The Development of Regional Dialect Locality Judgments and Language Attitudes Across the Life Span

Elizabeth A. McCullough
Ohio State University and University of Washington

Cynthia G. Clopper and Laura Wagner
Ohio State University

The development of language attitudes and perception of talker regional background was investigated across the life span (N = 240, age range = 4–75 years). Participants rated 12 talkers on dimensions of geographic locality, status, and solidarity. Children could classify some dialects by locality by age 6–7 years and showed adult-like patterns by age 8 years. Children showed adult-like status ratings for some dialects by age 4–5 years but were not fully adult-like until age 12 years. Solidarity ratings were more variable and did not exhibit a clear developmental trajectory, although some adult-like patterns were in place by age 6–7 years. Locality ratings were a significant but modest predictor of attitude ratings, suggesting that geographic knowledge is one contributor to language attitudes throughout development.

Language attitudes are the beliefs and opinions that individuals hold about different language varieties as well as the people who speak them. They have been studied in both adults and children for decades (e.g., Anisfeld & Lambert, 1964; Jones, 1950), but their developmental trajectory remains poorly understood, especially with respect to how they interact with children’s general appreciation that variation in the speech signal can provide information about where a talker is from. In the literature on adult language attitudes, a widespread assumption is that these attitudes arise from associations between language variation and geographic areas, such as countries or more local regional divisions (Giles & Niedzielski, 1998). That is, listeners know that people from a particular region (e.g., the Southern United States) talk in a particular way, and they have attitudes about those people; a given pronunciation is not intrinsically linked to specific attitudes but becomes linked to these attitudes by being associated in the listener’s mind with the people who produce it (Anisfeld, Bogo, & Lambert, 1962; Lambert, Hodgson, Gardner, & Fillenbaum, 1960; Nesdale & Rooney, 1996). Although other factors also play a role, a complete understanding of a listener’s language attitudes requires knowing which regional varieties are perceptually distinct and how those distinct varieties are linked to geographic areas for the listener (Preston, 1989). However, few studies with adults or children have incorporated direct tests of participants’ knowledge about the talkers’ regional backgrounds alongside tests of language attitudes (cf. Anisfeld & Lambert, 1964; Kinzler & DeJesus, 2013). Thus, although adults’ knowledge about regional language varieties includes both specific linguistic information and attitude information, the proposed cognitive mechanism connecting these different kinds of information has not been explicitly tested. Furthermore, children must learn language attitudes as part of acquiring their adult language system, but it is unclear when, or even if, that competence is connected to their ability to assess a talker’s geographic region of origin. The current work investigated the development of language attitudes across the life span, from ages 4 through 75 years, and also obtained explicit geographic locality judgments from all participants to allow for a direct comparison between geographic categories and language attitudes. The primary emphasis of the study is how children initially develop their competence in this domain, and thus children were densely sampled between the ages of 4 and 17 years. However, as the literature suggests that the process depends in part on learning cultural knowledge, sampling...
from three adult age groups was also included—after all, learning does not stop with adulthood and additional cultural and geographic knowledge can be gained from a full lifetime of experience.

**Adults’ Language Attitudes and Perception of Talkers’ Region of Origin**

Early work on adults’ language attitudes explicitly asked participants what they thought about people who spoke particular varieties (Jones, 1950), but later work shifted the focus to tests in which listeners hear samples of speech and evaluate talkers based on that speech (Lambert, Giles, & Picard, 1975; Lambert et al., 1960). These talker evaluations are then taken to reflect the listeners’ attitudes toward the group each talker represents. Research using this approach has often focused on measures of “status,” which reflect aspects of the social standing of a talker, including wealth, intelligence, success, and importance, as well as measures of “solidarity,” which reflect aspects of interpersonal interactions with the talker, including friendliness, honesty, reliability, and likability (Fuertes, Gottdiener, Martin, Gilbert, & Giles, 2012; Giles & Billings, 2004). This work has consistently shown that listeners rate talkers of more standard and prestigious language varieties as having higher status than talkers of less prestigious varieties. This pattern has been found across languages (e.g., Lambert et al., 1960 for English over French in Canada) as well as for both regional (Cheyne, 1970; Ladegaard, 1998; Levin, Giles, & Garrett, 1994; Luhman, 1990) and ethnic (Arthur, Farrar, & Bradford, 1974; Gallois & Callan, 1989; Nesdale & Roosey, 1990) varieties of English in Australia, the United Kingdom, and the United States. The same results are typically observed regardless of whether the listeners are speakers of a more standard variety or a less standard variety (Cheyne, 1970; Lambert et al., 1960). In a few rare cases, speakers of standard and nonstandard varieties are rated equally on measures of status (e.g., Lambert et al., 1975 for some varieties of American English and French; Nesdale & Roosey, 1990 for some ethnic varieties of Australian English), or the typical pattern is reversed so that nonstandard speakers are rated more positively than standard speakers (i.e., “positive prejudice” in Gallois & Callan, 1989 study on Hispanic Australian immigrants).

