novel objects (e.g., Clark, 1990; 1992; 1997; 2007; Golinkoff, Mervis & Hirsh-Pasek, 1994; Halberda, 2003; Markman, 1989; 1990; Merriman & Bowman, 1989). If they also learned that the Extension Speaker has higher front vowels than the Training Speaker, then they should interpret "tIpu" as the Extension Speaker's pronunciation of the trained object's label. Thus, toddlers would look longer to the trained object for this talker.

3.1.4 Coding of looking times

Looking time during the test phase was coded off-line using customized software (Brown University), frame-by-frame (1 frame = 33 msec). Looking proportions for the objects were determined for the baseline period and for the test period, which began 300 msec after test word onset to account for the time necessary to program an eye movement

in response to the first vowel in "tIpu" (e.g., Swingley & Aslin, 2002). Only the first 3 seconds of the test period were analyzed, to equate the length of the baseline and test periods (see Figure 6).

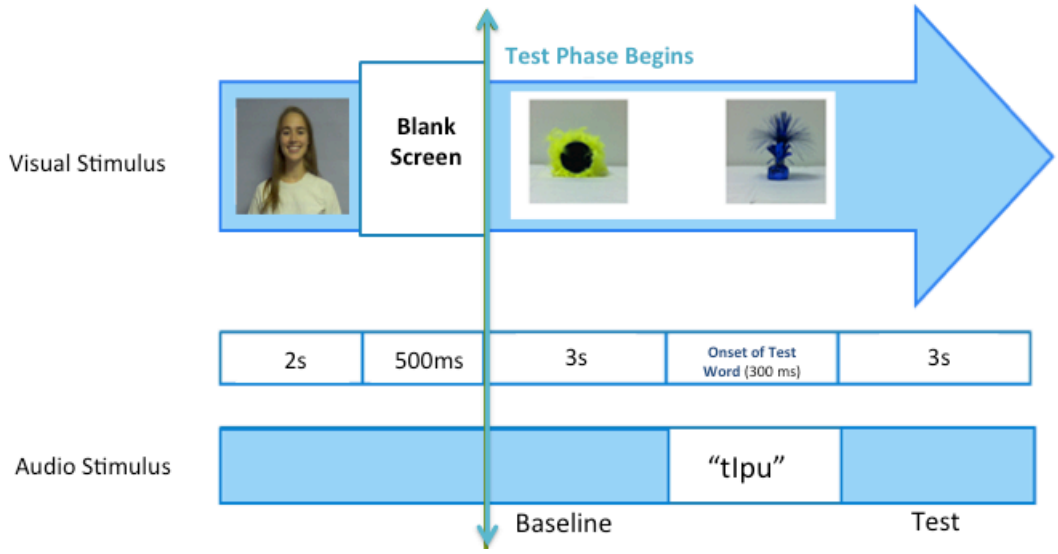

Figure 6. An example of a test trial in Experiment 4. This example depicts a trial for the Training Speaker.

3.1.5 Results

For both the baseline and test periods, the proportion of time toddlers looked at each object was computed (out of the total time looking at either object during the 3-second period). To assess toddlers' interpretation of the word "tIpu", a difference score was calculated for each trial using the looking proportions for each period (proportion target object$_{test}$-proportion target object$_{baseline}$).

A repeated measures ANOVA on these difference scores with the within-subjects factor of Speaker and between-subjects factor of test Order revealed a main effect of

Speaker, $F(1,38) = 0.566$, $p < .001$, and no main effect of order $F(1, 38) = 0.43$, $p = .515$, and no Speaker * Order interaction, $F(1,38) = 0.225$, $p = .638$ (Figure 7).

To determine the effect of labelling for each talker separately, one-sample t-tests compared difference scores for each speaker against chance (i.e., a difference score of 0). As predicted, when the Training Speaker said "tIpu", looking significantly decreased to the trained object, $t(39) = -2.16$, $p = .037$. In contrast, for the Extension Speaker, looking significantly increased to the trained object, $t(39) = 3.06$, $p = .004$. Thus, just as the 11-month-olds in Weatherhead & White (2016), toddlers increased their looking toward the untrained object when the Training Speaker said "tIpu", but increased their looking toward the trained object when the Extension Speaker said "tIpu".

This pattern of results demonstrates that toddlers learned the Training Speaker's label for the training object ("tEpu"), and by a process of disambiguation inferred the test label mapped onto the untrained object. Furthermore, toddlers tracked the talker-specific linguistic differences and used this information to predict the Extension Speaker's label for the trained object.

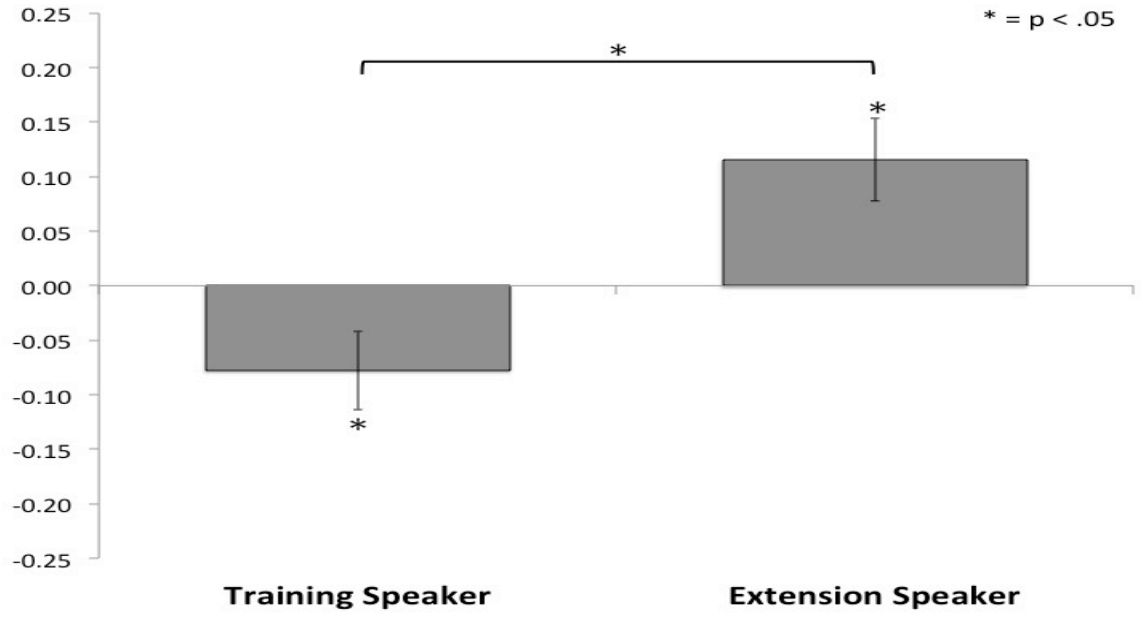

Figure 7. Difference scores and standard errors for each speaker in Experiment 4. Positive scores reflect an increase in looking to the trained object while negative scores reflect a decrease in looking to the trained object.

3.1.5 Discussion

This Experiment demonstrates that toddlers are able to track the subtle accent differences between two speakers. Using this talker-specific information, they were then able to predict how the Extension Speaker would pronounce a word they had not previously heard her say. This demonstrates that (1) toddlers recognized that the label for the trained object would be different as a function of speaker, and (2) toddlers used previous linguistic information to interpret future utterances by those speakers. This shows sophistication beyond that of the 11-month-old infants in Weatherhead & White (2016); 11-month-olds were only successful in the easier test order in which the Training

Speaker appeared first (which appeared to allow them to use their interpretation of the Training Speaker's label as a reference for interpreting the Extension Speaker's).

The ability to track talker-specific linguistic information is essential to the formation of group-level linguistic representations. That is, through tracking individual speakers' pronunciations, and their socioindexical information, a sociolinguistic category will emerge.

3.2 Experiment 5

Experiment 4 demonstrated that 2-year-olds are able to track the subtle accent differences between two speakers. Using this talker-specific information, they were then able to predict how the Extension Speaker would pronounce a word they had not previously heard her say. Thus, toddlers interpreted the test word as a function of who produced the word.

Experiment 5 investigates whether toddlers generalize this talker-specific accent information to members of the same race. To accomplish this, in Experiment 5, the Training Speaker and the Extension Speaker were different races (Caucasian and South Asian respectively). A third speaker was introduced immediately preceding the test trials (hereafter referred to as the Novel Speaker), who was the same race as the Extension Speaker. If toddlers generalize talker-specific accent information to individuals of the same race, then toddlers' interpretation of the Novel Speaker's "tIpu" should be the same as that of the Extension Speaker and different than that of the Training Speaker. This would demonstrate that toddlers' representations of the two speakers' utterances incorporated social group information, in this case race. Furthermore, it would suggest

that very little experience with the social group is needed before this information affects the interpretation of new speakers' utterances.

3.2.1 Participants

Twenty 24-to-26-month-olds were tested (9 females; mean age: 762; age range: 739-796 days). One additional participant was tested, but not included due to lack of attention during test.

3.2.2 Stimuli

Audio Stimuli

The audio stimuli were the same as in Experiment 4. An additional female speaker produced a 3rd test token of "tIpu".

Audiovisual Stimuli (Exposure Phase)

The audiovisual stimuli in Experiment 5 were almost identical to that of Experiment 4. In this experiment, the Training Speaker was Caucasian and the Extension Speaker was of another race (South Asian; see Figure 8). As in Experiment 4, the two talkers' productions systematically differed in the height of their front vowels: The Extension Speaker's front vowels were higher than the Training Speaker's. After some exposure to the differences in their accents, toddlers learned the label for a novel object from the Training Speaker ("tEpu"), but did not hear the Extension Speaker label it. Immediately preceding the test trials, a new speaker, the Novel Speaker, was introduced along with the other two speakers. The Novel Speaker simply waved to the child;

critically, she was never heard speaking nor did she interact with either the Training or Extension Speakers. The Novel Speaker was an other-race speaker that was the same race as the Extension Speaker (South Asian).

Figure 8. The Training Speaker (Left), Novel Speaker (Center), and Extension Speaker (Right) in Experiment 5.

3.2.3 Procedure

The procedure was identical to that of Experiment 4, with the addition of the Novel Speaker Introduction video described above in the Exposure Phase. At test there were three test trials, one for each talker (Novel Speaker, Training Speaker, and Extension Speaker). For all participants the Novel Speaker appeared first at test. The orders of the second and third trials were counterbalanced for the Training and Extension Speakers (i.e., half the participants saw the order Novel, Training, Extension, and the other half saw Novel, Extension, Training). The side on which the trained object appeared was counterbalanced across participants.

Based on Experiment 4, toddlers should map the novel label "tIpu" to the untrained object for the Training Speaker, and to the trained object for the Extension Speaker. If toddlers interpret a speaker's utterances based on social group membership, and use race as a cue to social group membership, they should interpret the Novel

Speaker's "tIpu" the same way as they do for the Extension Speaker (recall that the Novel and Extension Speakers are both South Asian, while the Training Speaker is Caucasian). Thus, for the Novel Speaker "tIpu" should refer to the trained object.

