AMERICAN PSYCHOLOGICAL ASSOCIATION

Developmental Psychology

ISSN: 0012-1649

2021, Vol. 57, No. 8, 1195–1209
https://doi.org/10.1037/dev0001032

Toddlers Link Social and Speech Variation During Word Learning

Drew Weatherhead[1] and Katherine S. White[2]
[1] Department of Psychology and Neuroscience, Dalhousie University
[2] Department of Psychology, University of Waterloo

Within a language, there is considerable variation in the pronunciations of words owing to social factors like age, gender, nationality, and race. In the present study, we investigate whether toddlers link social and linguistic variation during word learning. In Experiment 1, 24- to 26-month-old toddlers were exposed to two talkers whose front vowels differed systematically. One talker trained them on a word-referent mapping. At test, toddlers saw the trained object and a novel object; they heard a single novel label from both talkers. Toddlers responded differently to the label as a function of talker. The following experiments demonstrate that toddlers generalize specific pronunciations across speakers of the same race (Experiment 2), but not across speakers who are simply an unfamiliar race (Experiment 3). They also generalize pronunciations based on previous affiliative behavior (Experiment 4). When affiliative behavior and race are pitted against each other, toddlers' linguistic interpretations are more influenced by affiliative behavior (Experiment 5). These experiments suggest that toddlers attend to and link social and speech variation in their environment.

Keywords: accent processing, sociolinguistic development, speech perception, word learning

One of the most impressive accomplishments of early language acquisition is how quickly children acquire new words. Critical to this process is determining the mappings between words and their referents. Although much attention has been paid to the difficulty of determining referents, the acquisition of word-referent mappings is also made challenging by the variability of the word pronunciations themselves. A single referent may be labeled by different wordforms (e.g., the beginning of the word "pasta" has the same vowel as *pod* in the United States, but *pad* in Canada). In many cases, between-individual linguistic variability is systematically linked to social factors, including aspects of the speaker's identity (such as their geographic background or racial background). For example, both Southern American English and African American English include phonological, lexical, and grammatical differences from Standard American English (Fridland et al., 2014; Green, 2002).

How might this variability affect word learning? One possibility is that, when first acquiring words, children acquire only a single form of each word, perhaps the most frequent or salient form in their environment. Alternate forms would be acquired later, along with social information governing their usage. Alternatively, children may track multiple forms of a word from early in development, as well as the social context in which these various forms occur (see Johnson & White, 2020; for discussion of these issues). In the present study, we seek evidence that young children are tracking multiple word variants and linking them to social properties of speakers.

There is considerable evidence that adults track social information about speakers (such as their gender, age, nationality) and link it to variation in the pronunciations of words. This evidence comes from production data showing that speakers modify their use of different variants in different contexts (e.g., Babel, 2010; Bell, 1984; Craig & Grogger, 2012; Delvaux & Soquet, 2007; Eckert, 1988, 1989; Hay et al., 1999; Rickford & McNair-Knox, 1994), as well as perception data showing that providing information about a speaker's nationality or race can impact the way speech is perceived (Babel & Russell, 2015; Campbell-Kibler, 2007; Hay et al., 2010, 2006; Kang & Rubin, 2009; McGowan, 2015; Niedzielski, 1996; Rubin, 1992).

One indication that young children might not be tolerant of multiple forms of the same word is their tendency to assume that different labels map to different objects (Clark, 1990, 1997; Golinkoff et al., 1994; Halberda, 2003; Markman, 1989, 1990; Merriman et al., 1989). However, they appear to relax that assumption and map different labels to the same object when they know that speakers use different languages (Au & Glusman, 1990). This suggests that they understand that labels are conventions that are shared within, and are specific to, a language

Drew Weatherhead https://orcid.org/0000-0002-3649-2030
Katherine S. White https://orcid.org/0000-0003-2631-7711

The authors thank Ashley Blayney-Hoffer, Joel LeForestier, Emily McIntosh, Shaquille Sealy, and Ragini Suresh for help with recruitment and testing, as well as Emily Downs, Harpreet Grewal, Michelle Huh, Jane Klinger, Emily Shiu, Ayah Taji, Jaismine Wadhwa, and Darcy White for their contributions to the project. We also thank the many families who volunteered their time to participate. This work was funded by an operating Grant from the Natural Sciences and Engineering Research Council of Canada to Katherine S. White.

Correspondence concerning this article should be addressed to Drew Weatherhead, Department of Psychology and Neuroscience, Dalhousie University, 1355 Oxford Street, Halifax, NS B3H 4R2, Canada. Email: drew.weatherhead@dal.ca

community (Clark, 1992, 1997). Thus, when they have evidence that people belong to different language communities, they can use this information to selectively accept alternative labels.

Similarly, there is some evidence that by the age of 4 years, children may use multiple forms in their own speech. However, they do not always do so according to the adult patterns of usage. This is particularly true when the use of different forms is conditioned by social factors (Roberts, 1997; Smith et al., 2009), although there is some evidence that children's productions are sensitive to social context (e.g., Miller, 2013; Smith et al., 2007). But it is unclear just how early children begin to track multiple pronunciation variants in their perceptual input or what their expectations are about what kinds of social information condition this type of linguistic variation. That is, what types of information do children use to infer that speakers are part of the same linguistic community?

Artificial language experiments, in which both the type of linguistic variation and conditioning context can be precisely manipulated, have demonstrated that adults prioritize particular social cues (such as gender/sex) over others (such as age and race) when learning novel patterns of linguistic variation (Rácz et al., 2020). But adults have considerably more experience with social and linguistic variation than children, and their biases may reflect this prior experience. What little research exists using these paradigms with children has shown that by the age of 6 years, children can learn some patterns of variation—for example, the fact that the usage of a specific particle differs by speaker gender (Samara et al., 2017). However, they do not always extract the same patterns that adults do (Hudson Kam, 2015).

In the present study, we begin to examine how the links between linguistic and social variation are formed at the earliest stages of word learning. To do so, we conduct a simplified artificial language study to ask whether toddlers assume a link between pronunciation variation and two types of social information about speakers that they have been shown to be highly sensitive to: race and affiliative behavior.

Race

Infants are sensitive to race from a very early age. By 3 months, infants prefer to attend to same-race faces over other-race faces (Bar-Haim et al., 2006; Kelly et al., 2005). Throughout the course of their first year of life, infants become less capable of discriminating or recognizing other-race faces; at 9 months, infants categorize faces by race and are better at recognizing individual same-race faces than individual other-race faces (Anzures et al., 2010; Kelly et al., 2007, 2009) and they display different scanning patterns for same-race and other-race faces (Wheeler et al., 2011). These changes reflect the limited experience that infants in these studies had with other-race faces. For example, the majority of mono-racial infants' interactions are with people from the same mono-racial background as themselves (e.g., Rennels & Davis, 2008; Sugden et al., 2014).

There is some evidence that infants use race as a linguistic marker. For example, six-month-olds match familiar race faces to a familiar language and other-race faces to a novel language (Uttley et al., 2013). A more recent study suggests that 11-month-old White infants who have experience with individuals of Chinese origin associate both English and a Chinese language with an East Asian face, but do not match an unfamiliar language (Spanish) to an East Asian face (May et al., 2019), demonstrating that familiarity with a linguistic community can influence infants' race-language associations.

To date, only one study has investigated the role of race in infants' language processing at the word level. This study demonstrated that a speaker's race influenced 16-month-olds' recognition of familiar and unfamiliar pronunciations (e.g., "dog" vs. "dag") of familiar words (Weatherhead & White, 2018). Infants initially recognized only familiar pronunciations from a familiar-race speaker (and not the unfamiliar pronunciations). For the unfamiliar-race speaker, infants did not initially recognize either type of pronunciation (e.g., "dog" or "dag"), perhaps because they had no prior expectations about how the speaker should produce words. However, they rapidly learned to recognize both types of pronunciations from the unfamiliar-race speaker. This suggests that, by 16 months, infants have formed links between familiar-race speakers and familiar pronunciations of words, whereas they do not link familiar pronunciations to unfamiliar race speakers. This result is consistent with the possibility that infants and toddlers are tracking social information (in this case race) during word learning.

Affiliative Behavior

Infants use people's previous behavior to make inferences about their social affiliations. For example, infants use information about shared ritualistic behavior, shared evaluations or preferences, and previous independent interactions to make inferences about the social relationship between two individuals (e.g., Liberman et al., 2014, 2018; Powell & Spelke, 2013; Spokes & Spelke, 2017).

Infants also have general assumptions about the behavior of communicative partners. For example, six-month-old infants recognize that speech should be directed toward another person and not an inanimate object (Molina et al., 2004), and by 10 months infants expect conversation partners to look at one another (Beier & Spelke, 2012).

Beyond these general assumptions about communicative behavior, infants also appear to have more specific expectations about the relation between language and social behavior. Work in this area has focused on whether infants have behavioral expectations for individuals who speak the same language. For example, nine-month-olds look longer when two individuals who both speak the infants' native language have a negative interaction than when they have a positive interaction (Liberman et al., 2017). And 14-month-olds generalize food preferences across individuals who speak the same language but not across individuals who speak different languages (Liberman et al., 2016). This work suggests that infants expect individuals with similar linguistic patterns to have positive interactions and share the same conventional behavior. However, no previous work has asked the reverse, whether social relationships between speakers influence children's interpretation of language.

In addition to examining toddlers' attention to these two social properties independently, we also ask whether they treat one as more relevant for language processing than the other. The social relevance of particular speaker properties has been shown to affect word learning in adults (Rácz et al., 2020). What do toddlers do when race and affiliation information conflict? One could argue that affiliation is a stronger cue to a speaker's linguistic community, as it entails familiarity between individuals. Do toddlers

prioritize affiliative behavior over race when interpreting pronunciation variation?