By contrast, results pertaining to measures of solidarity are much less consistent. In some cases, speakers of language varieties that receive low status ratings receive high solidarity ratings, reflecting “covert prestige” (Gallois & Callan, 1989; Ladegaard, 1998; Levin et al., 1994; Luhman, 1990), especially if the listeners are speakers of the less prestigious variety (Cheyne, 1970). However, in other cases, speakers of standard and prestigious varieties are favored in ratings of both status and solidarity (Arthur et al., 1974 for standard English over Chicano English; Lambert et al., 1960 for English over French in Canada), or speakers of standard and nonstandard varieties receive comparable solidarity ratings (Lambert et al., 1975 for some varieties of American English and French; Nesdale & Roosey, 1990 for some ethnic varieties of Australian English). Thus, adult listeners reliably distinguish between standard or prestigious varieties and other language varieties on measures of status, but they have more variable attitudes in the solidarity domain.

Across studies, a wide variety of specific attitudes have been examined for both status and solidarity, including very general notions (e.g., “successful” and “general likability”) as well as a range of quite specific qualities (e.g., “wealthy” and “dependable”). It is notable, therefore, that despite the variety among the specific terms used to test each dimension, status judgments tend to be more consistent than solidarity judgments. One potential reason for the difference between status and solidarity judgments is the fact that across the range of specific terms, status and solidarity judgments tap different conceptual elements that implicate different kinds of information. Status concepts are properties of the talker and are more readily verifiable as facts about the world. These facts can be directly linked to group characteristics (e.g., the Northern United States is wealthier than the Southern United States), which are in turn captured by dialect differences between the groups. By contrast, solidarity concepts are properties of the relationship between the talker and the listener, and are therefore more subjective. These subjective attitudes may not be as easily linked to groups or, by extension, language varieties. Thus, the relation between status and solidarity judgments differs depending on the varieties under investigation.

Adults also exhibit explicit knowledge about the relation between language variation and region of origin. For example, Dutch adults classify talkers by their province of origin with above-chance accuracy, British adults successfully classify talkers by their local region of origin (Van Bezooijen & Gooskens, 1999), and Welsh adults perform above chance in regional forced choice classification of Welsh English (Williams, Garrett, & Coupland,
Patterns of variation in dialect classification further suggest that listeners make a perceptual distinction between local and nonlocal dialects (Baker, Eddington, & Nay, 2009; Clopper & Pisoni, 2004, 2006; Preston, 1993). Clopper et al. (2012) explicitly investigated this possibility by asking adult listeners to judge the likelihood that talkers were from Columbus, Ohio, where the experiment was conducted. The participants consistently and correctly judged talkers from two Midwestern dialects (Midland and Northern) as more likely to be from Ohio than talkers from New England and the South, confirming that adults can use information in the speech signal to make explicit judgments about local versus nonlocal dialects.

Given adults’ success in classifying talkers by region of origin, it is reasonable to assume that adults use their knowledge about those regions—including their stereotypes of people from those regions—to inform their attitudes about speakers of those dialects. That is, adults have all the necessary pieces of knowledge in place to support the conventional assumption that their language attitudes about regional dialect are mediated by their knowledge about regions and their stereotypes about the people who live there (Niedzielski & Preston, 2003).