3.2.4 Coding of looking times

See Section 3.1.4.

3.2.5 Results

A repeated measures ANOVA on the participants' calculated difference scores with the within-subjects factor of Speaker and between-subjects factor of test Order revealed a main effect of Speaker, $F(2,18) = 10.92$, $p < .001$, and no main effect of order, $F(2, 18) = 0.03$, $p = .871$, and no Speaker * Order interaction, $F(2,18) = 0.04$, $p = .965$. Paired sample t-tests revealed no difference in looking behaviour for the Novel Speaker and the Extension Speaker, $t(19) = 0.95$, $p = .355$, but significant differences between the Novel Speaker and the Training Speaker, $t(19) = 4.05$, $p = .001$, and the Training Speaker and the Extension Speaker, $t(19) = 4.03$, $p = .001$ (Figure 9).

To determine the effect of labelling for each talker separately, one-sample t-tests compared difference scores for each speaker against chance (where chance = a difference score of 0). As predicted, looking significantly decreased to the trained object when the Training Speaker said "tIpu", $t(19) = -3.15$, $p = .005$. In contrast, looking significantly increased to the trained object for both the Novel Speaker, $t(19) = 2.76$, $p = .013$, and the Extension Speaker, $t(19) = 2.79$, $p = .012$. Thus, when the Training Speaker said "tIpu", toddlers increased their looking toward the untrained object, but when the Extension

Speaker and the Novel Speaker said “tIpu”, they increased their looking toward the trained object.

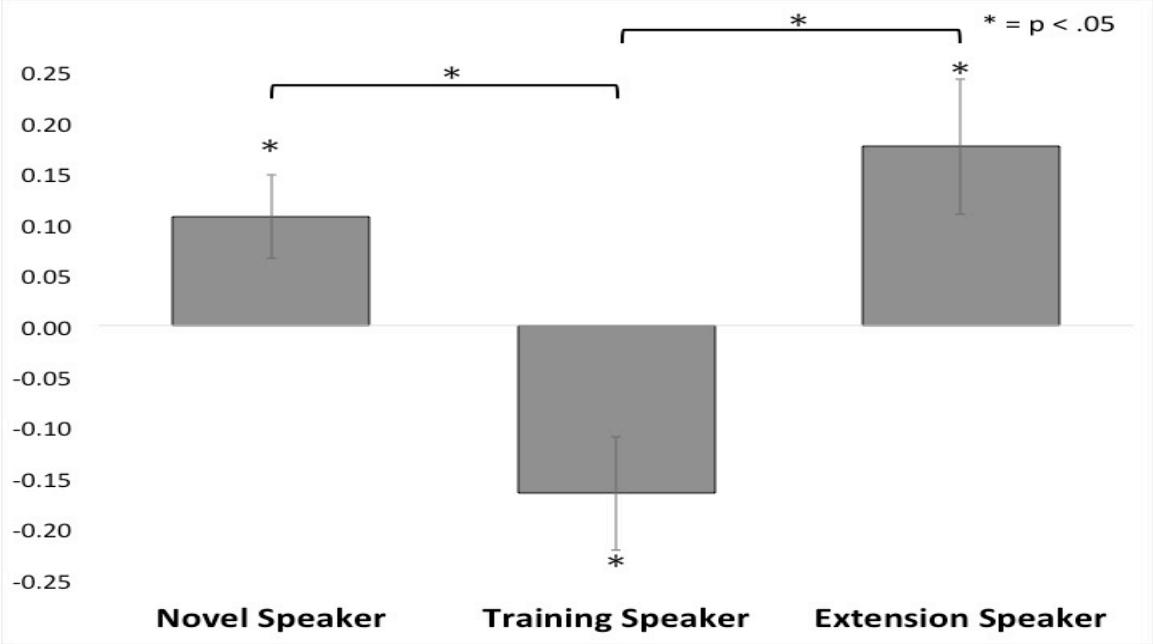

Figure 9. Difference scores and standard errors for each speaker in Experiment 5. Positive scores reflect an increase in looking to the trained object while negative scores reflect a decrease in looking to the trained object.

3.2.6 Discussion

Toddlers inferred that the Novel Speaker would produce words in the same way as the Extension Speaker, using their previous experience with an other-race speaker to guide their interaction with a speaker of that same race. Thus, even following short exposure to one speaker, toddlers will generalize their accent to new speakers of the same race. This suggests that social information about the speaker is being indexed to toddlers’ linguistic representations.

3.3 Experiment 6

Experiment 6 addresses how specific the social information is that is linked to toddlers' linguistic representations. The procedure was identical to Experiment 5 with one exception. Rather than being the same race as the Extension Speaker, the Novel Speaker was an other-race speaker who was a different race than the Extension Speaker (East Asian).

If toddlers simply think about race in terms of "like me"/"other" (i.e., Caucasian speakers talk like this and everyone else talks in some different way), they should interpret the Novel Speaker's pronunciations the same way as the Extension Speaker's pronunciations. If, however, the information they are tracking is more specific (i.e., they differentiate people of different races), they should not map either speaker's accent to the Novel Speaker, and thus perform at chance levels.

3.3.1 Participants

Nineteen 24-to-26-month-olds were tested (11 females and 8 males; mean age: 758; age range: 728-792 days). One additional participant was tested, but not included due to failure to complete the task.

3.3.2 Stimuli

Audio Stimuli

See Experiment 5.

Audiovisual Stimuli (Exposure Phase)

See Experiment 5. In this experiment, again the Training Speaker was Caucasian and the Extension Speaker was an other-race speaker (South Asian). Critically, the Novel Speaker was an other-race speaker who was a different race than the Extension Speaker (East Asian; Figure 10).

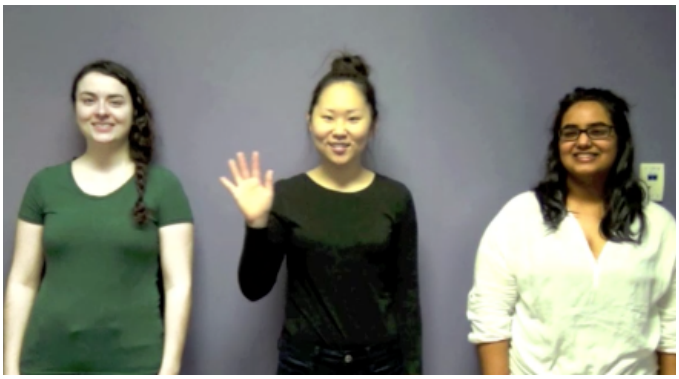

Figure 10. The Training Speaker (Left), Novel Speaker (Center), and Extension Speaker (Right) in Experiment 6.

3.3.3 Procedure

The procedure was identical to that of Experiment 5. Based on Experiments 4 and 5, toddlers should map the novel label "tIpu" onto the untrained object for the Training Speaker, and to the trained object for the Extension Speaker. If toddlers are simply tracking whether speakers are ingroup or outgroup members, then they should extend the Extension Speaker's accent to the Novel Speaker. If toddlers are tracking specific socio-indexical information about the speakers, they should not map either speaker's accent to the Novel Speaker, and thus perform at chance levels.

3.3.4 Coding of looking times

See Section 3.1.4.

3.3.5 Results

A repeated measures ANOVA on the participants' calculated difference scores with the within-subjects factor of Speaker and between-subjects factor of test Order revealed a main effect of Speaker, $F(2,17) = 9.16$, $p = .001$, and no main effect of order, $F(2, 17) = .03$, $p = .856$, and no Speaker * Order interaction, $F(2,17) = .53$, $p = .591$. Paired sample t-tests revealed a significant difference in looking behaviour between the Training Speaker and the Extension Speaker, $t(18) = 4.71$, $p < .001$, and between the Novel Speaker and the Training Speaker, $t(18) = 2.16$, $p = .044$. Additionally, there was a marginal difference between the Novel Speaker and the Extension Speaker, $t(18) = .193$, $p = .070$ (Figure 11).

To determine the effect of labelling for each talker separately, one-sample t-tests compared difference scores for each speaker against chance (where chance = a difference score of 0). As predicted, looking significantly decreased to the trained object when the Training Speaker said "tIpu" $t(18) = 3.475$, $p = .003$. Looking marginally increased to the trained object for the Extension Speaker, $t(18) = 1.83$, $p = .084$. However, for the Novel Speaker, toddlers' change in looking was at chance, $t(18) = .01$, $p = .994$. Thus, when the Novel Speaker said "tIpu" toddlers were unsure which object to direct their attention to.

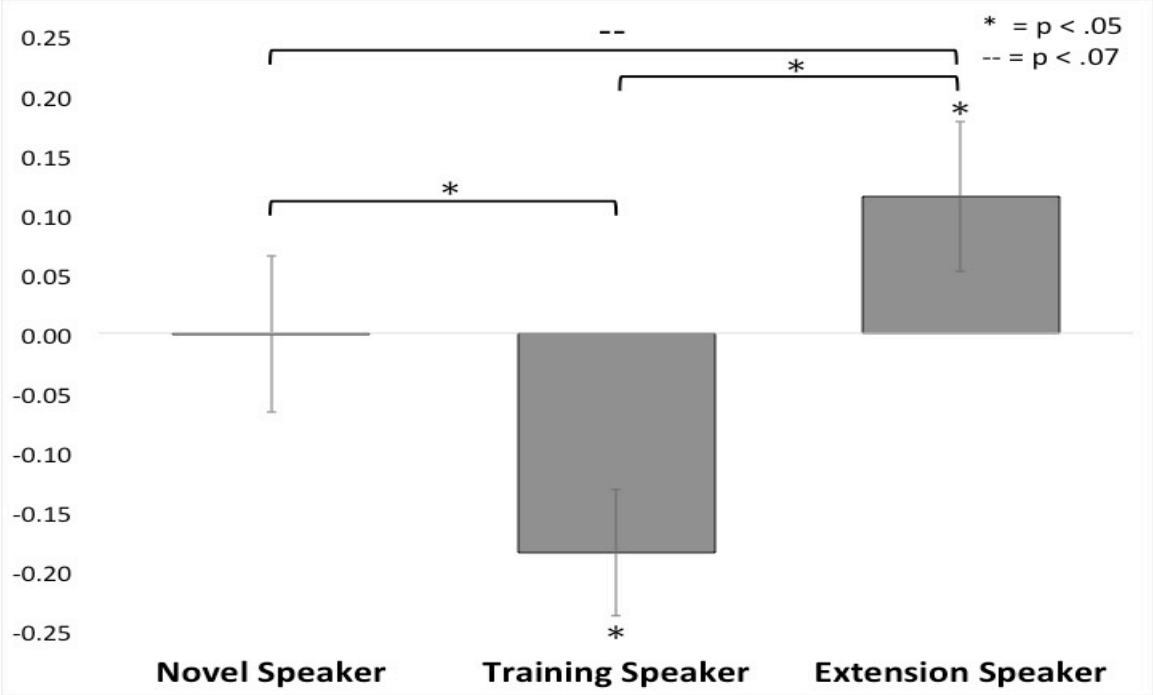

Figure 11. Difference scores and standard errors for each speaker in Experiment 6. Positive scores reflect an increase in looking to the trained object while negative scores reflect a decrease in looking to the trained object.