The Current Study

The aforementioned work suggests that infants and toddlers can track linguistic information in their environment, and that they are sensitive to social information, namely race and affiliation, as a delineator of social groups. In the current study, we ask whether toddlers (a) track multiple versions of the same words and (b) use race and affiliation as markers of which form a speaker will use. We focus on 24- to 26-month-olds because of their facility for word learning (e.g., Borgströmet al., 2015) and their ability to accommodate unfamiliar pronunciations of novel words after receiving exposure to the relevant linguistic variation (Schmale et al., 2011, 2012).

To investigate these questions, we introduced toddlers to two speakers with different patterns of word pronunciation, or accents (Weatherhead & White, 2016), and then asked how they interpreted the speech of a third individual. Following the procedure used in Weatherhead & White (2016), toddlers were presented with two talkers whose productions systematically differed in the height of their front vowels—a "Training" Speaker and an "Extension" Speaker. The Extension Speaker's pronunciations were always shifted relative to the Training Speaker's. For example, if the Training Speaker said *dɛlu*, the Extension Speaker said *dɪlu*. Toddlers heard this shift across two different front vowels (/ɛ/ to /ɪ/ and /ɪ/ to /i/) to ensure they learned the general phonetic characteristics of the difference between the accents, and not just a rule about a particular vowel (Maye et al., 2008). This type of vowel shift reflects real world variation; for example, the Canadian Vowel shift entails lowering and retraction of the three short front vowels: /ɪʔ/ (kit), /ɛ/ (dress), and /æ/ (trap; e.g., Boberg, 2019).

After some exposure to the differences in their accents, toddlers learned the label for a novel object from the Training Speaker (*tɛpu*) but did not hear the Extension Speaker label it. Thus, toddlers did not directly hear the Extension Speaker's label for the object, although if they were able to track the systematic difference in the accents, they could infer the Extension Speaker would label it *tɪpu*. At test, toddlers saw the trained object from earlier, and a new untrained object, and heard each talker use the label *tɪpu*. If toddlers successfully track the differences across the talkers, their interpretation of the test label *tɪpu* should differ as a function of the talker. That is, for the Training Speaker, toddlers should interpret *tɪpu* as referring to the untrained object, as she previously labeled the trained object *tɛpu.* Thus, we expect a disambiguation response wherein a phonetically different label is interpreted as a new label referring to a new object. For the Extension Speaker, toddlers should interpret *tɪpu* as referring to the trained object, as they should have inferred that *tɪpu* is her pronunciation of the word *tɛpu*, which previously referred to the trained object. Thus, toddlers should infer that the same wordform refers to different objects, depending on which speaker said it.

A disambiguation response is expected in response to the Training Speaker despite research demonstrating that toddlers and young children sometimes interpret familiar words differing in one phonological feature as referring to the familiar objects (e.g., Creel, 2012; Swingley, 2016; White & Morgan, 2008). We make this prediction for two reasons. First, other work has found that toddlers do reject close vowel mispronunciations as familiar labels (Mani & Plunkett, 2007, 2008, 2011; Weatherhead & White, 2018; White & Aslin, 2011). Second, in our methodology, toddlers are tested on a newly trained word for which they have heard numerous repetitions of the Training Speaker's pronunciation. Thus, when toddlers hear the Training Speaker utter *tɪpu* we anticipate that they will identify the vowel change as signaling a new word.

In Experiment 2, we determine whether toddlers extend specific accent properties to an individual who is the same race. To accomplish this, we introduced toddlers again to productions from a Training Speaker (White) and an Extension Speaker (South Asian). However, we then introduced a new silent individual (the Generalization Speaker) who was the same race as the Extension Speaker. At test, productions were heard from all three speakers. If toddlers generalize specific accent properties to individuals of the same race, then their interpretation of the Generalization Speaker's *tɪpu* should be the same as their interpretation for the Extension Speaker (i.e., the trained object). In Experiment 3, we address the specificity of the social information toddlers track about speakers by again introducing a Generalization Speaker. However, this time she differed in race from both the Training Speaker and Extension Speaker (she was East Asian). If toddlers infer that all outgroup members use the same pronunciation conventions, then their interpretation of the Generalization Speaker's *tɪpu* should be the same as their interpretation for the Extension Speaker (i.e., the trained object). However, if they are instead making more specific links between race and pronunciation, then their interpretation of the Generalization Speaker's *tɪpu* should not match either the Training or the Extension Speaker's.

In Experiment 4, we address whether toddlers link more abstract social information to pronunciation variation. In this experiment, both the Training and Extension speaker were White, and the Generalization Speaker (also White) affiliated with either the Training Speaker or the Extension Speaker (between subjects). If toddlers link abstract cues to group membership to pronunciation variation across words, then their interpretation of the Generalization Speaker's *tɪpu* should differ as a function of which speaker she affiliates with. That is, if she affiliates with the Training Speaker, *tɪpu* should refer to the untrained object, but if she affiliates with the Extension Speaker, *tɪpu* should refer to the trained object. Finally, in Experiment 5, we pitted race and affiliation against one another to determine which social cue toddlers more strongly consider. In this experiment, the Training Speaker was White and the Extension Speaker was South Asian, and the Generalization Speaker (also South Asian) affiliated with the Training Speaker. If toddlers prioritize affiliation, they should interpret *tɪpu* the same as they did for the Training Speaker (i.e., the untrained object). But if toddlers prioritize race, they should interpret *tɪpu* the same as they did for the Extension Speaker (i.e., the trained object).

Experiment 1: Can Toddlers Track Pronunciation Variation Across Talkers?

Method

Participants

Forty 24- to 26-month-olds were tested (16 females and 24 males; mean age: 749 days; age range: 724–795 days). This first

Table 1
Audio Stimuli Used During Exposure and Test

Word type	Training speaker	Extension speaker
Exposure pair 1	mɪto	mito
Exposure pair 2	dɛlu	dɪlu
Exposure pair 3	bɪmo	bimo
Object presentation event	tɛpu	
Test word	tɪpu	

experiment included a larger sample size to have adequate power to assess whether there were any effects of speaker order at test (as in Weatherhead & White, 2016). Three additional participants were tested, but not included, because of failure to attend to the screen during the postnaming period of test trials (2) or failure to attend to both objects during the baseline period of test trials (1). In the Waterloo, Canada region, 85% of residents are from White European backgrounds; East Asian and South Asian individuals account for the largest visible minority groups. Additionally, 75% of residents identify English as their mother tongue, 1% report French and 24% report another language. Toddlers in all five experiments were White and were monolingual English learners as indicated by parental report during the experimental session. All experiments received ethics approval from the Office of Research Ethics, at the University of Waterloo (protocol number: 16660, title: Specificity and flexibility in early word recognition).

In all five experiments, we required 20 participants for each condition (Experiments 1 and 4 had two conditions). Testing stopped after 20 participants met the language exposure criteria and showed sufficient attention during the familiarization period and test trials. This stopping criterion was selected based on previous work using the intermodal-preferential looking procedure to examine children's recognition of familiar words produced in familiar and unfamiliar accents (e.g., Weatherhead & White, 2018, which found medium-sized effects for paired sample tests, and medium to large effects for one-sample tests). A posthoc power analysis of Experiment 1 revealed that a sample size of 19 is needed to detect the main effect between the Training and Extension speaker (Faul et al., 2007, 2009). In Experiment 1, the sample size was doubled to 40 participants to allow for 20 participants in each speaker order at test, because of the order effects found in Weatherhead and White (2016). In Experiment 4, we again doubled the sample size to 40, to allow for 20 participants in each affiliation condition.

Stimuli

Audio Stimuli. The stimuli consisted of four pairs of CVCV (consonant, vowel, consonant, vowel) nonsense words (see Table 1), produced by two female native speakers of English. We created an artificial accent difference by having speakers' pronunciations of the same words differ by a systematic vowel shift. The pronunciations of words in a pair differed only in the first vowel (a front vowel, /ɛ/ as in bed, /ɪ/ as in bid, or /i/ as in bead), while the remaining sounds in each word were consistent across speakers. Three of the word pairs (*m[ɪ/i]to*, *d[ɛ/ɪ]lu*, and *b[ɪ/i]mo*) were presented during exposure without referents (exposure pairs). The Training Speaker also used the word *tɛpu* during exposure to label an object (object presentation event). The last word, *tɪpu*, was heard only at test. Stimuli were recorded in a sound-treated booth at a sampling rate of 44,100 Hz and equated for amplitude in Praat (Boersma, 2001). The audio stimuli for the exposure phase were inserted into the videos described below.

Audiovisual Stimuli (Exposure Phase). Both talkers, 24-year-old White 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,* during which they held two objects, an unfamiliar blue object and an unfamiliar yellow object, one in each hand. Each speaker drew attention to one object (the trained object) by waving it in their hand and looking at it, while the other object remained still in their other hand. In the Training Speaker's object presentation event, she held and waved the target object while labeling it *tɛpu* 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).