Children’s Language Attitudes and Perception of Talkers’ Region of Origin

Numerous studies have examined children’s language attitudes (Anisfeld & Lambert, 1964; Cremona & Bates, 1977; Day, 1980; Garrett, Coupland, & Williams, 1999; Giles, 1970; Giles, Harrison, Creber, Smith, & Freeman, 1983; Kinzler & DeJesus, 2013; Lambert, Anisfeld, & Yeni-Komshian, 1965; Lambert et al., 1975; Light, Richard, & Bell, 1978; Nesdale & Rooney, 1996), but it is difficult to observe a clear trajectory for how language attitudes develop throughout childhood due to substantial variation in the language varieties, the ranges of ages, and the types of measures used across studies. Nonetheless, a few generalizations emerge. First, even young children use linguistic variation as a basis for social judgments (Nardy, Chevrot, & Barbu, 2013). For example, preschool-aged children choose to be friends with speakers of their own language rather than speakers of a different language and with native rather than non-native speakers of their own language (Creel, 2017; Kinzler, Dupoux, & Spelke, 2007; Kinzler, Shutts, DeJesus, & Spelke, 2009; Souza, Byers-Heinlein, & Poulin-Dubois, 2013). They also preferentially trust information from native rather than non-native speakers (Kinzler, Corriveau, & Harris, 2011; Wagner, Dunfield, & Rohrbeck, 2014). Thus, the idea that language should be the basis for social judgments appears to be in place from an early age.

Second, children of various ages exhibit more positive attitudes about talkers who share their speech variety, regardless of the specific attitude being tested. For example, Hebrew- and Arabic-speaking teenagers in Israel favor talkers of their own cultural group on measures of both status and solidarity (Lambert et al., 1965). Similarly, kindergarteners in Hawaii rate speakers of the local Hawaiian Creole English higher than standard English speakers on measures of status and solidarity (Day, 1980), monolingual French (but not bilingual French–English) 10-year-olds in Montreal rate talkers more favorably on both status and solidarity when they speak French rather than English (Anisfeld & Lambert, 1964), and 6-year-old Italian children favor the local regional variety over standard Italian (Cremona & Bates, 1977). These results suggest that children’s attitudes may, at least initially, depend more on an egocentric judgment (i.e., is this person like me or not like me?) rather than specific geographic or sociocultural knowledge.

Third, as children get older, their attitudes about status move toward the adult state, which favors more standard or prestigious language varieties. For example, in the study by Cremona and Bates (1977), although the youngest children responded that the regional variety talker “spoke better,” this preference progressively shifted so that by 10 years old, children showed a nearly unanimous preference for the standard Italian speaker. Similar shifts toward linking talkers of the more standard language variety to higher status in older children were observed by Day (1980) for standard and Hawaiian Creole English, by Giles (1970) for an assortment of regional British and foreign accents in English, and by Kinzler and DeJesus (2013) for Northern and Southern varieties of American English. Moreover, other studies looking only at
isolated age groups have found that children older than 8–10 years typically make adult-like associations between high-prestige dialects and higher status in Australia, the United Kingdom, and the United States (Garrett et al., 1999; Light et al., 1978; Nesdale & Rooney, 1996).

Fourth, as in the adult literature, developmental data pertaining to solidarity judgments are more variable. Some studies have shown that older children judge speakers of less prestigious varieties more favorably on solidarity measures (Garrett et al., 1999 for Welsh varieties of English; Kinzler & DeJesus, 2013 for Southern American English; Lambert et al., 1975 for varieties of French), but other studies have found the reverse (Anisfeld & Lambert, 1964 for English over Canadian French; Day, 1980 for standard English over Hawaiian Creole English; Light et al., 1978 for educated over uneducated African American English). Presumably children, like adults, use different conceptual bases for their status and solidarity judgments, leading to this variability. Thus, the specific varieties being compared likely matter for children just as they matter for adults, although they may also matter differently across development as children develop more adult-like knowledge. In addition, as with the adult studies, a range of specific status and solidarity attitudes were tested with the children. However, as children are developing their conceptual (and lexical) repertoire at the same time as they are developing their dialect sensitivity, it is also possible that the differences across studies reflect differences in conceptual understanding of the specific terms used to obtain attitude judgments.