3.3.6 Discussion

Unlike the previous experiment, toddlers did not increase their looking to the Trained Object when the Novel Speaker said "tIpu". Additionally, there was a marginal difference in looking behaviour between the Novel Speaker and the Extension Speaker. These two findings together suggest that toddlers are not simply lumping all other race speakers into one large "other" category. Rather, it may be the case that toddlers track specific information about race, and this detail about race is indexed to their linguistic representations. This also suggests that in terms of linguistic variation, toddlers are doing

more than a simple "like me" vs. "other" comparison, and learning something about how specific types of people produce words.

However, the difference between the Novel Speaker and the Extension Speaker was marginal and not statistically significant. This appears to be largely driven by toddlers' performance during the Extension Speaker's trials, which was not as strong as in the previous experiment (Mean = .12(.28) compared to Mean = .16(.27)). This is likely due to the demands of the task. In a follow up study I will reduce the working memory load by having the Extension Speaker provide her label for the Trained Object during exposure ("tIpu"). In this easier version of the task, I predict that looking behaviour will significantly differ for the Novel Speaker and Extension Speaker.

3.4 General Discussion

In three experiments, I investigated the formation of sociolinguistic categories in toddlers. In Experiment 4, toddlers tracked the talker-specific features of two speakers, and used this information to guide their referential interpretations. Thus, toddlers are able to track linguistic variation at the individual-talker level. Experiment 5 demonstrates that social group membership is indexed to toddlers' linguistic representations, after little experience with a single group member. This information then influences their future interactions with members from that specific social group. In Experiment 6, toddlers did not generalize the Extension Speaker's pronunciations to a Novel Speaker of a different race. This suggests that toddlers are tracking not only specific linguistic information from individual talkers, but also specific physical information.

These experiments demonstrate that after learning talker-specific accent differences, toddlers generalize a single talker's accent to a member of the same social group (as defined by race). Toddlers had not heard the Novel Speaker talk prior to test, yet based on this social group information, extended talker-specific accent information to this new speaker. They were able to do this even though the Novel Speaker belonged to the same social group as the Extension Speaker. In this case, toddlers had not even heard the Extension Speaker's label for "tEpu"; thus, they not only predicted how the Extension Speaker would pronounce "tEpu", they also predicted that a member of her social group would pronounce it in the same way.

Importantly, toddlers only had very brief exposure to the Training and Extension Speakers, having only heard them pronounce 3 words (plus "tEpu" in the Training Speaker's case). Moreover, recall that prior to the exposure phase, toddlers had not had experience with the specific linguistic variation (i.e., the artificial accent), and no (or very little experience) with the social group of the Extension Speaker or the Novel Speaker (i.e., South Asian and East Asian speakers). Yet, even with very few examples of how each speaker talked, toddlers generalized talker specific accents to members of the same group. Thus, while Chapter 2 demonstrated that experience with a speaker is necessary for a link between social and linguistic information to be formed, Chapter 3 demonstrates that very little experience is needed before this information is applied to new speakers.

These experiments also addressed the specificity of toddlers' of the social information indexed to linguistic representations. As discussed earlier, toddlers could be tracking features of the speaker more generally (e.g., speaker not Caucasian), which could lead to an overly general category. Or they could be tracking specific features of

the speaker (e.g., speaker is racially Indian, tall, long hair, female, etc.), which may lead to an overly specific category. Experiment 5 demonstrates that toddlers were certainly tracking information about race. However, though both speakers were ethnically Indian, there were a number of other differences between the Novel Speaker and the Extension Speaker (such as height, body type, t-shirt colour, and the presence/absence of glasses). The fact that toddlers did generalize across these speakers suggests that the social information indexed to the word representations is not over-specified. However, in Experiment 6, toddlers did not interpret the Novel Speaker's pronunciations in the same way as the Extension Speaker's (or the Training Speaker's pronunciations for that matter). Thus, the social information indexed to word representations is not underspecified either, and is specific to the group level, at least in terms of race.

Overall, these results demonstrate that toddlers track linguistic variation at both the individual, and the social category level. Importantly, after learning talker-specific accent differences, toddlers generalize a single talker's accent to a member of the same social group. These results suggest that toddlers' linguistic representations incorporate social information, as they are first being formed.

4. The effect of affiliation on toddlers' speech processing

The previous two chapters demonstrate that infants and toddlers use race as an indicator of linguistic variation. The motivation for this experiment is to determine whether toddlers are sensitive to more abstract cues to linguistic group membership. While race is certainly a salient cue, it is not always an indicator of linguistic group membership. For example, Canadians and Russians might look very similar but there are many cultural differences between the two groups. Most importantly, these two social groups speak different languages. Another issue with using race as an indicator of linguistic group membership is that many Canadians who speak with native accents are not Caucasian. Thus, toddlers' assumptions that other-race speakers have different linguistic properties, or that all speakers of the same race speak in the same way, may not always be correct.

Experiment 7 was designed to determine whether toddlers are able to use more abstract cues to infer linguistic group membership, and whether this affects their processing of speech. Specifically, I investigate the role of affiliative behaviour on toddlers' word processing. Affiliation between two individuals should act as a strong indicator that they are members of the same social group. For example, infants readily affiliate with members of their own linguistic group over members of another (Kinzler et al., 2007). Recent work has even demonstrated that infants can reason about third-party affiliative relationships (i.e., whether two individuals, independent of the child, are affiliated with the same social group). For example, 9-month-olds expect that individuals who share the same food preferences will later have a positive interaction, while two individuals with opposing preferences will later have a negative interaction (Liberman et

al., 2014). Similarly, 9-month-olds are more surprised when two individuals who speak their native language have a negative interaction than when they have a positive interaction; but, when two individuals speak different languages (with one speaking their native language and the other speaking a foreign language), infants are more surprised when they later have a positive interaction than when they have a negative interaction (Liberman, Woodward, & Kinzler, 2016). These studies suggest that infants make judgements of affiliative relationships based on previous similar or dissimilar behaviour.

If more abstract social group information is indexed to toddlers' linguistic representations, then they should infer that two individuals who previously affiliated with each other would produce words in the same way.

4.1 Experiment 7

This experiment uses the same general methods from Chapter 3, with the exception that the third speaker (the Affiliated Speaker) did not share racial characteristics with either speaker, but instead chose to affiliate with, and wore the same t-shirt colour as, either the Training Speaker or the Extension Speaker (between subjects). At test, toddlers once again saw the trained object and an untrained object, and heard all three speakers use the label "tIpu". The Affiliated Speaker always appeared first at test. If toddlers generalize talker-specific accent information to members of the same social group, then toddlers' interpretation of the Affiliated Speaker's "tIpu" should differ as a function of which speaker she affiliated with.

4.1.1 Participants

Forty 24-to-26-month-olds were tested (18 females and 22 males; mean age: 763 days; age range: 730-786 days). Four additional participants were tested, but not included due to lack of attention during test (3), or difference scores exceeding 2.5 standard deviations from the mean of either speaker (1). Participants were randomly assigned to one of two affiliation conditions (20 per condition).

4.1.2 Stimuli

Audio Stimuli

See Section 3.1.2.

Audiovisual Stimuli (Exposure Phase)

See Section 3.1.2. There was one additional silent video that occurred immediately before the test phase. In this video a third speaker (hereafter referred to as the Affiliated Speaker) stood between the Training and Extension speaker. For half the participants the Affiliated speaker wore the same colour t-shirt as the Training Speaker; she looked at both speakers, and then waved at the Training Speaker, and the two embraced (see Figure 12). For the other half of the participants the Affiliated speaker wore the same colour t-shirt as the Extension Speaker; again she looked at both speakers, and then waved at the Extension Speaker, and the two embraced.

4.1.3 Procedure

The procedure was identical to that of Section 3.1.3, with the addition of the affiliation video described above in the Exposure Phase. At test there were three test trials, one for each talker (Affiliated Speaker, Training Speaker, and Extension Speaker).

For all participants the Affiliated Speaker appeared first at test. The order of the second and third trials was counterbalanced for the Training and Extension Speakers (i.e., half the participants saw the order Affiliated, Training, Extension, and the other half saw Affiliated, Extension, Training). As previously mentioned, whether the Affiliated Speaker affiliated with the Training or Extension Speaker was balanced across participants. The side on which the trained object appeared was counterbalanced across participants.

Based on the results of the previous chapter, toddlers should map the novel label "tIpu" onto the untrained object for the Training Speaker, and to the trained object for the Extension Speaker. If toddlers interpret a speaker's utterances based on social group membership, and use affiliative behaviour as a cue to social group membership, they should interpret the Affiliated Speaker's "tIpu" differently, depending on which speaker she affiliated with. When she affiliates with the Training Speaker, "tIpu" should be interpreted as referring to the untrained object. However, when she affiliates with the Extension Speaker, "tIpu" should be interpreted as referring to the trained object.

4.1.4 Coding of looking times

See Section 3.1.4.

Figure 12. Visual depiction of the affiliation event in Experiment 7.

4.1.5 Results

Overall Analysis

A repeated measures ANOVA on the participants' calculated difference scores with the within-subjects factor of Speaker (Affiliated, Training, Extension) and between-subjects factor of Condition (affiliates with Training Speaker vs. affiliates with Extension Speaker) revealed a main effect of Speaker, $F(2,38) = 15.23$, $p < .001$, a main effect of Condition, $F(2, 38) = 4.28$, $p = .045$, and a Speaker * Condition interaction, $F(2, 38) = 4.20$, $p = .019$. Due to the Speaker * Condition interaction, analyses for each condition were run separately.

Affiliation to the Training Speaker

A repeated measures ANOVA on the participants' calculated difference scores with the within-subjects factor of Speaker and between-subjects factor of test Order revealed a main effect of Speaker, $F(2,18) = 13.42$, $p < .001$, and no main effect of Order, $F(2, 18) = 0.58$, $p = .456$, and no Speaker * Order interaction, $F(2,18) = 0.55$, $p = .584$. Paired sample t-tests revealed no difference in looking behaviour for the Affiliated Speaker and the Training Speaker, $t(19) = 0.11$, $p = .912$, but significant differences between the

Affiliated Speaker and the Extension Speaker, $t(19) = 3.82$, $p = .001$, and the Training Speaker and the Extension Speaker, $t(19) = 4.20$, $p < .001$ (Figure 13).