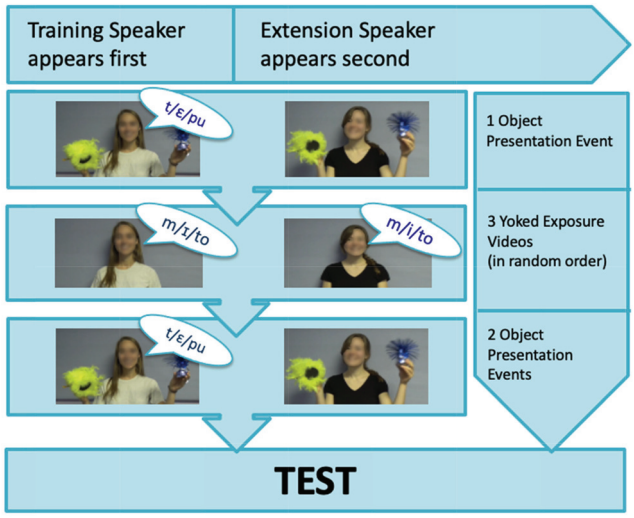
Procedure

The participant sat on his or her parent's lap approximately 1.5 ft. from a 36 × 21-inch plasma screen TV in a sound-treated testing room. A camera under the TV recorded the child's looking behavior. 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 et al., 1993) at approximately 65 dB. 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 (recall, only the training speaker labeled the trained object). Next, the three pairs of yoked exposure videos (e.g., *mɪto-mito*) were presented in random order (see Table 1). These pairs highlighted the vowel difference between the talkers. Finally, the object presentation event pair was presented again twice. In total, toddlers saw the Training Speaker label the trained object 9 times (see Figure 1 for schematic of Exposure Phase). 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 phase. In Experiment 1, there were two test trials, one per talker. Each trial

Figure 1
Schematic of the Exposure Phase for all Experiments



Note. The exposure phase begins with one object presentation event, followed by the three exposure events, followed by two more object presentation events. In each event, the Training Speaker is seen first (approximately 6 seconds), followed by the Extension Speaker (approximately 6 seconds). In each event, the speaker is alone on the screen. We present them together in the figure to highlight the alternation (note that the specific Training and Extension speakers depicted for the object presentation and exposure events are for Experiments 1 and 4). See the online article for the color version of this figure.

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 an additional 8 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 (*tɪpu*; see Figure 2 for schematic of the test phase). 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 learn the Training Speaker's label for the trained object, *tɛpu*, during the exposure phase, then the novel label *tɪpu* should be mapped to the untrained object for this talker. Previous work suggests that toddlers will interpret a familiar label varying by a front vowel (e.g., *dag* instead of *dog*) as a new label belonging to a new object (Weatherhead & White, 2018). Thus, a similar disambiguation response is expected for the Training Speaker. If they also track the pronunciation differences between the speakers, then they should interpret *tɪpu* as the Extension Speaker's pronunciation of the trained object's label and look longer to the trained object for this talker.

Coding of Looking Times

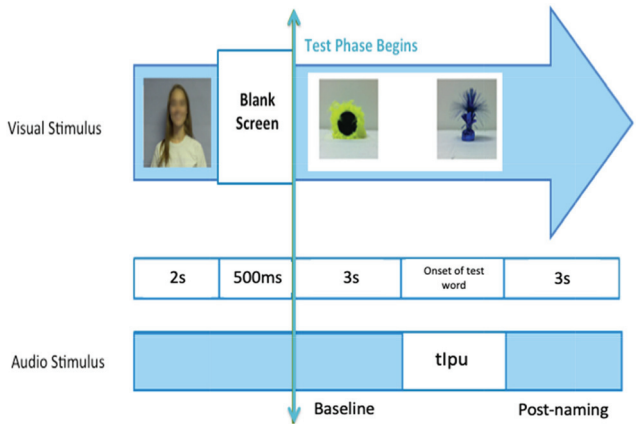
Looking time during the test phase was coded off-line using customized software (James Morgan, Brown University), frame-by-frame (1 frame = 33 ms). Looking proportions for the objects were determined for the baseline period and for the postnaming period, which began 267 ms after test word onset to account for the time necessary to program an eye movement in response to the auditory stimulus (e.g., Swingley & Aslin, 2002). Only the first 3 seconds of the postnaming period were analyzed, to equate the length of the baseline and postnaming periods. Postnaming windows of 3 s and longer have been used in previous studies using looking time measures of toddlers' responses to newly learned words (e.g., Bion et al., 2013; Booth & Waxman, 2009; Houston-Price et al., 2005)

Results and Discussion

For both the baseline and postnaming periods, the proportion of time toddlers looked at each object was computed (out of the total time looking at either object during the three-second period). To assess toddlers' interpretation of the novel test word *tɪpu*, a difference score was calculated for each trial using the looking proportions for each period (proportion trained $\text{object}_{\text{postnaming}}$– proportion trained $\text{object}_{\text{baseline}}$). This measure was selected because it indicates the change in looking toward the trained object after labeling occurred. Such a measure eliminates concern that toddlers had a preference for a particular object or side of the screen (and has been used in many studies of toddler word recognition; Kalashnikova et al., 2018; Kandhadai et al., 2017; Mather & Plunkett, 2011; White & Aslin, 2011; White & Morgan, 2008). We include the proportions for the baseline phase, the postnaming phase, and the calculated difference scores for all experiments in Table 2.

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, $F(1, 38) = 9.44$, $p = .004$, $\eta^2 = .199$, no main effect of Order, $F(1, 38) = .69$, $p = .411$, and no Speaker × Order interaction, $F(1, 38) = .19$, $p = .664$ (see Figure 3).

Figure 2
Schematic of a Test Trial



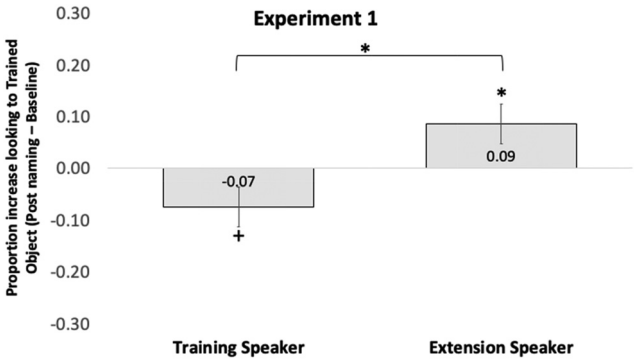
Note. An image of the speaker appears alone on the screen for two seconds, followed by images of the trained and untrained object on either side of the screen. Objects are onscreen for 3 seconds before the test word is uttered (baseline period) and remain on screen for another 3 seconds (postnaming period). There was one test trial for each speaker, for a total of two test trials in Experiment 1, and three test trials in Experiments 2–5. See the online article for the color version of this figure.

Table 2
Means and Standard Deviations of Proportion Looking to the Trained Object for the Baseline Phase and Postnaming Phase, and Postnaming − Baseline Difference Scores, for Each Experiment

Experiment	Baseline	Postnaming	Difference score
Experiment 1			
Training speaker	.55 (.14)	.48 (.24)	−.07 (.24)
Extension speaker	.53 (.17)	.61 (.24)	.09 (.24)
Experiment 2			
Training speaker	.54 (.16)	.38 (.20)	−.17 (.24)
Extension speaker	.49 (.20)	.67 (.17)	.18 (.29)
Generalization speaker	.51 (.13)	.62 (.17)	.11 (.17)
Experiment 3			
Training speaker	.49 (.17)	.36 (.29)	−.13 (.31)
Extension speaker	.45 (.16)	.56 (.18)	.11 (.29)
Generalization speaker	.54 (.14)	.51 (.20)	−.03 (.31)
Experiment 4			
Affiliation to Training speaker			
Training speaker	.54 (.17)	.41 (.20)	−.13 (.16)
Extension speaker	.47 (.17)	.64 (.26)	.17 (.33)
Generalization speaker	.51 (.10)	.37 (.13)	−.14 (.18)
Affiliation to Extension speaker			
Training speaker	.56 (.15)	.45 (.27)	−.11 (.34)
Extension speaker	.45 (.17)	.63 (.21)	.18 (.27)
Generalization speaker	.45 (.11)	.61 (.18)	.16 (.24)
Experiment 5			
Training speaker	.50 (.19)	.38 (.20)	−.14 (.12)
Extension speaker	.52 (.18)	.61 (.16)	−.12 (.23)
Generalization speaker	.57 (.11)	.43 (.16)	.09 (.22)

To determine the effect of labeling for each speaker 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 *tɪpu*, looking marginally decreased to the trained object, $t(39) = .97$, $p = .056$, $d = .309$. In contrast, for the Extension Speaker, looking significantly increased to the trained object, $t(39) = 2.25$, $p = .030$, $d = .355$. Thus, just as the 11-month-olds in Weatherhead & White (2016), toddlers increased their looking toward the untrained object when the Training Speaker said *tɪpu*, but increased their looking toward the trained object when the Extension Speaker said *tɪpu*. The lack of an order effect shows sophistication beyond that of the 11-month-old infants in Weatherhead and White (2016), who were only successful in the easier test order in which the Training Speaker appeared first.

Figure 3
Difference Scores and Standard Errors for Experiment 1



Note. Speaker is on the *x* axis. The *y* axis gives the difference score (postnaming − baseline) for the proportion looking to the trained object. Positive scores indicate increased looking to the trained object; negative scores indicate increased looking to the untrained object.
$^{+}p < .1$. $^{*}p < .05$.

This pattern of results demonstrates that toddlers learned the Training Speaker's label for the training object (*tɛpu*) and inferred that the novel test label mapped onto the untrained object for this talker. More importantly, toddlers tracked the linguistic differences across talkers and used this information to infer the Extension Speaker's label for the trained object. It is important to note that Experiment 2 of Weatherhead and White (2016) rules out the possibility that participants simply learned to respond differently to the two speakers (without learning the features of their accents). In that experiment, 11-month-olds responded similarly to the two speakers when the test label had only back vowels. Thus, during training infants had specifically learned that front vowels differed across the speakers.

Experiment 2: Do Toddlers Generalize Pronunciations Based on Race?

Experiment 1 demonstrated that 2-year-olds are able to track the subtle pronunciation differences between two speakers, and later interpreted a test word differently as a function of who produced it. Experiment 2 investigates whether toddlers generalize pronunciation variation to individuals of the same race. To accomplish this, in Experiment 2, the Training Speaker and the Extension Speaker were different races (White and South Asian, respectively). A third speaker was introduced immediately preceding the test trials

(hereafter referred to as the Generalization Speaker); she was the same race as the Extension Speaker. If toddlers generalize pronunciation variants to individuals of the same race, then their interpretation of the Generalization Speaker's *tɪpu* should be the same as that for the Extension Speaker and different than that for the Training Speaker. Such a pattern would suggest that toddlers' representations of the two speakers' utterances are linked to the speakers' race in some way.