In contrast to work on language attitudes, only a few studies have examined children’s ability to classify talkers by regional dialect using explicit reference to actual geographic places. Williams et al. (1999) found that adolescents performed above chance in a forced choice dialect classification task in Wales, although success rates were modest and less accurate than adults’ overall. In the United States, Kinzler and DeJesus (2013) played speech samples produced by talkers from two dialect regions to 5- and 6-year-olds and 9- and 10-year-olds and asked them which of the talkers sounded like they were “from here” (i.e., from the location of testing). Children in the North selected Northern talkers significantly more often than Southern talkers, and the older children exhibited greater accuracy than the younger children. However, children in the South, even those in the older age group, performed at chance. These results suggest that children’s knowledge about regional dialect variation develops with age but that other factors, perhaps including dialect exposure or community attitudes, also influence development. Recent work by Creel (2017) similarly found that 7-year-old American children could correctly classify Dutch-accented talkers as being not “from here,” but she found no correlation between children’s ability to make that locality judgment and the strength of their preference to be friends with a native talker relative to a Dutch-accented talker. Thus, the nature of the connection between children’s locality assessments and their socially oriented language judgments is unclear.

Other studies examining perceptual dialect abilities in 5- and 6-year-olds have relied on the connection between geographic locality and social groupings by asking children to group talkers into teams or families based on dialect (Girard, Floccia, & Goslin, 2008 with regional varieties of French; Wagner, Clopper, & Pate, 2014 with regional and second-language varieties of English). These studies have consistently shown that young children cannot use a regional dialect difference as a basis for grouping talkers, although they succeed when foreign-accented speech is contrasted with the child’s home dialect. Children’s difficulties in these studies may be due to task demands, such as the need to link dialect differences to an ad hoc category structure. In recent work by Jones, Yan, Wagner, and Clopper (2017), 4- through 17-year-olds were asked to group talkers from four regional dialect groups using a free classification task. This method removes the more difficult task demands of previous work and reveals that children as young as 4–5 years old are able to distinguish at least one adult-like regional dialect group.

Thus, the existing literature on children’s dialect classification paints a very underspecified picture. By age 4–5 years, children have the basic ability to group talkers by regional dialect if the task is easy enough, but they do not exhibit adult-like competence until at least adolescence. To the extent that language attitudes are connected to regional dialect classification, children’s extended difficulties in classifying talkers by region of origin should have some impact on the developmental trajectory of their attitudes toward those talkers.

**Experimental Overview**

The goal of the current study was to explore the relation between developmental trajectories in language attitudes and regional dialect perception across the life span using the same regional varieties. This approach allows for a direct comparison
of performance across children and adults of different ages, which a meta-analysis of the existing literature does not permit, given the range of varieties examined across studies. The experiment also involves within-participant language attitude and geographic locality ratings, allowing for an explicit test of the relation between these two kinds of perceptual responses for both children and adults.

Child participants (ages 4–17 years) were densely sampled in this experiment because the primary research question concerns the developmental trajectory over childhood to adult-like locality and language attitude judgments about regional dialects and the potential mediating connections between those judgments. Adult participants (ages 18–75 years) were more sparsely sampled, but sufficient numbers were tested to allow for an examination of lifelong learning in this domain. To the extent that geographic and sociocultural knowledge continues to be learned throughout the life span, changes in how regional dialects are judged as a result of more life experience are expected.

All participants listened to the speech of talkers from four American English dialect regions: Midland, Northern, and Southern talkers, and further, New England talkers as being from a different place than those judgments. Adult participants (ages 18–75 years) were more sparsely sampled, but sufficient numbers were tested to allow for an examination of lifelong learning in this domain. To the extent that geographic and sociocultural knowledge continues to be learned throughout the life span, changes in how regional dialects are judged as a result of more life experience are expected.

Given adults’ overall success in previous dialect classification studies (Clopper & Pisoni, 2004; Van Bezooijen & Gooskens, 1999; Williams et al., 1999), adults’ locality judgments were predicted to veridically reflect the geographic regions where the dialects are spoken. Specifically, Midland and Northern talkers were expected to be judged as more likely to be from Ohio than Southern and, especially, New England talkers. Adults’ attitudes were also expected to be strongly related to their locality judgments, given that listeners’ assessments of a talker’s region of origin are assumed to underpin their language attitudes. However, as noted above, multiple adult groups were tested to allow for the possibility that even adults’ representations and attitudes about geographic regions might change as more experience is accumulated with age.