To determine the effect of labelling for each talker separately, one-sample t-tests compared difference scores for each speaker against chance (where chance = a difference score of 0). As predicted, looking significantly decreased to the trained object when the Training Speaker said "tIpu", $t(19) = -3.24$, $p = .004$, and when the Affiliated Speaker said "tIpu", $t(19) = -3.77$, $p = .001$. In contrast, for the Extension Speaker, looking significantly increased to the trained object, $t(19) = 2.47$, $p = .023$. Thus, when the Training Speaker and the Affiliated Speaker said "tIpu", toddlers increased their looking toward the untrained object, but when the Extension Speaker said "tIpu", they increased their looking toward the trained object. Thus, toddlers interpreted words from the Affiliated Speaker in the same way as they did for Training Speaker.

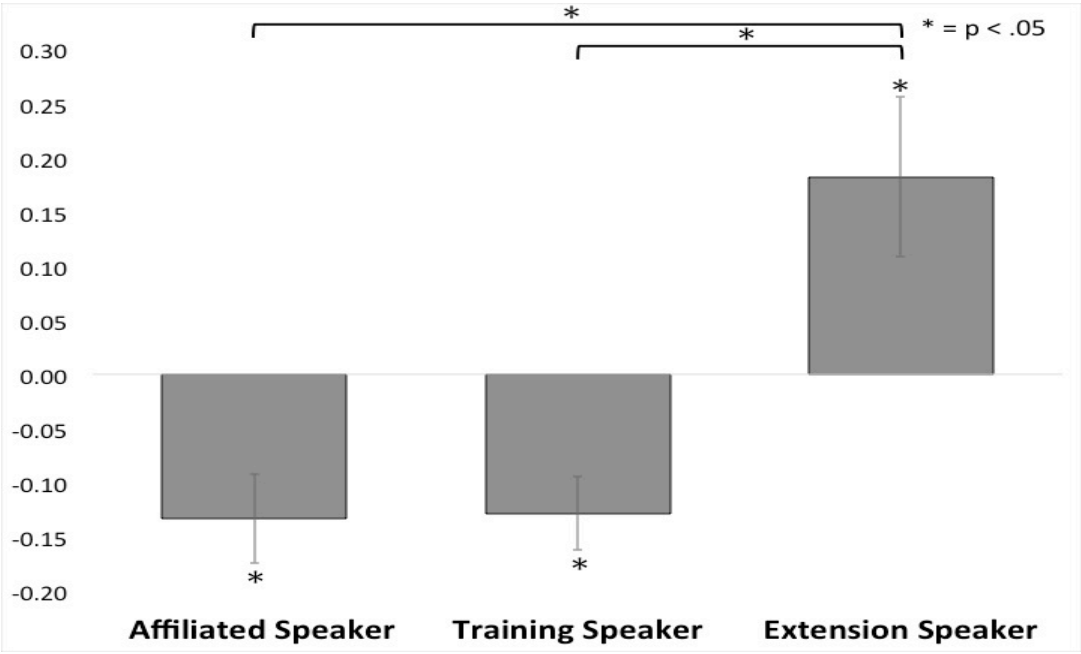

Figure 13. Difference scores and standard errors for each speaker in Experiment 7 (Affiliation with Training Speaker condition). Positive scores reflect an increase in

looking to the trained object while negative scores reflect a decrease in looking to the trained object.

Affiliation to the Extension Speaker

A repeated measures ANOVA on these difference scores with the within-subjects factor of Speaker and between-subjects factor of test Order revealed a main effect of Speaker, $F(2,18) = 7.27$, $p = .002$, a marginal effect of order, $F(2, 18) = 4.15$, $p = .057$, and no Speaker * Order interaction, $F(2,18) = 0.60$, $p = .555$. Paired sample t-tests reveal no difference in looking behaviour for the Affiliated Speaker and the Extension Speaker, $t(19) = 0.42$, $p = .680$, but significant differences between the Affiliated Speaker and the Training Speaker, $t(19) = 3.14$, $p = .005$, and the Training Speaker and the Extension Speaker, $t(19) = 3.27$, $p = .006$ (Figure 14).

To determine the effect of labelling for each talker separately, one-sample t-tests compared difference scores for each speaker against chance. As predicted, looking significantly increased to the trained object when the Extension Speaker said "tIpu", $t(19) = 3.32$, $p = .004$, and when the Affiliated Speaker said "tIpu" $t(19) = 3.05$, $p = .007$. In contrast, for the Training Speaker, looking did not significantly change, $t(19) = 1.57$, $p = .113$. Thus, when the Extension Speaker and the Affiliated Speaker said "tIpu", toddlers increased their looking toward the untrained object; but while toddlers showed a different pattern of looking for the Training Speaker, looking did not significantly increase to the untrained object. Overall, toddlers interpreted words from the Affiliated Speaker in the same way as they did for Extension Speaker.

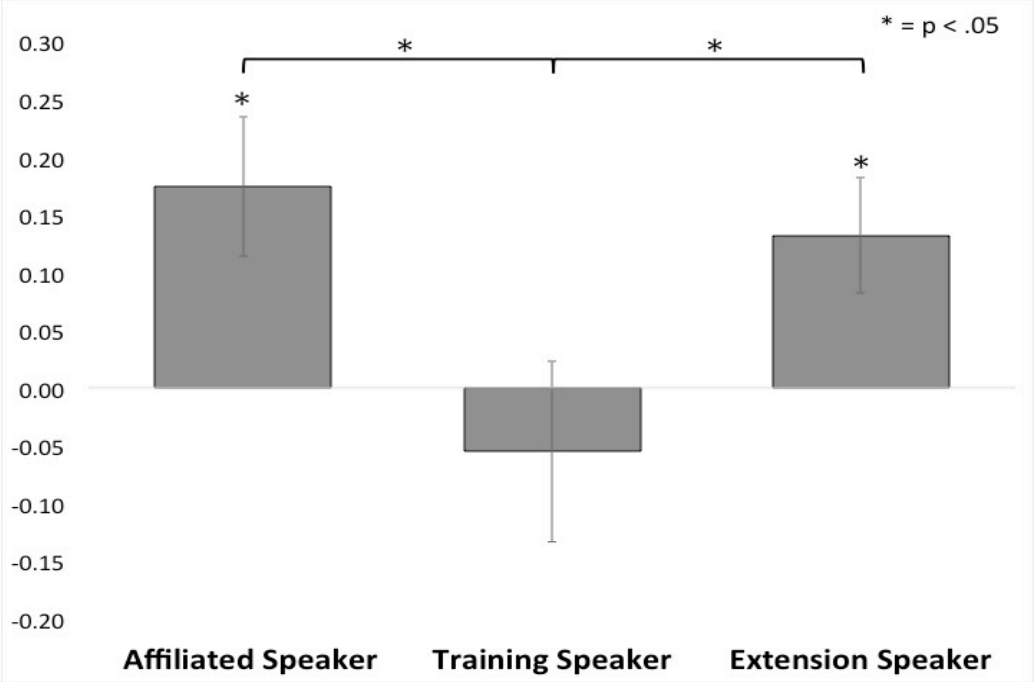

Figure 14. Difference scores and standard errors for each speaker in Experiment 7 (Affiliation with Extension Speaker condition). Positive scores reflect an increase in looking to the trained object while negative scores reflect a decrease in looking to the trained object.

4.1.6 Discussion

This experiment demonstrates that toddlers use previous affiliative behaviour as a predictor of linguistic group membership. That is, toddlers' referential interpretations for the Affiliated Speaker relied on which speaker she was shown to affiliate with. Thus, toddlers use both visually salient (Chapter 3) and more abstract (Chapter 4) cues when predicting how a speaker will produce words.

This finding reinforces the findings of the previous chapter, that social information is indexed to toddlers' linguistic representations. The current study expands on this to also demonstrate that the social information linked to toddlers' linguistic

representations can be based on more abstract social relationships as well. An interesting question for future research is whether race or affiliative behaviour is more heavily weighted when the two are pitted against each other. As previously mentioned, race, although a particularly salient social group marker, is not always indicative of linguistic variation. A more reliable indicator of an individual's linguistic group membership is the group they actually choose to be a part of. An empirical study that combines aspects of Chapters 3 and 4 could answer this question. For example, in an adaption of the procedure of Chapter 3, the Extension Speaker could be an other-race speaker and the Training Speaker could be Caucasian. An Affiliated Speaker, of the same race as the Extension Speaker, could then affiliate with the training speaker. Thus, toddlers would receive conflicting information, as the Affiliated Speaker would be the same race as the Extension Speaker, but affiliated with the Training Speaker. If toddlers recognize that affiliation is more important than race for language-based social groups, then they should treat the Affiliated Speaker's "tIpu" the same way as the Training Speaker's. However, if toddlers view race information as being a stronger predictor of language variation, then they should treat the Affiliated Speaker's "tIpu" the same way as the Extension Speaker's.

Overall, Chapter 5 demonstrates that 2-year-olds generalize linguistic properties across social group members, based on previous affiliative behaviour. This demonstrates that more abstract social information is indexed to toddlers' word representations.

5. Discussion

5.1 Summary

In three chapters I demonstrate that social group information plays a role in infants' and toddlers' speech processing. In Chapter 2, I showed that social information is indexed to 16-month-old infants' linguistic representations. When encountering a speaker of an unfamiliar race, from a social group they have no previous experience with, infants do not use linguistic features from their own social group for that speaker. In fact, infants must gain some experience with the speaker before recognizing her pronunciations. Chapter 3 demonstrates that the link between race and linguistic information is formed very quickly, even after only very little experience with a single group member. This information then influences toddlers' future interactions with members from the same race, but not just any speaker from a novel social group. This suggests that the social information being indexed to linguistic representations is highly specific. Finally, Chapter 4 shows that toddlers also use abstract cues such as previous affiliative behaviour to determine which social group an individual belongs to, and to predict how she will pronounce words. Thus, social information is indexed to toddlers' linguistic representations, and is used when interpreting new speakers' utterances.

5.2 Implications

As previously discussed, social information influences adults' speech processing. Adults are sensitive to linguistic differences, and make rich social inferences on the basis of them (e.g., Giles & Billings, 2004; Gluszek & Dovidio, 2010; Labov, 2006). Likewise, social information about a speaker influences their speech perception (e.g., Drager, 2006;

Niedzielski, 1999). Young children also appear to be sensitive to linguistic variation when making social inferences (Kinzler & DeJesus, 2013; Weatherhead et al., 2016). Additionally, sociophonetic variation can be observed in their own productions both across language varieties (e.g., Jacewicz Fox, & Salmons, 2011), and within a dialect community (e.g., Foulkes, Docherty, & Watt, 2005; Smith, Durham, & Fortune, 2007). Thus, even in early childhood children have some representation of sociolinguistic information. The current work is the first to suggest that social information is indexed to their linguistic representations at the earliest stages of language acquisition.

Foulkes and Docherty (2006) suggest that because sociolinguistic categories are formed through experience with variation, forming associations between social indices and linguistic properties may take an extended amount of time. However, I show that very little experience is needed before toddlers use social information during speech processing. Foulkes and Docherty (2006) also predict that those social categories that are most frequently encountered and that are most transparent (e.g. gender, ethnicity) would be the first to be acquired. Social categories that are not transparent, or are arbitrarily defined, would be last to be acquired. However, in Chapter 4, I show that toddlers are sensitive to more subtle, cues in addition to the very obvious ones. Thus, not only is the link between social and linguistic information present earlier than previously thought, it may be more sophisticated as well.