Method

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, because of failure to attend to the screen during the entire test phase. Toddlers were White and were monolingual English learners, with very little exposure to racial diversity, as indicated by parental report during the experimental session.

Stimuli

Audio Stimuli. The same four pairs of CVCV nonsense words as in Experiment 1 were produced by two female native speakers of English. An additional female, native-English speaker provided a third token of *tɪpu* to be used for the Generalization Speaker's test trial. The audio stimuli for the exposure phase were inserted into the videos described below.

Audiovisual Stimuli (Exposure Phase). Both talkers were recorded against the same backdrop and wore different colored t-shirts (white and green). The talkers recorded the same exposure videos and object presentation events as in Experiment 1. Critically, the Training Speaker was White and the Extension Speaker was South Asian. Immediately preceding the test trials, a new speaker, the Generalization Speaker, was introduced. The Generalization Speaker stood equidistant from the two other speakers, and waved to the toddler; critically, she was not heard speaking during Exposure, nor did she interact with either the Training or Extension Speaker (see Figure 4). The Generalization Speaker was of the same South Asian origin as the Extension Speaker.

Procedure

The procedure was identical to that of Experiment 1, with the addition of the introduction video described above and a third test trial (there was one test trial for each talker). For all participants, the Generalization Speaker appeared first at test. Whether the Training or Extension Speaker was presented next was counterbalanced across children (i.e., half of the participants had the order Generalization, Training, Extension, and the other half had Generalization, Extension, Training). The side on which the trained object appeared was also counterbalanced across participants.

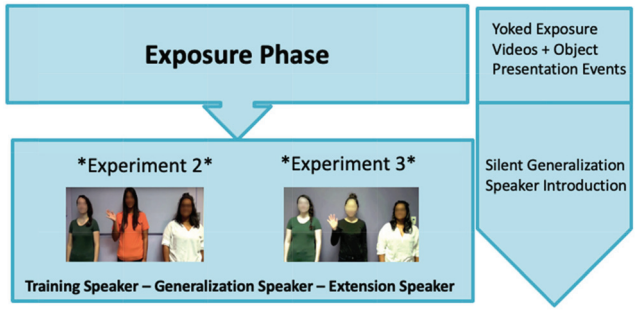
Coding of Looking Times

See Experiment 1.

Results and Discussion

A repeated measures ANOVA on the participants' difference scores with the within-subjects factor of Speaker and between-subjects factor of test Order revealed a main effect of Speaker, $F(2, 18) = 10.46$, $p < .001$, $\eta^2 = .368$, no main effect of Order, $F(2, 18) = .04$, $p = .853$, and no Speaker × Order interaction, $F(2, 18) = .03$, $p = .967$. Paired sample *t*-tests revealed no difference in looking behavior for the Generalization Speaker and the Extension Speaker, $t(19) = .96$, $p = .351$, but significant differences between the Generalization Speaker and the Training Speaker, $t(19) = 4.00$, $p = .001$, $d = .895$, and between the Training Speaker and the Extension Speaker, $t(19) = 3.86$, $p = .001$, $d = .862$ (see Figure 5).

Figure 4
The Silent Introduction to the Generalization Speaker in Experiments 2 and 3



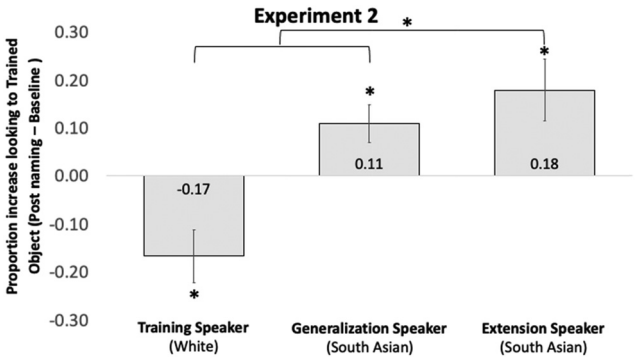
Note. In both experiments, the three speakers appeared on the screen at once, with the Training Speaker (White) to the left, the Extension Speaker (South Asian) to the right, and the Generalization Speaker (South Asian in Experiment 2; East Asian in Experiment 3) in the middle. The Generalization Speaker stood silently and waved to the camera. See the online article for the color version of this figure.

To determine the effect of labeling 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 the novel test word *tɪpu*, $t(19) = -3.05$, $p = .007$, $d = .681$. In contrast, looking significantly increased to the trained object for both the Generalization Speaker, $t(19) = 2.79$, $p = .012$, $d = .624$, and the Extension Speaker, $t(19) = 2.78$, $p = .012$, $d = .623$. Thus, when the Training Speaker said *tɪpu*, toddlers increased their looking toward the untrained object, but when the Extension Speaker and the Generalization Speaker said *tɪpu,* they increased their looking toward the trained object. The fact that toddlers showed the same pattern of looking for the Generalization Speaker and the Extension Speaker suggests that they linked the linguistic variation to race during the familiarization phase and then used this link to determine how to interpret the pronunciation of the new speaker. This finding is also noteworthy because toddlers had not even heard the Extension Speaker's label for the trained object; thus, they not only inferred how the Extension Speaker would pronounce *tɛpu*, they also inferred that a new person of the same race would pronounce it in the same way.

Experiment 3: How Specific Is the Race Information Being Encoded?

Experiment 3 addresses the nature of the information toddlers are tracking about race in this context. The procedure was identical to Experiment 2 with one exception. Rather than being the same race as

Figure 5
Difference Scores and Standard Errors for Experiment 2



Note. Speaker is on the *x* axis. The *y* axis gives the difference score (postnaming − baseline) for the proportion looking to the trained object. Positive scores indicate increased looking to the trained object; negative scores indicate increased looking to the untrained object.
* $p < .05$.

the Extension Speaker, the Generalization Speaker (East Asian) was a different race than the Extension Speaker (she was South Asian).

If toddlers simply think about race in terms of same/different (e.g., familiar-race, or ingroup, speakers talk one way and everyone else talks in some different way), they should interpret the Generalization Speaker's pronunciations in the same way as the Extension Speaker's pronunciations. If, however, the information they are tracking is more specific (i.e., they encode race-specific information), they should not generalize either speaker's accent to the Generalization Speaker, and thus perform at chance for this speaker.

Method

Participants

Twenty 24- to 26-month-olds were tested (11 females; mean age: 758; age range: 728–792 days). One additional participant was tested, but not included, because of failure to complete the task. Toddlers were White and were monolingual English learners, with very little exposure to racial diversity, as indicated by parental report during the experimental session.

Stimuli

Audio Stimuli. See Experiment 2.

Audiovisual Stimuli (Exposure Phase). See Experiment 2. In this experiment, the Training Speaker was again White and the Extension Speaker was of an unfamiliar race (South Asian). Critically, the Generalization Speaker was of an unfamiliar race, but a different race than the Extension Speaker (she was East Asian).

Procedure

The procedure was identical to that of Experiment 2. Based on Experiments 1 and 2, toddlers should map the novel label *tɪpu* 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 (familiar race) or outgroup (unfamiliar race) members, then they should extend the Extension Speaker's accent to the Generalization Speaker. If toddlers are tracking more specific socioindexical information about the speakers, they should not extend either speaker's pronunciations to the Generalization Speaker, and thus perform at chance levels for this speaker.

Coding of Looking Times

See Experiment 1.

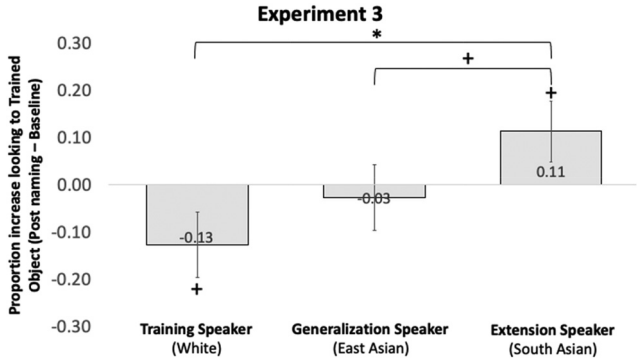
Results and Discussion

A repeated measures ANOVA on the participants' difference scores with the within-subjects factor of Speaker and between-subjects factor of test Order revealed a main effect of Speaker, $F(2, 17) = 3.62$, $p = .037$, $\eta^2 = .167$, no main effect of Order, $F(2, 17) = .16$, $p = .696$, and no Speaker × Order interaction, $F(2, 17) = .12$, $p = .883$. Paired sample *t*-tests revealed a significant difference in looking behavior between the Training Speaker and the Extension Speaker, $t(19) = 3.27$, $p = .004$, $d = .728$, and a marginal difference between the Generalization Speaker and the Extension Speaker, $t(19) = 1.95$, $p = .065$, $d = .436$. There was no difference in looking behavior for the Generalization Speaker and the Training Speaker, $t(19) = .89$, $p = .386$ (see Figure 6).

To determine the effect of labeling 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 decreased to the trained object when the Training Speaker said the novel test word *tɪpu,* $t(19) = 1.84$, $p = .082$, $d = .410$, and increased to the trained object for the Extension Speaker, $t(19) = 1.75$, $p = .096$, $d = .392$, though both effects were marginal. For the Generalization Speaker, toddlers' change in looking did not differ from chance, $t(19) = .41$, $p = .689$. Thus, toddlers did not increase their looking to the Trained Object when the Generalization Speaker said *tɪpu*, unlike in the previous experiment, suggesting they were unsure which object to direct their attention to.

These findings suggest that toddlers are not simply tracking race in terms of whether speakers are of a familiar or unfamiliar race.