To the extent that geographic knowledge is a critical mediator of language attitudes, shifts in children’s attitude ratings should parallel shifts in their geographic ratings. Previous work (Jones et al., 2017) showed that children’s geographic classification of these regional dialects took nearly a decade to become adult like, with different dialects showing different trajectories to adult competence. This protracted and uneven course of development across the regional dialects affords the opportunity to find specific links between the improvement of geographic knowledge for a particular dialect and language attitudes for that dialect. However, given that children’s understanding of the concepts encoded by the attitudes may also be developing, there is the further possibility that children will show complex interactions among the specific attitudes, the regional dialects, and the geographical associations to those dialects. To preview the findings, the results suggest a connection between language attitudes and geographic associations with language variation throughout the life span, although mismatches between attitudes and their relation to locality judgments suggest that other factors also contribute to language attitude judgments.

**Method**

**Participants**

Two hundred and forty monolingual native speakers of American English between 4 and 75 years old were recruited from the visitors at a
science center in Columbus, Ohio, in 2014–2015. As shown in Table 1, participants were divided into 10 age groups, with 24 individuals per group. Samples of this size have yielded significant results in closely related work in the past (e.g., Kinzler & DeJesus, 2013). As is typical for participants drawn from this particular pool of science center visitors, the overwhelming majority of the participants were White and from high socioeconomic backgrounds. Table 1 summarizes the residential history of the participants in each of the 10 age groups and shows that the majority of the participants were lifetime residents of the Midwestern United States, including the Midland and Northern dialect regions, a small minority were lifetime residents of the Southern region, and the remaining participants lived in other regions or had lived in multiple dialect regions. None were lifetime residents of the New England region. Across all participants, 76% had lived in Ohio for some period of time. Data from 93 additional listeners were excluded for not meeting the native language requirement (n = 17), difficulty hearing the stimuli (n = 2), technical problems with the experiment software (n = 5), or not completing the task (n = 69). Many of the participants who did not complete the task (n = 46) were 4- and 5-year-olds who were unable to sit attentively through the complete experimental session, opting out early to experience other activities, exhibits, or shows in the science center.

### Stimulus Materials

The stimulus materials were produced by 12 female talkers from the TIMIT corpus (Fisher, Doddington, & Goudie-Marshall, 1986), including three talkers from each of four regions in the United States (Midland, North, South, and New England). The sentence “She had your dark suit in greasy wash water all year” was produced by each of the talkers in the TIMIT corpus, and was selected as the stimulus sentence for the current study. This sentence was designed to reveal dialect-specific variation in pronunciation that may be perceptible to listeners. The selected talkers were identified as representatives of their broad dialect regions, given their production of documented features of these major dialects (Labov et al., 2006) in this target sentence. For instance, all three of the selected Southern talkers said “gready” rather than “greasy,” and all three of the selected New England talkers omitted the postvocalic /r/ in the words “dark” and “year.” Although each of these dialects exhibits subregional variation, characteristics of the broad regions were shared by all three selected talkers within each dialect. Because all of the talkers produced the same sentence, only phonological dialect variation was available to the listeners; no lexical or syntactic differences were present.

The four regional varieties were chosen because they differ from one another in terms of prestige and proximity to the testing location. The Midland dialect is spoken in the region that includes the testing location in Columbus, Ohio, and is high prestige (Niedzielski & Preston, 2003). The Northern and Southern dialects are both spoken in regions close to the testing location but differ in their social evaluation. The Northern dialect is high prestige, and while it is phonetically distinct from the Midland dialect, the differences are not socially marked; when individuals discuss differences between these two varieties, they often ascribe them to individual talker characteristics rather than to the dialect overall (Campbell-Kibler, 2012). The Southern dialect is lower in prestige and is associated with prominent cultural stereotypes (Campbell-Kibler, 2007; Niedzielski & Preston, 2003). The New England dialect is relatively low in prestige (Niedzielski & Preston, 2003) and is spoken in a region that is farther from the testing location. Given their residential histories (see Table 1), the listeners likely have more extensive exposure to the Midland, Northern, and Southern varieties than to the New England variety.