5.3 Incorporating social information in speech processing

Earlier I discussed the influence of social information on adults' speech processing, with factors such as race, nationality, gender, etc. affecting sound and word

processing. The findings of this dissertation suggest that even in infancy and toddlerhood, social information is linked to linguistic representations. Thus, a complete model of developing speech perception should incorporate these social influences. In this section I briefly outline two popular types of models of adult speech perception that have attempted to incorporate sociolinguistic information. I discuss both in terms of my findings.

The first type of model I will discuss are exemplar models (Boomershine, 2005; Foulkes & Docherty, 2006; Hay & Drager, 2007; Pierrehumbert, 2002). In exemplar models of speech perception, the lexicon exists as a distribution of stored tokens, or exemplars, corresponding to particular words. Each exemplar contains both linguistic information, and talker-specific information, such as the talker's acoustic and phonetic detail. Some accounts suggest that linguistic information is eventually abstracted away from particular exemplars, and forms a more general template for speech typical of that category (e.g., Pierrehumbert, 2006), but others do not (e.g., Goldinger, 1998). During speech perception, exemplars are activated based on their similarity to the input, both in terms of phonological patterning and indexical dimensions, and the appropriate linguistic category is selected based on the activation of the exemplar distribution.

Many have suggested that these exemplars also store social information about the speaker (e.g., Foulkes & Docherty, 2006; Hay & Drager, 2007). For example, a common feature of Japanese-Accented English is the pronunciation of the phoneme /l/ as [r]. In Japanese, the phoneme /l/ does not exist, and /r/ has a tap articulation (unlike English). As a result, English words with /l/ will fall in different clusters of exemplars, corresponding with (in this case) whether they are produced by North American or

Japanese speakers. If these exemplars also include information about social characteristics of speakers (e.g., Caucasian vs. Asian ethnicity), this type of account can help explain why, for example, American listeners are poorer at comprehending native-Accented English when shown a picture of a Chinese individual, or why vowel classifications vary by whether the listener was told that the speaker is an Australian or a New Zealander.

In these models, direct experience hearing a particular kind of variation is needed in order to form sociolinguistic categories (e.g., Foulkes & Docherty, 2006), although the amount of experience needed to form a category is unclear. This thought brings us to the current studies. The results of Chapter 2 demonstrate what happens when infants have no or few exemplars of a social category. In the first block, infants did not accept even the familiar pronunciations for words from an other-race speaker, demonstrating that they needed some experience with the speaker's pronunciations before they could effectively interpret her speech. The following chapters show that after exposure to only one speaker from a category, toddlers apply their pronunciations to a new speaker of the same type (Chapter 3) or speakers who have been shown to affiliate with that speaker (Chapter 4).

A second approach to modelling speech perception is through Bayesian approaches (Clayards, Tanenhaus, Aslin, & Jacobs, 2008; Feldman, Griffiths, & Morgan, 2009; Kleinschmidt & Jaeger, 2015; Norris & McQueen, 2008). Bayesian models, like exemplar models, rely on emerging patterns in the environment. However, in Bayesian models, a linguistic category is represented by a probabilistic distribution, rather than a distribution of stored tokens.

Under this framework, listeners have probabilistic beliefs based on linguistic regularities observed in their previous interactions. Thus, rather than having access to the distribution of exemplars, decisions are based on a belief about the distribution. It should be noted that a "belief" in this context does refer to a conscious belief, but a probability interpretation (i.e., how likely an outcome is based on the available evidence). A Bayesian model could easily incorporate social information, as there are also regularities in how talkers vary within a language, or within a social group. As listeners interact with their environment, their beliefs are updated through experience. For example, upon first encountering a Japanese speaker, a listener may form a belief that, for this type of speaker, /r/ sounds like /l/. When they encounter another Japanese speaker with the same linguistic properties the belief is strengthened. However, when they encounter a Japanese speaker that does not share that linguistic property the belief is weakened. If 99 out of 100 Japanese Speakers produce /r/ like /l/, then the listener will have very strong expectations about how speaker 101 will produce words. However, if 50 out of 100 Japanese Speakers produce /r/ like /l/, the listener will have very uncertain expectations.

The results from Chapters 2-4 could be very easily integrated into a Bayesian account as well. In Chapter 2, upon first encountering the other-race speaker, infants have a very uncertain belief about how she will produce words. After some exposure to the speaker they are able to adapt to her pronunciations. In Chapters 3 and 4, after some experience with a speaker, they form a belief about how speakers from that social group will produce words and apply that belief to new speakers of the same race (Chapter 3), or speakers who previously affiliated with that social group (Chapter 4).

Both Exemplar models and Bayesian models are appealing, as there have been a number of studies that suggest infants track these kinds of information. For example, a number of studies have demonstrated that infants fail to recognize familiarized wordforms across different affects (i.e., a word learned in a happy affect is not recognized later when produced in sad affect; Singh, Morgan, White, 2004), speakers of different genders (Houston & Jusczyk, 2000), or differences in amplitude or pitch (Singh, White, Morgan, 2008). Thus, young infants require high perceptual similarity across tokens for recognition to occur, which suggests their linguistic representations index acoustic information specific to the speaker. At the same time, studies have demonstrated that infants are capable of learning linguistic information by observing statistical distributions in their input. For example, they use transitional probabilities to segment fluent speech into words (e.g., Saffran, Aslin, & Newport, 1996), track phonological and phonotactic regularities (e.g., Seidl, Cristia, Bernard, & Onishi, 2009; White, Pepperkamp, Kirk, & Morgan, 2008), and attend to distributional cues when learning phonetic categories (e.g., Maye, Werker, & Gerken, 2002). Thus, both types of models are plausible and could account for the results observed in the current studies.

5.4 Alternative explanation

The way I have framed the results thus far has been in terms of sociolinguistic categories. That is, social information is linked to linguistic representations, which in turn affects speech perception. However, it is possible that infants' and toddlers' behaviour in the current study also reflects their metalinguistic knowledge. That is, they may have an expectation that language conventions are held specifically within a socially

defined linguistic group. These two interpretations are not mutually exclusive; however, while the first is undoubtedly supported by the current findings, the latter is a richer interpretation. Below I outline the evidence for the latter interpretation in other work, and describe the results of the current studies under this interpretation.

5.4.1 Children's appreciation of the conventionality of language

As previously mentioned, young language learners initially struggle to recognize familiar words when they are produced with an accent, unless they are given sufficient exposure previously (Best et al., 2009; van Heugten & Johnson, 2014; van Heugten & Johnson, 2015; White & Aslin, 2011). Infants and toddlers may struggle with accented speech because they assume that new word forms correspond to new referents (Clark, 1990; 1992; 1997; 2007; Golinkoff et al., 1994; Halberda, 2003; Markman, 1989; 1990; Merriman & Bowman, 1989; Shukla, White & Aslin, 2011). One explanation for this is that children understand that the forms of words are used in conventional ways in their community. "Bottle" has an established meaning within the child's language community. Thus when encountering something like "bettle", children infer that the speaker wishes to express a different meaning (if the speaker had intended to refer to a bottle they would have used the conventional label).

Work with pre-school and school-aged children has suggested that children infer that word meanings are limited to individuals who share the same conventions or who share the same knowledge (Diesendruck, 2005; Schell, 2016). For example, in Diesendruck (2005), 3 and 4-year-old Hebrew-English bilingual children were taught labels in English for one of two novel objects and later tested by a puppet who was either

a monolingual Hebrew speaker or, like the children, a Hebrew-English bilingual speaker. When tested by the monolingual Hebrew puppet, children interpreted the novel label as referring to either of the objects, suggesting that they inferred the Hebrew puppet would not know the English labels. In contrast, when the puppet was bilingual, children were more likely to assume a novel label referred to the object unnamed earlier by the English speaker. These results indicated that children do not expect speakers of different languages to have knowledge of the word meanings outside their language. Another demonstration of this point is that children do not show a disambiguation response across languages or accents (Au & Glusman, 1990; Weatherhead & White, 2016).

Work showing selective trust or endorsement may also reflect an understanding of conventional practices. For example, 4-year-old children readily accept the false testimony of a race-and-accent in-group speaker for what happened during an event, even when children themselves observed firsthand what happened; but, under the same circumstance they reject the false testimony when it is provided by a race-and-accent outgroup speaker (McDonald & Ma, 2016). These results suggest that children are more likely to learn from members of their perceived social group than individuals from another social group. These results are consistent with the notion that children's behaviour is driven by an understanding that the information provided by one's own social group has more relevance than information provided by an outgroup member. These types of effects are observed in non-linguistic situations as well - children selectively learn non-linguistic information (e.g., how to use a novel object) from native-accented speakers over foreign-accented speakers (Howard, Henderson, Carrazza, & Woodward, 2015; Kinzler, Corriveau, & Harris, 2011). Children at this age have also

demonstrated an understanding of conventional differences in other domains. For example, older children recognize that individuals from different social groups may have different conventional practices (e.g., Schmidt, Rakoczy, & Tomasello, 2011; 2012). In fact, even preverbal infants have some expectation that members of the same social group will act in a consistent way while members from different groups will act in an inconsistent way (Powell & Spelke, 2013).

5.4.2 Toddlers' appreciation of conventionality

The results of this dissertation could also be interpreted as infants and toddlers having some appreciation of the fact that their beliefs about how words are pronounced are only relevant for their own linguistic community, and that these pronunciation conventions should not necessarily extend to other linguistic communities. For Chapter 2, it could be that infants may expect other-race speakers to talk in novel ways. These expectations could have been fairly general – for instance, assuming that looking novel leads to talking in any novel way. However, infants in Chapter 2 went beyond simple novelty-novelty matching, and expected other-race speakers to speak in some systematic novel way. When encountering the speaker for the first time they were unsure what pronunciation conventions to assign to her and waited for evidence to determine the nature of the difference. In other words, they expected that a speaker from a novel social group *could* produce words in a way different from their own linguistic community. In Chapters 3 and 4, after learning about one speaker's pronunciations, they applied this pronunciation convention to new speakers of the same social group. This is true for both racially-defined groups (Chapter 3) and affiliatively-defined groups (Chapter 4). Again,

these results may reflect toddlers' expectations that members of the same social group should talk in the same way (i.e., Experiments 5 and 7), and those from different social groups should not talk in the same way (i.e., Experiment 6).

The results of the current studies clearly suggest that at a minimum, social information is indexed to infants' and toddlers' linguistic representations. Infants and toddlers would be unable to respond to different types of speakers differently if they did not have distinct linguistic representations for different social groups. However, whether they also have more metalinguistic expectations, as discussed above, remains to be seen. It is certainly plausible that a combination of the two explanations is driving the results. But, for the time being, I endorse the more conservative explanation.