Figure 6
Difference Scores and Standard Errors for Experiment 3



Note. Speaker is on the *x* axis. The *y* axis gives the difference score (postnaming − baseline) for the proportion looking to the trained object. Positive scores indicate increased looking to the trained object; negative scores indicate increased looking to the untrained object.
+ $p < .1$. * $p < .05$.

Rather, it appears that toddlers track specific information about race, and this specific social information is linked to linguistic variation.

Experiment 4: Do Toddlers Generalize Pronunciations Based on Previous Affiliative Behavior?

The previous two experiments demonstrate that toddlers link race to linguistic variation at the word level. The motivation for Experiment 4 is to determine whether more abstract cues, like affiliative behavior, influence toddlers' generalizations of pronunciation variants. Experiment 4 uses the same general methods from Experiments 2 and 3, with the exception that the third speaker (the Generalization Speaker) was the same race as both speakers, but chose to affiliate with, and wore the same t-shirt color as, either the Training Speaker or the Extension Speaker (between subjects). If social group membership influences toddlers' assumptions about shared pronunciations, and if affiliation is a cue to group membership, then they should interpret the Generalization Speaker's test word *tɪpu* as a function of who she affiliated with previously.

Method

Participants

Forty 24- to 26-month-olds were tested (18 females and 22 males; mean age: 763 days; age range: 730–786 days). We doubled the sample size in this experiment to have 20 participants in each affiliation condition, as we were uncertain whether toddlers would succeed in both of these conditions. Four additional participants were tested, but not included, due to failure to attend to the screen during the postnaming period of test trials. Participants were randomly assigned to one of two affiliation conditions (20 per condition). Toddlers were White and were monolingual English learners as indicated by parental report during the experimental session.

Stimuli

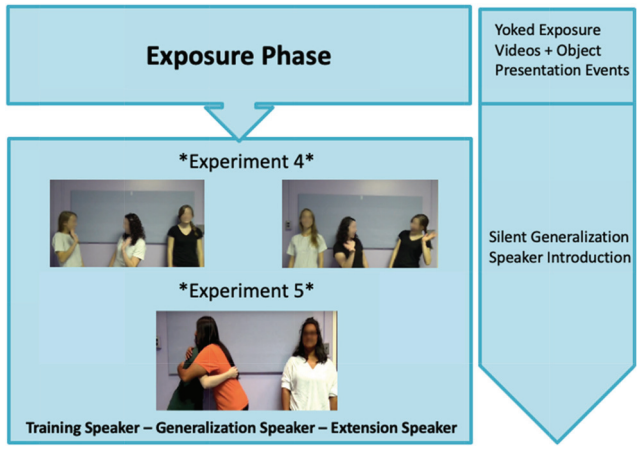
Audio Stimuli. See Experiment 2.

Audiovisual Stimuli (Exposure Phase). See Experiment 2. All speakers were White. An Affiliation introduction video was added to the end of the familiarization phase. For half the participants, the Generalization Speaker wore the same color t-shirt as the Training Speaker; she looked at both speakers, and then silently waved at the Training Speaker in a highly excited manner, and the two embraced for a hug which lasted approximately two seconds (see Figure 7). For the other half of the participants, the Generalization Speaker wore the same color t-shirt as the Extension Speaker; again, she looked at both speakers, and then silently waved at the Extension Speaker, and the two embraced. T-shirt color thus added another perceptual cue to group membership (e.g., Dunham et al., 2011; Patterson & Bigler, 2006).

Procedure

See Experiment 2. Based on the previous experiments, toddlers should map the novel label *tɪpu* 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 affiliative behavior as a cue to social group membership, they should interpret the Generalization Speaker's *tɪpu* differently depending on which speaker she affiliated with. When she has affiliated with the Training Speaker, *tɪpu* should be interpreted as referring to the untrained object. However, when she has affiliated with the Extension Speaker, *tɪpu* should be interpreted as referring to the trained object.

Figure 7
The Silent Introduction to the Generalization Speaker in Experiments 4 and 5



Note. Following the Exposure Phase, all three speakers appeared on the screen at once, with the Training speaker to the left, the Extension Speaker to the right, and the Generalization Speaker in the middle. In Experiment 4 all speakers were the same race (White), and in Experiment 5 the Extension Speaker and the Generalization Speaker were the same race (South Asian) whereas the Training Speaker was a different race (White). In both experiments, the Generalization Speaker looked to both speakers and then waved in an excited manner to one of the two speakers, and those two speakers then embraced. See the online article for the color version of this figure.

Coding of Looking Times

See Experiment 1.

Results and Discussion

Overall Analysis

A repeated measures ANOVA on the participants' calculated difference scores with the within-subjects factor of Speaker (Generalization, Training, Extension) and between-subjects factor of condition (affiliation with Training Speaker vs. affiliation with Extension Speaker) revealed a main effect of Speaker, $F(2, 38) = 13.96$, $p < .001$, $\eta^2 = .269$, a main effect of Condition, $F(2, 38) = 4.51$, $p = .04$, $\eta^2 = .106$, and a Speaker × Condition interaction, $F(2, 38) = 3.97$, $p = .023$, $n^2 = .095$. Because of the Speaker × Condition interaction, analyses for each condition were run separately.

Affiliation to the Training Speaker

A repeated-measures ANOVA on participants' postnaming – baseline difference scores with the within-subjects factor of

Speaker and between-subjects factor of test Order revealed a main effect of Speaker, $F(2, 18) = 12.33$, $p < .001$, $\eta^2 = .407$, no main effect of Order, $F(2, 18) = .04$, $p = .485$, and no Speaker × Order interaction, $F(2, 18) = .60$, $p = .557$. Paired sample *t*-tests revealed no difference in looking behavior for the Generalization Speaker and the Training Speaker, $t(19) = .13$, $p = .902$, but significant differences between the Generalization Speaker and the Extension Speaker, $t(19) = 3.71$, $p = .001$, $d = .830$, and the Training Speaker and the Extension Speaker, $t(19) = 4.04$, $p = .001$, $d = .903$ (see Figure 8).

To determine the effect of labeling for each talker separately, one-sample *t*-tests compared difference scores against chance (a difference score of 0). As predicted, looking significantly decreased to the trained object when the Training Speaker said the novel test word *tɪpu*, $t(19) = 3.35$, $p = .003$, $d = .814$. Toddlers also significantly decreased their looking to the trained object when the Generalization Speaker said *tɪpu*, $t(19) = 3.62$, $p = .002$, $d = .757$. Thus, toddlers interpreted words from the Generalization Speaker in the same way as they did for Training Speaker. In contrast, for the Extension Speaker, looking significantly increased to the trained object, $t(19) = 2.32$, $p = .032$, $d = .519$.

Affiliation to the Extension Speaker

A repeated-measures ANOVA on participants' postnaming − baseline difference scores with the within-subjects factor of Speaker and between-subjects factor of test Order revealed a main effect of Speaker, $F(2, 18) = 6.35$, $p = .004$, $\eta^2 = .261$, no main effect of Order, $F(2, 18) = .486$, $p = .495$, and no Speaker × Order interaction, $F(2, 18) = .14$, $p = .874$. Paired sample *t*-tests reveal no difference in looking behavior for the Generalization Speaker and the Extension Speaker, $t(19) = .36$, $p = .723$, but significant differences between the Generalization Speaker and the Training Speaker, $t(19) = 2.94$, $p = .008$, $d = .657$, and the Training Speaker and the Extension Speaker, $t(19) = 3.04$, $p = .007$, $d = .791$ (see Figure 8).

To determine the effect of labeling for each talker separately, one-sample *t*-tests compared difference scores against chance. As predicted, looking significantly increased to the trained object when the Extension Speaker said *tɪpu*, $t(19) = 3.05$, $p = .007$, $d = .677$. It also increased significantly to the trained object when the Generalization Speaker said *tɪpu*, $t(19) = 2.94$, $p = .008$, $d = .656$. In contrast, for the Training Speaker, looking did not significantly change from baseline, $t(19) = 1.40$, $p = .178$. Thus, toddlers interpreted words from the Generalization Speaker in the same way as they did for the Extension Speaker.

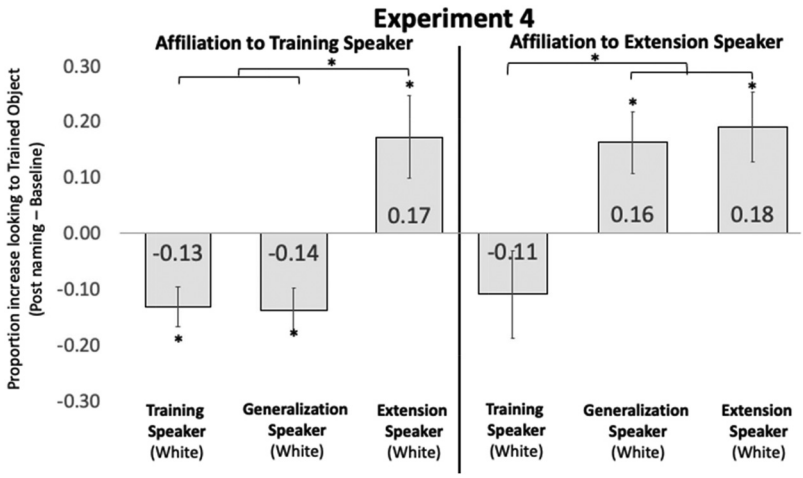
Overall, these results suggest that toddlers tracked linguistic and affiliative behavior during familiarization. Previous affiliation was later the basis for interpreting a pronunciation from a new speaker.

Experiment 5: Do Actions Speak Louder Than Race?

The previous four experiments demonstrate that toddlers link race and previous affiliative behavior to variation in the pronunciations of words. The motivation for Experiment 5 is to determine whether toddlers privilege one type of information over the other. Although race is certainly a salient social property, it is not a reliable cue to linguistic variation. For example, there are many instances in which two speakers with the same racial background use variant forms of words, and in which two speakers of different racial backgrounds use the same forms. A more reliable indicator of linguistic variation might be affiliative behavior, as it suggests familiarity between two speakers. However, this would require toddlers to prioritize affiliation over the highly visually salient property of race.

Experiment 5 uses the same general methods as Experiment 4, with the exception that the third speaker (the Generalization

Figure 8
Difference Scores and Standard Errors for Experiment 4



Note. Speaker is on the *x* axis. The *y* axis gives the difference score (postnaming − baseline) for the proportion looking to the trained object. Positive scores indicate increased looking to the trained object; negative scores indicate increased looking to the untrained object.
* $p < .05$.

Speaker) shared racial characteristics with only the Extension Speaker but affiliated with Training Speaker. If race is a privileged property along which toddlers expect linguistic variation, then they should interpret the Generalization Speaker's *tɪpu* the same way they do for the Extension Speaker. However, if previous affiliative behavior is privileged, then they should interpret the Generalization Speaker's *tɪpu* like the Training Speaker's *tɪpu*.

Method

Participants

Twenty 24- to 26-month-olds were tested (10 females; mean age: 759 days; age range: 737–793 days). Three additional participants were tested, but not included, because of failure to attend to the screen during the postnaming period of test trials (2) and technical errors (1). Toddlers were White and were monolingual English learners, with very little exposure to racial diversity, as indicated by parental report during the experimental session.

Stimuli

Audio Stimuli. See Experiment 2.

Audiovisual Stimuli (Exposure Phase). See Experiment 4 (see Figure 7). In the Affiliation Video, the Generalization Speaker stood between the Training and Extension speaker, looked at both speakers, and then silently waved at the Training Speaker in a highly excited manner, and the two embraced for a hug, which lasted approximately two seconds. The Generalization Speaker and the Extension Speaker were both South Asian, while the Training Speaker was White.

Procedure

Same as Experiments 2–4.

As in the previous experiments, toddlers should map the novel label *tɪpu* to the untrained object for the Training Speaker, and to the trained object for the Extension Speaker. For the Generalization Speaker, toddlers' behavior provides a test of whether race or affiliation is a stronger cue to them about a speaker's potential linguistic pattern. If race is stronger, looking should increase to the trained object, whereas if affiliation is stronger, looking should increase to the untrained object.

Coding of Looking Times

See Experiment 1.

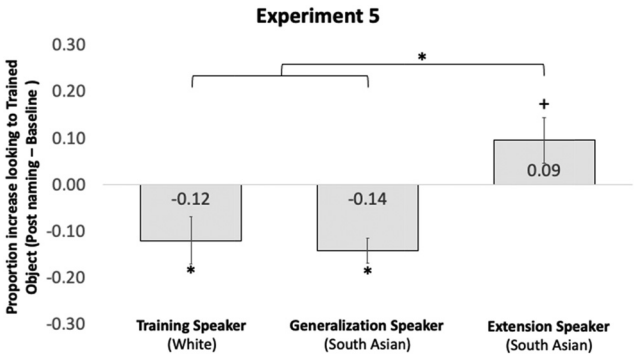
Results and Discussion

A repeated measures ANOVA on participants' test-baseline difference scores with the within-subjects factor of Speaker (Generalization, Training, and Extension), and between-subjects factor of test Order revealed a main effect of Speaker, $F(2, 18) = 10.29$, $p < .001$, $\eta^2 = .364$, and no main effect of Order, $F(2, 18) = .02$, $p = .891$, but a marginal Speaker × Order interaction, $F(2, 18) = 3.23$, $p = .0511$.[1]

Paired sample *t*-tests revealed no difference in looking behavior for the Generalization Speaker and the Training Speaker, $t(19) = .37$, $p = .713$, but significant differences between the Generalization Speaker and the Extension Speaker, $t(19) = 4.96$ $p < .001$, $d = 1.112$, and the Training Speaker and the Extension Speaker, $t(19) = 2.85$, $p = .010$, $d = .740$ (see Figure 9).

Figure 9

Difference Scores and Standard Errors for Experiment 5



Note. Speaker is on the *x* axis. The *y* axis gives the difference score (postnaming − baseline) for the proportion looking to the trained object. Positive scores indicate increased looking to the trained object; negative scores indicate increased looking to the untrained object.

$^{+}$ $p < .1$. * $p < .05$.

To determine the effect of labeling for each talker separately, one-sample *t*-tests compared difference scores against chance (a difference score of 0). As predicted, looking significantly decreased to the trained object when the Training Speaker said the novel test word *tɪpu*, $t(19) = 2.39$, $p = .028$, $d = .536$. Toddlers also significantly decreased their looking to the trained object when the Generalization Speaker said *tɪpu*, $t(19) = 5.23$, $p < .001$, $d = 1.182$. Thus, toddlers interpreted words from the Generalization Speaker in the same way as they interpreted words from Training Speaker. In contrast, for the Extension Speaker, looking marginally increased to the trained object, $t(19) = 1.92$, $p = .070$, $d = .431$. Thus, when previous affiliative behavior and speaker race are pitted against each other, toddlers privilege affiliation.

General Discussion

We encounter a great deal of diversity in our linguistic environments, even within the same language. In some cases, this linguistic variation is attributable to idiosyncratic differences across talkers. In other cases, it can be linked to group-level social factors such as nationality, socioeconomic status, and ethnicity (e.g., Labov, 2006). In five experiments, we demonstrate that two-year-olds track speech variation across talkers and use this information to interpret future pronunciations from those talkers, as well as talkers who appear to be members of the same social group—even if they have never heard these individuals speak previously. The fact that toddlers generalized the learned pronunciations to new individuals suggests that they viewed these pronunciations as group-level conventions, rather than idiosyncratic features of individual speakers. Importantly, these results demonstrate that very early in development, children simultaneously track multiple

[1] The interaction in Experiment 5 was unexpected based on the previous experiments. Given the small sample size in each order, we are reluctant to interpret it further.

forms of the same new word, as well as the social context in which these forms occur.

A variety of studies have shown that children make social judgments and inferences on the basis of accent (e.g., Kinzler et al., 2007; Kinzler & DeJesus, 2013a, 2013b; Liberman et al., 2016; Weatherhead et al., 2016, 2018); suggesting that children treat accent as an indicator of an individual's social group membership. These types of inferences can be quite specific. For example, children infer that speakers with the same accent have similar cultural, but not personal, preferences (Weatherhead et al., 2016). The current study demonstrates the reverse: that social properties of speakers influence toddlers' processing of similar sounding pronunciation variants. These results together suggest that from an early age, children appreciate the links between linguistic and social variation. Critically, these links are nuanced and specific: children infer that some types of social variation are more relevant to a speaker's pronunciations than others, and that pronunciation variation is more relevant to certain types of social information (i.e., cultural preferences vs. personal preferences).

Sociolinguistic Development

This is the one of the first studies to investigate which types of social information influence toddlers' word-referent interpretations. Our results show that salient physical properties of the speaker such as race, as well as abstract social information such as the speaker's previous affiliative interactions, influence which pronunciations are generalized to that speaker.

Experiment 2 demonstrates that toddlers privileged information about race over other visible speaker characteristics; although both speakers were of Southeast Asian origin, there were a number of other salient differences between the Generalization Speaker and the Extension Speaker (such as height, body type, t-shirt color, 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 overspecified. At the same time, in Experiment 3 (in which the Generalization Speaker was a different race from both the Training and Extension Speakers), toddlers did not interpret the Generalization Speaker's pronunciations in the same way as the Extension Speaker's, despite both speakers being from a race unfamiliar to the participants. The fact that toddlers did not increase their looking to either object suggests that they were not making any assumptions about how the novel Generalization Speaker should pronounce the word. This is consistent with the findings of Weatherhead and White (2018), in which younger toddlers initially failed to map words from an unfamiliar race speaker to either presented object. The difference between the Experiment 3 results and those of Experiment 2 suggests that the social information indexed to word representations is neither overspecified nor underspecified, and is specific to the group level, at least in terms of race.

Critically, toddlers in this study had very little exposure to racial diversity, and specifically to individuals of South East Asian and East Asian backgrounds, as suggested by parental report in Experiments 2, 3 and 5. Additionally, the words used by the speakers in this study were novel. Thus, it is not the case that toddlers need previous experience with the specific linguistic and social properties of speakers to learn and generalize this information to new speakers. Our results show that toddlers consider both race and affiliative behavior to be relevant dimensions for interpreting linguistic variability. But when and how do these links develop? One possibility is that infants have innate biases to track certain types of information during speech processing. Another possibility is that learners track all types of salient information as they encounter individuals, and through experience abstract the pertinent social information. The answer to this question is beyond the scope of the present study. The fact that our toddlers had little exposure to other races and accents (as per parent report) makes our findings consistent with the former explanation. At the very least, our results suggest that toddlers do not need contrastive information in the environment to form links between linguistic patterns and these social properties. Moreover, the fact that one type of information (affiliative behavior) was prioritized over the other (race) suggests there is some hierarchical organization in which some types of social information are considered more relevant (to linguistic variation, or more broadly) than others. Whether this organization is innate or learned through experience is an open question for future research.

One concern may be that because toddlers had little exposure to racial diversity, they could have had trouble differentiating speakers of the same race, particularly in Experiments 2 and 5, in which the Generalization and Extension Speakers were both from a South Asian background. However, we believe that this possibility is unlikely because both speakers appeared on screen at the same time, along with the Training Speaker, prior to the test trials. Importantly, they were seen wearing different colored t-shirts, and differed on a number of characteristics such as height, weight, and hair style. Furthermore, toddlers in Experiment 5 interpreted the Generalization Speaker's pronunciations as being consistent with the Training Speaker (who was a different race). Thus, it is unlikely that in Experiment 2 their responses were due to confusion about talker identity. However, it is still possible that, because toddlers had little exposure to racial diversity, they were unsure how to interpret the Generalization Speaker in Experiment 3. Indeed, one way to interpret the null result is that toddlers were uncertain how to classify the East Asian speaker, and that this is why they did not generalize the Training or the Extension Speaker's pronunciations to the Generalization Speaker. Even so, the difference in results across Experiments 2 and 3 demonstrate that there is some level of specificity in how toddlers link language variation to race.