5.5 Remaining Questions

One remaining question from the current studies is what social information is indexed to infants' linguistic representations. Infants do not have a great deal of experience with different linguistic groups and therefore may not recognize what social information is relevant to linguistic differences. Certain information about speakers might be particularly salient to infants, but not relevant to certain linguistic differences. Just because two speakers are wearing the same colour of shirt, or both have reading glasses, does not mean they will necessarily have the same linguistic properties. However, infants may still index this irrelevant speaker information when learning a new linguistic variety. For example, an English infant may coincidentally only be exposed to French speakers who have blonde hair. The infant may erroneously infer that having blonde hair is a relevant social factor for this particular linguistic pattern.

Alternatively, certain social information may not be salient to infants, but very relevant to linguistic differences. For example, geographic background is a highly relevant predictor of linguistic group membership. Children as young as 3 years old infer that speakers who speak with the same accent are from the same place, and 4- and 5-year-olds infer that speakers with different accents are from different places (Weatherhead, White, & Friedman, 2016). Furthermore, Canadian 4-, 5-, and 6-year-olds infer that speakers who share their native accent were born in Canada, while speakers with foreign accents were born far away (Weatherhead, Friedman, & White, 2017). However, whether knowing that someone is Canadian vs. foreign affects children' processing of that person's speech is unknown. This is clearly the case in adults, as information about a person's nationality affects classification of their vowels (Niedzielski, 1999; Hay et al., 2006; Hay & Drager, 2010). While infants would not have a concept of geographic background in this way, they may have some concept of location more generally as a social group boundary. Whether a non-obvious property like the environmental context the speaker was speaking in is indexed to their utterances is a question I intend to investigate in future research.

A final question that remains, which was raised in the discussion of Chapter 4, is the extent to which different types of social information are valued in children's linguistic processing. Chapters 3 and 4 demonstrate that toddlers' linguistic processing is sensitive to both race and previous affiliative behaviour when they are presented in isolation. If these two indicators were pitted against each other, which would be more heavily weighted? In this particular comparison, affiliative behaviour should be weighed more heavily. As previously mentioned, race is not always indicative of linguistic variation. A

more reliable indicator of an individual's linguistic group membership is the group they actually choose to be a part of.

However, whether toddlers recognize this may be influenced by how much exposure they have to racial diversity. Toddlers who have very little experience with racial diversity may weigh race heavily as a linguistic predictor, as it is very visually salient. Thus, they may use the race of the Affiliated Speaker to guide their linguistic expectations. In contrast, those who have a lot of experience with racial diversity may have an appreciation that race is not always a reliable linguistic marker, and therefore use the Affiliated Speaker's behaviour to predict how she will speak (consistent with Kinzler & Dautel, 2012).

A final possibility is that perhaps toddlers do not even consider race in the presence of alternative cues; that is, regardless of prior experience with diversity, perhaps toddlers do not heavily weigh race in their linguistic expectations. This would be consistent with the theory that race is not an essential social property. Those who subscribe to this viewpoint argue that race was not a factor denoting group membership evolutionarily, as neighbouring groups historically would not look different in terms of race and individuals were unlikely to have travelled more than 40 miles (Baker, 2002). Differences in race would have little value in distinguishing members from neighbouring coalitions, and thus, it is unlikely that there were cognitive adaptations to preferentially encode this dimension (Cosmides, Tooby, & Kurzban, 2003; Kurzban, Tooby, & Cosmides, 2001). Evidence that race does not a play a strong role in infants' and toddlers' social preferences, but does in older children and adults, supports this claim (Baron & Banaji, 2006; Dovidio, Kawakami, Gaertner, 2002; Kinzler & Spelke, 2011).

Overall, there is a possibility that race may not be considered as strongly as previous affiliative behaviour. However, infants and toddlers are still extremely sensitive to race as a cue to linguistic group membership as we saw in Chapter 1, and race itself is a very salient speaker feature (Bar-et al., 2006; Kelly et al., 2005). Thus, it is unclear which of the two cues would win when pitted against each other.

Concluding Remarks

Speech perception is a complex, multifaceted process. Factors, such as race, gender, age, social class, and nationality can influence how we perceive someone's speech. Thus, in order to have a more complete understanding of how this process works we must account for social factors affecting perception. The findings of my dissertation demonstrate that social information is indexed to the linguistic representations of even the youngest language learners. These findings make the need for a more comprehensive model of perception even more necessary, as social information appears to play a role in perception right from the beginning.

References

Aboud, F. E. (1988). *Children and prejudice. The development of ethnic awareness and identity*. Oxford, UK: Blackwell

Anisfeld, M., Bogo, N., & Lambert, W. E. (1962). Evaluational reactions to accented English speech. *The Journal of Abnormal and Social Psychology*, *65*(4), 223.

Anzures, G., Quinn, P. C., Pascalis, O., Slater, A. M., & Lee, K. (2010). Categorization, categorical perception, and asymmetry in infants' representation of face race. *Developmental Science*, *13*(4), 553-564. doi: 10.1111/j.1467-7687.2009.00900.xView

Au, T. K. F., & Glusman, M. (1990). The principle of mutual exclusivity in word learning: to honor or not to honor? *Child development*, *61*(5), 1474-1490.

Babel, M., & Russell, J. (2015). Expectations and speech intelligibility. *The Journal of the Acoustical Society of America*, *137*(5), 2823-2833.

Bahrick, L. E., Netto, D., & Hernandez-Keif, M. (1998). Intermodal perception of adult and child faces and voices by infants. *Child development*, *69*(5), 1263-1275.

Baker, M. C. (2002). *The atoms of language: The mind's hidden rules of grammar*. Basic books, NY.

Bar-Haim, Y., Ziv, T., Lamy, D., & Hodes, R. M. (2006). Nature and nurture in own-race face processing. *Psychological Science*, *17*(2), 159-163.

Baron, A. S., & Banaji, M. R. (2006). The development of implicit attitudes: Evidence of race evaluations from ages 6 and 10 and adulthood. *Psychological Science*, *17*(1), 53-58.

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

Bell, A. (1984). Language style as audience design. *Language in society*, *13*(02), 145-204.

Bell A. (2001). Back in style: reworking audience design. In Style and Sociolinguistic Variation, ed. P Eckert, JR Rickford, pp. 139–69. Cambridge, UK: Cambridge University Press

Bent, T. (2014). Children's perception of foreign-accented words. *Journal of Child Language*, *41*(06), 1334-1355.

Best, C. T., Tyler, M. D., Gooding, T. N., Orlando, C. B., & Quann, C. A. (2009). Development of phonological constancy: Toddlers' perception of native- and Jamaican-accented words. *Psychological Science*, *20*(5), 539–42. doi:10.1111/j.1467-9280.2009.02327.x

Bigler, R. S., Jones, L. C., & Lobliner, D. B. (1997). Social categorization and the formation of intergroup attitudes in children. *Child development*, *68*(3), 530-543.

Boersma, P. & Weenink, D. (2013). Praat: doing phonetics by computer [Computer program]. Version 5.3.51, retrieved 2 June 2013 from http://www.praat.org/

Boomershine, A. R. (2005). *Perceptual processing of variable input in Spanish: An exemplar-based approach to speech perception* (Doctoral dissertation, The Ohio State University).

Bosch, L., & Sebastian-Galles, N. (1997). Native language recognition abilities in 4 month old infants from monolingual and bilingual environments. *Cognition*, 65, 33-69.

Bresnahan, M. J., Ohashi, R., Nebashi, R., Liu, W. Y., & Shearman, S. M. (2002). Attitudinal and affective response toward accented English. *Language & Communication*, *22*(2), 171-185.

Byers-Heinlein, K., Burns, T. C., & Werker, J. F. (2010). The roots of bilingualism in newborns. *Psychological Science*.

Byrne, D., & Nelson, D. (1965). The effect of topic importance and attitude similarity-dissimilarity on attraction in a multistranger design. *Psychonomic Science*, *3*(1-12), 449-450.

Campbell-Kibler, K. (2007). Accent,(ING), and the social logic of listener perceptions. *American speech*, *82*(1), 32-64.

Cargile, A. C., & Giles, H. (1997). Understanding language attitudes: Exploring listener affect and identity. *Language & Communication*, *17*(3), 195-217.

Clark, E. V. (1990). On the pragmatics of contrast. *Journal of child language*,*17*(02), 417-431.

Clark, E. V. (1992). Conventionality and contrast: Pragmatic principles with lexical consequences. *Lehrer and Kittay, 1992a*, 171-188.

Clark, E. V. (1997). Conceptual perspective and lexical choice in acquisition.*Cognition*, *64*(1), 1-37.

Clark, E. V. (2007). Conventionality and contrast in language and language acquisition. *New directions for child and adolescent development*, *115*, 11-23.

Clayards, M., Tanenhaus, M. K., Aslin, R. N., & Jacobs, R. A. (2008). Perception of speech reflects optimal use of probabilistic speech cues. *Cognition*, *108*(3), 804-809.

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, and Computers*, *25*(2), 257-271. [retrieved from http://psy.cns.sissa.it]

Colombo, J., & Bundy, R. S. (1981). A method for the measurement of infant auditory selectivity. *Infant Behavior and Development*, *4*, 219-223.

Cosmides, L., Tooby, J., & Kurzban, R. (2003). Perceptions of race. *Trends in Cognitive Sciences*, *7*(4), 173-179.

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, & Computers*, *28*(1), 125-127.

Dailey, R. M., Giles, H., & Jansma, L. L. (2005). Language attitudes in an Anglo-Hispanic context: The role of the linguistic landscape. *Language & Communication*, *25*(1), 27-38.

Devos, T., & Banaji, M. R. (2005). American= White? *Journal of Personality and Social Psychology*, *88*(3), 447.

Diesendruck, G. (2005). The principles of conventionality and contrast in word learning: An empirical examination. *Developmental Psychology*, *41*(3), 451.

Drager, K. (2006). From bad to bed: the relationship between perceived age and vowel perception in New Zealand English. Te Reo. 48:55–68.

Dornic, S., Nystedt, L. Laaksonen, T., & Arnberg, L. (1989). Evaluational reactions to speech: The role of ethnic-linguistic status. *Perceptual and Motor Skills*, 69, 307-317.

Dovidio, J. F., Kawakami, K., & Gaertner, S. L. (2002). Implicit and explicit prejudice and interracial interaction. *Journal of personality and social psychology*, *82*(1), 62.

Feldman, N. H., Griffiths, T. L., & Morgan, J. L. (2009). The influence of categories on perception: explaining the perceptual magnet effect as optimal statistical inference. *Psychological review*, *116*(4), 752.

Foulkes, P., & Docherty, G. J., (2006). The social life of phonetics and phonology. *Journal of Phonetics*. 34(4):409– 38

Foulkes, P., Docherty, G. J., & Watt, D. (2005). Phonological variation in child-directed speech. *Language*, *81*(1), 177-206.

Gelman, S. A., Collman, P., & Maccoby, E. E. (1986). Inferring properties from categories versus inferring categories from properties: The case of gender. *Child Development*, 396-404.