One limitation of the current study is that the races explored were White, South Asian, and East Asian, and that participants were from white monolingual backgrounds. Future work using more racially and linguistically diverse participants and stimuli are necessary to develop a fully develop a full understanding of how these sociolinguistic links develop. Additionally, the present study provided limited linguistic information about the Training and Extension speakers. How speaker race influences toddlers' comprehension of fluent speech in familiar and unfamiliar accents is an open question for future research.

Overall, like adults, the potential relevance of the social information appears to influence the degree to which toddlers weigh it during learning (Rácz et al., 2020). The fact that affiliative behavior outweighed other more visibly salient social information demonstrates that, at least early in development, affiliation is a strong linguistic marker. We speculate that this may be because affiliation

indicates a communicative relationship between two individuals (Liberman et al., 2016).

Toddlers' Understanding of the Conventionality of Accent

Language has communicative power because people assume that the mappings between words and meanings will be consistent among individuals who share the same knowledge (Clark, 1990, 2007; Kalish & Sabbagh, 2007). Older children appear to appreciate this conventionality of language (Clark, 1990; Clark & Clark, 1979), expecting that word meanings are limited to individuals who share the same linguistic conventions or who share the same knowledge (Diesendruck, 2005). The current study demonstrates that toddlers appear to recognize that individuals within a social group adhere to a certain pronunciation for a lexical item, while individuals outside the group may use other pronunciations for that very same item.

In other words, the current study suggests that toddlers recognize that, *even within the same language*, there can be variations in pronunciation across social groups. Because the pronunciations used in this study were systematically related and not distinct words (as have been used in previous research), they more closely resemble real world *accent* variation, rather than language differences. Our findings may also have implications for other work investigating children's recognition of accented words. Toddlers and children often fail to recognize words produced in unfamiliar accents, at least without exposure (e.g., Bent, 2014; Mulak et al., 2013; Nathan et al., 1998; Newton & Ridgway, 2016; van Heugten et al., 2015; White & Aslin, 2011). One possible contributor to children's difficulty may be the assumption that a new speaker is a member of their own social group. Our results suggest that children may show greater flexibility in processing words in novel accents if they are given some indication that the speaker is from a different social group.

Conclusion

Within a language, there is considerable variation in the pronunciations of words, which co-occurs with variation in social factors like gender, nationality, and race. A critical question for language acquisition is when learners begin to track this variability in their linguistic environment and determine the factors governing its occurrence. We demonstrate for the first time that social information influences two-year-olds' word learning and processing. Importantly we demonstrate that toddlers treat both race and affiliative behavior as socially relevant cues to linguistic variation.

References

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. https://doi.org/10.1111/j.1467-7687.2009.00900.x

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. https://doi.org/10.2307/1130757

Babel, M. (2010). Dialect divergence and convergence in New Zealand English. *Language in Society*, *39*(4), 437–456. https://doi.org/10.1017/S0047404510000400

Babel, M., & Russell, J. (2015). Expectations and speech intelligibility. *The Journal of the Acoustical Society of America*, *137*(5), 2823–2833. https://doi.org/10.1121/1.4919317

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. https://doi.org/10.1111/j.1467-9280.2006.01679.x

Beier, J. S., & Spelke, E. S. (2012). Infants' developing understanding of social gaze. *Child Development*, *83*(2), 486–496. https://doi.org/10.1111/j.1467-8624.2011.01702.x

Bell, A. (1984). Language style as audience design. *Language in Society*, *13*(2), 145–204. https://doi.org/10.1017/S004740450001037X

Bent, T. (2014). Children's perception of foreign-accented words. *Journal of Child Language*, *41*(6), 1334–1355. https://doi.org/10.1017/S0305000913000457

Bion, R. A., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word–object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*, *126*(1), 39–53. https://doi.org/10.1016/j.cognition.2012.08.008

Boberg, C. (2019). A closer look at the short front vowel shift in Canada. *Journal of English Linguistics*, *47*(2), 91–119. https://doi.org/10.1177/0075424219831353

Boersma, P. (2001). Praat: A system for doing phonetics by computer. *Glot International*, *5*(9/10), 341–345.

Booth, A. E., & Waxman, S. R. (2009). A horse of a different color: Specifying with precision infants' mappings of novel nouns and adjectives. *Child Development*, *80*(1), 15–22. https://doi.org/10.1111/j.1467-8624.2008.01242.x

Borgström, K., von Koss Torkildsen, J., & Lindgren, M. (2015). Substantial gains in word learning ability between 20 and 24 months: A longitudinal ERP study. *Brain and Language*, *149*, 33–45. https://doi.org/10.1016/j.bandl.2015.07.002

Campbell-Kibler, K. (2007). Accent, (ING), and the social logic of listener perceptions. *American Speech*, *82*(1), 32–64. https://doi.org/10.1215/00031283-2007-002

Clark, E. V. (1992). Conventionality and contrast: Pragmatic principles with lexical consequences. In A. Lehrer & E. F. Kittay (Eds.), *Frames, fields, and contrasts: New essays in semantic and lexical organization* (pp. 171–188). LEA Publishers.

Clark, E. V. (2007). Conventionality and contrast in language and language acquisition. *New Directions for Child and Adolescent Development*, *115*, 11–23. https://doi.org/10.1002/cd.179

Clark, E. V. (1990). On the pragmatics of contrast. *Journal of Child Language*, *17*(2), 417–431. https://doi.org/10.1017/S0305000900013842

Clark, E. V. (1997). Conceptual perspective and lexical choice in acquisition. *Cognition*, *64*(1), 1–37. https://doi.org/10.1016/S0010-0277(97)00010-3

Clark, E. V., & Clark, H. H. (1979). When nouns surface as verbs. *Language*, *55*(4), 767–811. https://doi.org/10.2307/412745

Cohen, J. D., MacWhinney, B., Flatt, M., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments *Behavior Research Methods, Instruments, & Computers*, *25*(2), 257–271. https://doi.org/10.3758/BF03204507

Craig, H. K., & Grogger, J. T. (2012). Influences of social and style variables on adult usage of African American English features. *Journal of Speech, Language, and Hearing Research*, *55*(5), 1274–1288. https://doi.org/10.1044/1092-4388(2012/11-0055)

Creel, S. C. (2012). Phonological similarity and mutual exclusivity: On-line recognition of atypical pronunciations in 3-5-year-olds. *Developmental Science*, *15*(5), 697–713. https://doi.org/10.1111/j.1467-7687.2012.01173.x

Delvaux, V., & Soquet, A. (2007). The influence of ambient speech on adult speech productions through unintentional imitation. *Phonetica*, *64*(2-3), 145–173. https://doi.org/10.1159/000107914

Diesendruck, G. (2005). The principles of conventionality and contrast in word learning: An empirical examination. *Developmental Psychology*, *41*(3), 451–463. https://doi.org/10.1037/0012-1649.41.3.451

Dunham, Y., Baron, A. S., & Carey, S. (2011). Consequences of "minimal" group affiliations in children. *Child Development*, *82*(3), 793–811. https://doi.org/10.1111/j.1467-8624.2011.01577.x

Eckert, P. (1988). Adolescent social structure and the spread of linguistic change. *Language in Society*, *17*(2), 183–207. https://doi.org/10.1017/S0047404500012756

Eckert, P. (1989). The whole woman: Sex and gender differences in variation. *Language Variation and Change*, *1*(3), 245–267. https://doi.org/10.1017/S095439450000017X

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–1160. https://doi.org/10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/BF03193146

Fridland, V., Kendall, T., & Farrington, C. (2014). Durational and spectral differences in American English vowels: Dialect variation within and across regions. *The Journal of the Acoustical Society of America*, *136*(1), 341–349. https://doi.org/10.1121/1.4883599

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*(1), 125–155. https://doi.org/10.1017/S0305000900008692

Green, L. (2002). A descriptive study of African American English: Research in linguistics and education. *International Journal of Qualitative Studies in Education*, *15*(6), 673–690. https://doi.org/10.1080/0951839022000014376

Halberda, J. (2003). The development of a word-learning strategy. *Cognition*, *87*(1), B23–B34. https://doi.org/10.1016/S0010-0277(02)00186-5

Hay, J., Drager, K., & Warren, P. (2010). Short-term exposure to one dialect affects processing of another. *Language and Speech*, *53*(Pt. 4), 447–471. https://doi.org/10.1177/0023830910372489

Hay, J., Jannedy, S., & Mendoza-Denton, N. (1999). Oprah and/ay: Lexical frequency, referee design and style. *Proceedings of the 14th International Congress of Phonetic Sciences* (pp. 1389–1392). University of California.

Hay, J., Nolan, A., & Drager, K. (2006). From fush to feesh: Exemplar priming in speech perception. *Linguistic Review*, *23*(3), 351–379. https://doi.org/10.1515/TLR.2006.014

Houston-Price, C., Plunkett, K., & Harris, P. (2005). 'Word-learning wizardry' at 1;6. *Journal of Child Language*, *32*(1), 175–189. https://doi.org/10.1017/S0305000904006610

Hudson Kam, C. L. (2015). The impact of conditioning variables on the acquisition of variation in adult and child learners. *Language*, *91*(4), 906–937. https://doi.org/10.1353/lan.2015.0051

Johnson, E. K., & White, K. S. (2020). Developmental sociolinguistics: Children's acquisition of language variation. *Wiley Interdisciplinary Reviews: Cognitive Science*, *11*(1), e1515. https://doi.org/10.1002/wcs.1515

Kalashnikova, M., Escudero, P., & Kidd, E. (2018). The development of fast-mapping and novel word retention strategies in monolingual and bilingual infants. *Developmental Science*, *21*(6), e12674. https://doi.org/10.1111/desc.12674

Kalish, C. W., & Sabbagh, M. A. (2007). Conventionality and cognitive development: Learning to think the right way *New Directions for Child and Adolescent Development*, *2007*(115), 1–9. https://doi.org/10.1002/cd.178

Kandhadai, P., Hall, D. G., & Werker, J. F. (2017). Second label learning in bilingual and monolingual infants. *Developmental Science*, *20*(1), e12429. https://doi.org/10.1111/desc.12429

Kang, O., & Rubin, D. L. (2009). Reverse linguistic stereotyping: Measuring the effect of listener expectations on speech evaluation. *Journal of Language and Social Psychology*, *28*(4), 441–456. https://doi.org/10.1177/0261927X09341950

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. https://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. https://doi.org/10.1111/j.1467-9280.2007.02029.x

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. https://doi.org/10.1111/j.1467-7687.2005.0434a.x

Kinzler, K. D., & DeJesus, J. M. (2013a). Children's sociolinguistic evaluations of nice foreigners and mean Americans. *Developmental Psychology*, *49*(4), 655–664. https://doi.org/10.1037/a0028740

Kinzler, K. D., & DeJesus, J. M. (2013b). Northern = smart and Southern = nice: The development of accent attitudes in the United States. *Quarterly Journal of Experimental Psychology*, *66*(6), 1146–1158. https://doi.org/10.1080/17470218.2012.731695

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), 12577–12580. https://doi.org/10.1073/pnas.0705345104

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

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), 966–971. https://doi.org/10.1037/a0034481

Liberman, Z., Kinzler, K. D., & Woodward, A. L. (2018). The early social significance of shared ritual actions. *Cognition*, *171*(2018), 42–51. https://doi.org/10.1016/j.cognition.2017.10.018

Liberman, Z., Woodward, A. L., & Kinzler, K. D. (2017). Preverbal infants infer third-party social relationships based on language. *Cognitive Science*, *41*(Suppl. 3), 622–634. https://doi.org/10.1111/cogs.12403

Liberman, Z., Woodward, A. L., Sullivan, K. R., & Kinzler, K. D. (2016). Early emerging system for reasoning about the social nature of food. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(34), 9480–9485. https://doi.org/10.1073/pnas.1605456113

Mani, N., & Plunkett, K. (2007). Phonological specificity of vowels and consonants in early lexical representations. *Journal of Memory and Language*, *57*(2), 252–272. https://doi.org/10.1016/j.jml.2007.03.005

Mani, N., & Plunkett, K. (2008). Fourteen-month-olds pay attention to vowels in novel words. *Developmental Science*, *11*(1), 53–59. https://doi.org/10.1111/j.1467-7687.2007.00645.x

Mani, N., & Plunkett, K. (2011). Does size matter? Subsegmental cues to vowel mispronunciation detection. *Journal of Child Language*, *38*(3), 606–627. https://doi.org/10.1017/S0305000910000243

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. https://doi.org/10.1207/s15516709cog1401_4

Mather, E., & Plunkett, K. (2011). Mutual exclusivity and phonological novelty constrain word learning at 16 months. *Journal of Child Language*, *38*(5), 933–950. https://doi.org/10.1017/S0305000910000401

May, L., Baron, A. S., & Werker, J. F. (2019). Who can speak that language? Eleven-month-old infants have language-dependent expectations regarding speaker ethnicity. *Developmental Psychobiology*, *61*(6), 859–873. https://doi.org/10.1002/dev.21851

Maye, J., Aslin, R. N., & Tanenhaus, M. K. (2008). The weckud wetch of the wast: Lexical adaptation to a novel accent. *Cognitive Science*, *32*(3), 543–562. https://doi.org/10.1080/03640210802035357

McGowan, K. B. (2015). Social expectation improves speech perception in noise. *Language and Speech*, *58*(Pt. 4), 502–521. https://doi.org/10.1177/0023830914565191

Merriman, W. E., Bowman, L. L., & MacWhinney, B. (1989). The mutual exclusivity bias in children's word learning *Monographs of the Society for Research in Child Development*, *54*(3-4), 1–132. https://doi.org/10.2307/1166130

Miller, K. (2013). Acquisition of variable rules:/s/-lenition in the speech of Chilean Spanish-speaking children and their caregivers. *Language Variation and Change*, *25*(3), 311–340. https://doi.org/10.1017/S095439451300015X

Molina, M., Van de Walle, G. A., Condry, K., & Spelke, E. S. (2004). The animate-inanimate distinction in infancy: Developing sensitivity to constraints on human actions. *Journal of Cognition and Development*, *5*(4), 399–426. https://doi.org/10.1207/s15327647jcd0504_1

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), 2064–2078. https://doi.org/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*(2), 343–365. https://doi.org/10.1017/S0305000998003444

Newton, C., & Ridgway, S. (2016). Novel accent perception in typically-developing school-aged children. *Child Language Teaching and Therapy*, *32*(1), 111–123. https://doi.org/10.1177/0265659015578464

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

Patterson, M. M., & Bigler, R. S. (2006). Preschool children's attention to environmental messages about groups: Social categorization and the origins of intergroup bias. *Child Development*, *77*(4), 847–860. https://doi.org/10.1111/j.1467-8624.2006.00906.x

Powell, L. J., & Spelke, E. S. (2013). Preverbal infants expect members of social groups to act alike. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(41), E3965–E3972. https://doi.org/10.1073/pnas.1304326110

Rácz, P., Hay, J. B., & Pierrehumbert, J. B. (2020). Not all indexical cues are equal: Differential sensitivity to dimensions of indexical meaning in an artificial language. *Language Learning*, *70*(3), 848–885. https://doi.org/10.1111/lang.12402

Rennels, J. L., & Davis, R. E. (2008). Facial experience during the first year *Infant Behavior and Development*, *31*(4), 665–678. https://doi.org/10.1016/j.infbeh.2008.04.009

Rickford, J. R., & McNair-Knox, F. (1994). Addressee-and topic-influenced style shift: A quantitative sociolinguistic study. In D. Biber, & E. Finegan (Eds.), *Sociolinguistic perspectives on register variation* (pp. 235–276). Oxford Press.

Roberts, J. (1997). Acquisition of variable rules: A study of (-t, d) deletion in preschool children. *Journal of Child Language*, *24*(2), 351–372. https://doi.org/10.1017/S0305000997003073

Rubin, D. L. (1992). Non-language factors affecting undergraduates' judgments of nonnative English-speaking teaching assistants. *Research in Higher Education*, *33*(4), 511–531. https://doi.org/10.1007/BF00973770

Samara, A., Smith, K., Brown, H., & Wonnacott, E. (2017). Acquiring variation in an artificial language: Children and adults are sensitive to socially conditioned linguistic variation. *Cognitive Psychology*, *94*(2017), 85–114. https://doi.org/10.1016/j.cogpsych.2017.02.004

Schmale, R., Cristia, A., & Seidl, A. (2012). Toddlers recognize words in an unfamiliar accent after brief exposure. *Developmental Science*, *15*(6), 732–738. https://doi.org/10.1111/j.1467-7687.2012.01175.x

Schmale, R., Hollich, G., & Seidl, A. (2011). Contending with foreign accent in early word learning. *Journal of Child Language*, *38*(5), 1096–1108. https://doi.org/10.1017/S0305000910000619

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. https://doi.org/10.1017/S0954394507070044

Smith, J., Durham, M., & Fortune, L. (2009). Universal and dialect-specific pathways of acquisition: Caregivers, children, and t/d deletion. *Language Variation and Change*, *21*(1), 69–95. https://doi.org/10.1017/S0954394509000039

Spokes, A. C., & Spelke, E. S. (2017). The cradle of social knowledge: Infants' reasoning about caregiving and affiliation. *Cognition*, *159*(2017), 102–116. https://doi.org/10.1016/j.cognition.2016.11.008

Sugden, N. A., Mohamed-Ali, M. I., & Moulson, M. C. (2014). I spy with my little eye: Typical, daily exposure to faces documented from a first-person infant perspective. *Developmental Psychobiology*, *56*(2), 249–261. https://doi.org/10.1002/dev.21183

Swingley, D. (2016). Two-year-olds interpret novel phonological neighbors as familiar words. *Developmental Psychology*, *52*(7), 1011–1023. https://doi.org/10.1037/dev0000114

Swingley, D., & Aslin, R. N. (2002). Lexical neighborhoods and the word-form representations of 14-month-olds. *Psychological Science*, *13*(5), 480–484. https://doi.org/10.1111/1467-9280.00485

Uttley, L., de Boisferon, 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. https://doi.org/10.1177/0165025412467583

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. https://doi.org/10.1080/15475441.2013.879636

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. https://doi.org/10.1080/15475441.2015.1024835

Weatherhead, D., & White, K. S. (2018). And then I saw her race: Race-based expectations affect infants' word processing. *Cognition*, *177*(2018), 87–97. https://doi.org/10.1016/j.cognition.2018.04.004

Weatherhead, D., Friedman, O., & White, K. S. (2018). Accent, language, and race: 4-6-year-old children's inferences differ by speaker cue. *Child Development*, *89*(5), 1613–1624. https://doi.org/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*(2016), 171–178. https://doi.org/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. https://doi.org/10.1371/journal.pone.0018621

White, K. S., & Aslin, R. N. (2011). Adaptation to novel accents by toddlers. *Developmental Science*, *14*(2), 372–384. https://doi.org/10.1111/j.1467-7687.2010.00986.x

White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, *59*(1), 114–132. https://doi.org/10.1016/j.jml.2008.03.001

Received May 11, 2020
Revision received February 16, 2021
Accepted March 24, 2021 ■