Giles, H., & Billings, A. (2004). *Language attitudes*. In A. Davies & E. Elder (Eds.), *The handbook of applied linguistics* (pp. 187–209). Oxford, UK: Blackwell.

Giles, H., Henwood, K., Coupland, N., Harriman, J., & Coupland, J. (1992) Language attitudes and cognitive mediation. *Human Communication Research*, 18, 500–27.

Gluszek, A., & Dovidio, J. F. (2010). The way they speak: A social psychological perspective on the stigma of nonnative accents in communication. *Personality and Social Psychology Review*, *14*(2), 214-237.

Goldinger, S. D. (1998). Echoes of echoes? An episodic theory of lexical access. *Psychological review*, *105*(2), 251.

Golinkoff, R. M., Mervis, C. B., & Hirsh-Pasek, K. (1994). Early object labels: The case for a developmental lexical principles framework. *Journal of child language*, *21*(01), 125-155.

Halberda, J. (2003). The development of a word-learning strategy. *Cognition*, 87, B23-B34.

Halle, P. & de Boysson-Bardies, B. (1994). Emergence of an early lexicon: Infant's recognition of words. *Infant Behavior and Development*, 17, 119-129.

Hay, J., Nolan, A., & Drager, K. (2006). From fush to feesh: Exemplar priming in speech perception. *The Linguistic Review*, *23*(3), 351-379.

Hay, J., & Drager, K. (2007). Sociophonetics. *Annual Review of Anthropology*, *36*(1), 89–103. doi:10.1146/annurev.anthro.34.081804.120633

Hay, J., & Drager, K. (2010). Stuffed toys and speech perception. *Linguistics*, *48*(4), 865-892. doi: https://doi.org/10.1515/ling.2010.027

Hay, J., Jannedy, S., & Mendoza-Denton, N., (1999). Oprah and /ay/: lexical frequency, referee design and style. Proceedings at the International Congress of Phonetic Science, 14th, San Francisco, pp. 1389–92.

Henrich, N., & Henrich, J. P. (2007). *Why humans cooperate: A cultural and evolutionary explanation*. Oxford University Press.

Hirschfeld, L. A. (1996). *Race in the making*. Cambridge, MA: MIT Press

Hirschfeld, L. A., & Gelman, S. A. (1997). What young children think about the relationship between language variation and social difference. *Cognitive Development*, *12*(2), 213–238. doi:10.1016/S0885-2014(97)90014-9

Houston, D. M., & Jusczyk, P. W. (2000). The role of talker specific information in word segmentation by infants. *Journal of Experimental Psychology: Human Perception and Performance, 26,* 1570–1582.

Howard, L. H., Henderson, A. M., Carrazza, C., & Woodward, A. L. (2015). Infants' and Young Children's Imitation of Linguistic In-Group and Out-Group Informants. *Child development*, *86*(1), 259-275.

Huygens, I. & Vaughn, G. M. (1984) Language attitudes, ethnicity, and social class in New Zealand. *Journal of Multilingual and Multicultural Development*, 4, 207–24.

Jahoda, M. (1961). A social-psychological approach to the study of culture. *Human Relations*, 14, 1961, 23-30. doi:10.1177/001872676101400102

Jacewicz, E., Fox, R. A., & Salmons, J. (2011). Vowel change across three age groups of speakers in three regional varieties of American English. *Journal of phonetics*, *39*(4), 683-693.

Johnson K, Strand EA, D'Imperio M. 1999. Auditory-visual integration of talker gender in vowel perception. *Journal of phonetics*. *27*, 359–84.

Jusczyk, P., & Aslin, R. (1995). Infants' detection of the sound patterns of words I fluent speech. *Cognitive Psychology*, 29, 1-23.

Jusczyk, P., Cutler, A., & Redanz, N. (1993). Infants' preference for predominant stress patterns of English words. *Child Development*, 64, 675-687.

Jusczyk, P. W., Friederici, A. D., Wessels, J. M., Svenkerud, V. Y., & Jusczyk, A. M. (1993). Infants' sensitivity to the sound patterns of native language words. *Journal of memory and language*, *32*(3), 402.

Kang, O., & Rubin, D. L. (2009). Reverse linguistic stereotyping: Measuring the effect of listener expectations on speech evaluation. *Journal of Language and Social Psychology*.

Katz, P. A., & Kofkin, J. A. (1997). Race, gender, and young children. *Developmental psychopathology: Perspectives on adjustment, risk, and disorder*, 51-74.

Kelly, D. J., Liu, S., Lee, K., Quinn, P. C., Pascalis, O., Slater, A. M., & Ge, L. (2009). Development of the other-race effect during infancy: Evidence toward universality? *Journal of Experimental Child Psychology*, *104*(1), 105-114. doi: dx.doi.org/10.1016/j.jecp.2009.01.006

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Ge, L., & Pascalis, O. (2007). The other-race effect develops during infancy evidence of perceptual narrowing. *Psychological Science*, *18*(12), 1084-1089.

Kelly, D. J., Quinn, P. C., Slater, A. M., Lee, K., Gibson, A., Smith, M., Ge, L., & Pascalis, O. (2005). Three-month-olds, but not newborns, prefer own-race faces. *Developmental Science*, *8*(6), F31-F36. doi: 10.1111/j.1467-7687.2005.0434a.x

Kinzler, K. D., & Dautel, J. B. (2012). Children's essentialist reasoning about language and race. *Developmental science*, *15*(1), 131-138.

Kinzler, K. D., Dupoux, E., & Spelke, E. S. (2007). The native language of social cognition. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(30), 12577–80. doi:10.1073/pnas.0705345104

Kinzler, K. D., Corriveau, K. H., & Harris, P. L. (2011). Children's selective trust in native‐accented speakers. *Developmental science*, *14*(1), 106-111.

Kinzler, K. D., & DeJesus, J. M. (2013). Children's sociolinguistic evaluations of nice foreigners and mean Americans. *Developmental Psychology*, *49*(4), 655–64. doi:10.1037/a0028740

Kinzler, K. D., Shutts, K., DeJesus, J., & Spelke, E. S. (2009). Accent trumps race in guiding children's social preferences. *Social Cognition*, *27*(4), 623. doi:10.1521/soco.2009.27.4.623.

Kinzler, K. D., Shutts, K., & Spelke, E. S. (2012). Language-based social preferences among children in South Africa. *Language learning and development*, *8*(3), 215-232.

Kinzler, K. D., & Spelke, E. S. (2011). Do infants show social preferences for people differing in race?. *Cognition*, *119*(1), 1-9.

Kircher, M., & Furby, L. (1971). Racial preferences in young children. *Child Development*, 2076-2078.

Kleinschmidt, D. F., & Jaeger, T. F. (2015). Robust speech perception : Recognize the familiar, generalize to the similar, and adapt to the novel. *Psychological Review*, *122*(2), 148–203. doi:10.1037/a0038695

Kuczaj, S. A. (1982). *Language development: Syntax and semantics*, 1. Hillsdale, NJ: Psychology Press.

Kuczaj, S. & Harbaugh, B. (1982). What children think about the speaking capacities of other persons and things. (as cited in Kuczaj, S. A. (1982). *Language development: Syntax and semantics*, 1. Hillsdale, NJ: Psychology Press.)

Kuhl, P., Williams, K., Lacerda, F., Stevens, K., & Lindblom, B. (1992). Linguistic experience alters phonetic perception in infants by 6 months of age. *Science*, 255. 606-608.

Kuhlmeier, V., Wynn, K., & Bloom, P. (2003). Attribution of dispositional states by 12-month-olds. *Psychological science*, *14*(5), 402-408.

Kurzban, R., Tooby, J., & Cosmides, L. (2001). Can race be erased? Coalitional computation and social categorization. *Proceedings of the National Academy of Sciences*, *98*(26), 15387-15392.

Labov, W. (1966). The Social Stratification of English in New York City. Washington, DC: Center of Applied Linguistics.

Labov, W. (1972). Sociolinguistic Patterns. Philadelphia: University of Pennsylvania Press.

Labov, W. (2006). A sociolinguistic perspective on sociophonetic research. *Journal of Phonetics*, *34*(4), 500–515. doi:10.1016/j.wocn.2006.05.002

Ladefoged, P., & Broadbent, D. E. (1957). Information conveyed by vowels. *The Journal of the Acoustical Society of America*, *29*(1), 98-104.

Liberman, Z., Kinzler, K. D., & Woodward, A. L. (2014). Friends or foes: Infants use shared evaluations to infer others' social relationships. *Journal of experimental psychology: general*, *143*(3), 966.

Liberman, Z., Woodward, A. L., & Kinzler, K. D. (2016). Preverbal Infants Infer Third-Party Social Relationships Based on Language. *Cognitive Science*.

Maccoby, E. E., & Jacklin, C. N. (1987). Gender segregation in childhood. *Advances in child development and behavior*, *20*, 239-287.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction*. MIT Press.

Markman, E. M. (1990). Constraints children place on word meanings. *Cognitive Science*, *14*(1), 57-77.

Martin, C. L., Fabes, R. A., Evans, S. M., & Wyman, H. (1999). Social cognition on the playground: Children's beliefs about playing with girls versus boys and their relations to sex segregated play. *Journal of Social and Personal Relationships*, *16*(6), 751-771.

Mascaro, O., & Csibra, G. (2012). Representation of stable social dominance relations by human infants. *Proceedings of the National Academy of Sciences*, *109*(18), 6862-6867.

Mascaro, O., & Csibra, G. (2014). Human infants' learning of social structures the case of dominance hierarchy. *Psychological science*, *25*(1), 250-255.

Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, 82, B101–B111.

McCall, R. B., & Kennedy, C. B. (1980). Subjective uncertainty, variability of experience, and the infant's response to discrepancies. *Child development*, 285-287.

McDonald, K. P., & Ma, L. (2016). Preschoolers' credulity toward misinformation from ingroup versus outgroup speakers. *Journal of Experimental Child Psychology*, *148*, 87-100.

Mcgowan, K. B. (2015). Social Expectation Improves Speech Perception in Noise. *Language and Speech*. doi:10.1177/0023830914565191

Merriman, W. E., & Bowman, L. L. (1989). The mutual exclusivity bias in children's word learning. *Monographs of the Society for Research in Child Development*, 54 (3/4).

Moon, C., Cooper, R. P., & Fifer, W. P. (1993). Two-day-olds prefer their native language. *Infant behavior and development*, *16*(4), 495-500.

Mulak, K. E., Best, C. T., Tyler, M. D., Kitamura, C., & Irwin, J. R. (2013). Development of phonological constancy: 19-month-olds, but not 15-month-olds, identify words in a non-native regional accent. *Child Development*, *84*(6), 2064–78. doi:10.1111/cdev.12087

Nathan, L., Wells, B., & Donlan, C. (1998). Children's comprehension of unfamiliar regional accents: a preliminary investigation. *Journal of Child Language*, *25*, 343–365.

Newton, C., & Ridgway, S. (2016). Novel accent perception in typically-developing school-aged children. *Child Language Teaching and Therapy*, *32*(1), 111-123.

Niedzielski, N. (1996). Acoustic analysis and language attitudes in Detroit. *University of Pennsylvania Working Papers in Linguistics*, *3*(1), 7.

Norris, D., & McQueen, J. M. (2008). Shortlist B: a Bayesian model of continuous speech recognition. *Psychological review*, *115*(2), 357.

Okumura, Y., Kanakogi, Y., Takeuchi, S., & Itakura, S. (2014). Twelve-month-old infants show social preferences for native-dialect speakers. *Shinrigaku kenkyu: The Japanese journal of psychology*, *85*(3), 248-256.

Paltridge, J., & Giles, H. (1984). Attitudes towards speakers of regional accents of French: Effects of regionality, age and sex of listeners. *Linguistische berichte*, (90), 71-85.

Penny, R., Barrett, M., & Lyons, E. (2001). *Children's naïve theories of nationality: A study of Scottish and English children's national inclusion criteria.* Poster presented at the 10th European Conference on Developmental Psychology, Uppsala University, Uppsala, Sweden.

Piaget, J., & Weil, A. M. (1951). The development in children of the idea of the homeland and of relations with other countries. *International Social Science Bulletin*, *3*(3), 561-578.

Pierrehumbert J. (2002). Word-specific phonetics. In Laboratory Phonology 7, ed. C Gussenhoven, N Warner, pp. 101–39. Berlin/New York: Mouton de Gruyter

Pierrehumbert, J. (2006). The statistical basis of an unnatural alternation. *Laboratory phonology*, *8*, 81-107.

Phillipson, R., & Skutnabb-Kangas, T. (1994). English, panacea or pandemic. Proceedings *from Language issues in English-medium universities: a global concern*, University of Hong Kong.

Powell, L. J., & Spelke, E. S. (2013). Preverbal infants expect members of social groups to act alike. *Proceedings of the National Academy of Sciences*, *110*(41), E3965-E3972.

Quinn, P. C., Yahr, J., Kuhn, A., Slater, A. M., & Pascalis, O. (2002). Representation of the gender of human faces by infants: A preference for female. *Perception*, *31*(9), 1109-1121.

Rakić, T., Steffens, M. C., & Mummendey, A. (2011). Blinded by the accent! The minor role of looks in ethnic categorization. *Journal of Personality and Social Psychology*, *100*(1), 16–29. doi:10.1037/a0021522

Rickford, J. R., & McNair-Knox, F. (1994). Addressee-and topic-influenced style shift: A quantitative sociolinguistic study. *Sociolinguistic perspectives on register*, 235-76.

Roberts, J. (1997). Acquisition of variable rules: a study of (-t, d) deletion in preschool children. *Journal of Child Language*, 24(2), 351-372.

Roberts, J., & Labov, W. (1995). Learning to talk Philadelphian: Acquisition of short a by preschool children. *Language Variation and Change*, 7(1), 101-112.

Rubin, D. L. (1992). Nonlanguage factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education*, 33, 511–531.

Ryan, E., Carranza, M. A., & Moffie, R. W. (1977). Reactions toward varying degrees of accentedness in the speech of Spanish-English bilinguals. *Language and speech*, *20*(3), 267-273.

Saffran, J. R., Newport, E. L., & Aslin, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of memory and language*, *35*(4), 606-621.

Saffran, J. R., Werker, J. F., & Werner, L. A. (2006). The infant's auditory world: Hearing, speech, and the beginnings of language. *Handbook of child psychology*.

Sanefuji, W., Ohgami, H., & Hashiya, K. (2006). Preference for peers in infancy. *Infant behavior and development*, *29*(4), 584-593.

Schell, V. (2016). Preschoolers restrict the scope of labels within their own linguistic group (Masters Thesis). Retrieved from qspace.library.queensu.ca

Schmale, R., & Seidl, A. (2009). Accommodating variability in voice and foreign accent: Flexibility of early word representations. *Developmental Science*, *12*(4), 583-601.

Schmale, R., Cristia, A., Seidl, A., & Johnson, E. K. (2010). Developmental changes in infants' ability to cope with dialect variation in word recognition. *Infancy*, *15*(6), 650-662.

Schmale, R., Cristia, A., & Seidl, A. (2012). Toddlers recognize words in an unfamiliar accent after brief exposure. *Developmental Science*, *15*(6), 732-738.

Schmale, R., Seidl, A., & Cristià, A. (2015). Mechanisms underlying accent accommodation in early word learning: evidence for general expansion. *Developmental Science*, *18*(4), 664–670. doi:10.1111/desc.12244

Schmidt, M. F., Rakoczy, H., & Tomasello, M. (2011). Young children attribute normativity to novel actions without pedagogy or normative language. *Developmental science*, *14*(3), 530-539.

Schmidt, M. F., Rakoczy, H., & Tomasello, M. (2012). Young children enforce social norms selectively depending on the violator's group affiliation. *Cognition*, *124*(3), 325-333.

Seidl, A., Cristià, A., Bernard, A., & Onishi, K. H. (2009). Allophonic and phonemic contrasts in infants' learning of sound patterns. *Language Learning and Development*, *5*(3), 191-202.

Seyfarth, R. M., & Cheney, D. L. (2012). The evolutionary origins of friendship. *Annual review of psychology*, *63*, 153-177

Seyfarth, R. M., & Cheney, D. L. (2013). Affiliation, empathy, and the origins of theory of mind. *Proceedings of the National Academy of Sciences*, *110* (Supplement 2), 10349-10356.

Shell, M. (2001). Language wars. *CR: The New Centennial Review*, *1*(2), 1-17.

Sherif, M., Harvey, O., White, B. J., Hood, W., & Sherif, C.. (1961). Intergroup cooperation and conflict: The robbers cave experiment. *Norman, OK: University of Oklahoma Book Exchange*.

Shukla, M., White, K. S., & Aslin, R. N. (2011). Prosody guides the rapid mapping of auditory word forms onto visual objects in 6-mo-old infants. *Proceedings of the National Academy of Sciences*, *108*(15), 6038-6043.

Shutts, K., Kinzler, K. D., McKee, C. B., & Spelke, E. S. (2009). Social information guides infants' selection of foods. *Journal of Cognition and Development*, *10*(1-2), 1-17.

Singh, L., Morgan, J. L., & White, K. S. (2004). Preference and processing: The role of speech affect in early spoken word recognition. *Journal of Memory and Language*, *51*(2), 173-189.

Singh, L., White, K. S., & Morgan, J. L. (2008). Building a word-form lexicon in the face of variable input: Influences of pitch and amplitude on early spoken word recognition. *Language Learning and Development*, 4, 157-178.

Smith, J., Durham, M., & Fortune, L. (2007). "Mam, my trousers is fa'in doon!": Community, caregiver, and child in the acquisition of variation in a Scottish dialect. *Language Variation and Change*, *19*(01), 63-99.

Souza, A. L., Byers-Heinlein, K., & Poulin-Dubois, D. (2013). Bilingual and monolingual children prefer native-accented speakers. *Frontiers in psychology*, *4*, 953.

Sparks, A. (1996). *Tomorrow is Another Country: the inside story of South Africa's road to change*. University of Chicago Press, IL.

Spokes, A. C., & Spelke, E. S. (2016). Children's Expectations and Understanding of Kinship as a Social Category. *Frontiers in psychology*, *7*.

Spokes, A. C., & Spelke, E. S. (2017). The cradle of social knowledge: Infants' reasoning about caregiving and affiliation. *Cognition*, *159*, 102-116.

Stewart, M. A., Ryan, E. B., & Giles, H. (1985). Accent and social class effects on status and solidarity evaluations. *Personality and Social Psychology Bulletin*, *11*(1), 98-105.

Strand EA, Johnson K. 1996. Gradient and visual speaker normalization in the perception of fricatives. In Natural Language Processing and Speech Technology, ed. D Gibbon, pp. 14–26. Berlin: Mouton

Swingley, D. (2005). 11-month-olds' knowledge of how familiar words sound. *Developmental science*, *8*(5), 432-443.

Swingley, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, *13*(5), 480-484.

Tincoff, R., & Jusczyk, P. W. (1996). Are word-final sounds perceptually salient for infants. In *Fifth Conference on Laboratory Phonology, Evanston, IL*.

Uttley, L., de Boisferon, A. H., Dupierrix, E., Lee, K., Quinn, P. C., Slater, A. M., & Pascalis, O. (2013). Six-month-old infants match other-race faces with a non-native language. *International Journal of Behavioral Development*, *37*(2), 84–89. doi:10.1177/0165025412467583

Van Heugten, M., & Johnson, E. K. (2014). Learning to contend with Accents in Infancy: Benefits of brief speaker exposure. *Journal of Experimental Psychology: General*, *143*, 340-350.

Van Heugten, M., Krieger, D. R., & Johnson, E. K. (2015). The developmental trajectory of toddlers' comprehension of unfamiliar regional accents. *Language Learning and Development*, *11*(1), 41-65.

Vouloumanos, A., & Werker, J. F. (2004). Tuned to the signal: the privileged status of speech for young infants. *Developmental science*, *7*(3), 270-276.

Vouloumanos, A., & Werker, J. F. (2007). Listening to language at birth: Evidence for a bias for speech in neonates. *Developmental science*, *10*(2), 159-164.

Vouloumanos, A., Hauser, M. D., Werker, J. F., & Martin, A. (2010). The tuning of human neonates' preference for speech. *Child development*, *81*(2), 517-527.

Wagner, L., Clopper, C. G., & Pate, J. K. (2014). Children's perception of dialect variation. *Journal of child language*, *41*(5), 1062.

Weatherhead, D., & White, K. S. (2016). He Says Potato, She Says Potahto: Young Infants Track Talker-Specific Accents. *Language Learning and Development*, *12*(1), 92-103.

Weatherhead, D., Friedman, O., White, K. S. (2017). Accent, Language, and Race: 4-6-year-olds' inferences differ by speaker cue. *Child Development.* Online early view. 10.1111/cdev.12797

Weatherhead, D., White, K. S., & Friedman, O. (2016). Where are you from ? Preschoolers infer background from accent. *Journal of Experimental Child Psychology*, *143*, 171-178. doi:10.1016/j.jecp.2015.10.011

Wheeler, A., Anzures, G., Quinn, P. C., Pascalis, O., Omrin, D. S., & Lee, K. (2011). Caucasian infants scan own-and other-race faces differently. *PloS one*, *6*(4), e18621.

White, K. S., & Aslin, R. N. (2011). Adaptation to novel accents by toddlers. *Developmental Science*, *14*(2), 372–84. doi:10.1111/j.1467-7687.2010.00986.x

White, K. S., Peperkamp, S., Kirk, C., & Morgan, J. L. (2008). Rapid acquisition of phonological alternations by infants. *Cognition*, *107*(1), 238-265.

Werker, J. F., & Tees, R. C. (1984). Cross-language speech perception: Evidence for perceptual reorganization during the first year of life. *Infant Behavior and Development*, 7, 49–63.