### Procedure

Data were collected in a private research space inside the science center. Participants were recruited from the floor of the science center by being asked if they (or their child) wanted to participate in a
“language experiment.” The testing space itself is a fishbowl-style room where other museum visitors can watch the participants engaging in the experiment. Each participant was seated at a desktop computer and wore headphones. To minimize distractions, the computers were positioned so that all participants faced toward a wall with their back to the public-facing glass of the laboratory.

For each trial, participants were presented with the stimulus sentence produced by one talker and were asked to rate the talker on one of five dimensions. One dimension focused on geographic locality (“from Ohio”), two dimensions focused on status attitudes (“smart” and “rich”), and two dimensions focused on solidarity attitudes (“honest” and “friendly”). Responses were selected from a 5-point rating scale displayed on the computer screen, on which all options were explicitly labeled. For example, the labels for the “smart” dimension were “very smart,” “a little smart,” “maybe or maybe not smart,” “not really smart,” and “not at all smart.” For the “from Ohio” dimension, the labels ranged from “very much like she’s from Ohio” to “not at all like she’s from Ohio.” Ratings were blocked by dimension, and each participant rated all 12 talkers on a single attribute before proceeding to the next attribute. The order of presentation of the five blocks and of the 12 talkers within each block was randomized separately for each participant. The experiment was preceded by a brief practice block involving rating the loudness of each of three male Midland talkers from the TIMIT corpus saying the same target sentence to allow the participants to become comfortable with the task. For all practice and test trials, the rating scale was displayed on the computer screen during the auditory stimulus presentation.

Adult participants filled out a brief demographic questionnaire after the rating blocks. Child participants (ages 4–17) filled out the demographic questionnaire prior to the rating blocks, which allowed the experimenter to evaluate the participant’s comfort level in reading from and interacting with a computer. For children ages 8–11 who needed help in completing the demographic questionnaire, the experimenter read aloud the instructions displayed on the computer screen. For all children ages 4–7, beginning with the demographic questionnaire and continuing through all rating blocks, the experimenter read aloud the instructions displayed on the computer screen and elicited the participant’s responses verbally; the experimenter then recorded the participant’s responses in the experiment interface on the computer. If participants seemed overwhelmed by choosing from five response options, the 5-point rating scale was divided into a series of binary questions (i.e., asking “smart or not smart?” and if the child said “smart,” following up with “very smart or a little smart?”), resulting in few ratings at the midpoint of the 5-point scale for the youngest age group.

**Analysis**

A series of linear mixed effects regression models was used to explore responses for each of the five rating scales. The dependent variable was each listener’s rating for each talker on a given attribute. The omnibus model for each attribute included fixed effects of age group (as a numerical variable representing age in years: 4, 6, 8, etc.), talker dialect (coded using sum contrasts), and their interaction, as well as a random intercept and talker dialect slope for each listener and a random intercept and age group slope for each talker, that is, Rating = AgeGroup × TalkerDialect + (1 + TalkerDialect|Listener) + (1 + AgeGroup|Talker). Age group was entered as a numerical variable to veridically capture the larger gaps in age between the adult groups than the child groups, allowing for the straightforward interpretation of any observed changes in adulthood. To confirm that this characterization of age did not mask developmental effects in childhood, parallel analyses were also conducted on the subset of data from children (ages 4–17). Significance of the fixed effects in the omnibus models was determined using log-likelihood comparisons. When talker dialect was significant in an omnibus model, Tukey’s tests were conducted to assess the significance of pairwise talker dialect differences. To further explore the developmental trajectory of performance in these tasks, a series of post hoc models was created for each age group. These post hoc models included a fixed effect of talker dialect, a random intercept and talker dialect slope for each listener, and a random intercept for each talker, that is, Rating = TalkerDialect + (1 + TalkerDialect|Listener) + (1|Talker). Tukey’s tests were used to assess the significance of pairwise talker dialect differences within each age group.

To explore the relation between locality judgments and language attitudes, a series of models were created to predict language attitude ratings from locality ratings, with a random locality slope for each listener and random talker and listener intercepts, that is, Rating = LocalityRating + (1 + LocalityRating|Listener) + (1|Talker). The fit of these
models, defined as the correlation between the fitted and observed values, was compared to the fit of the omnibus models with listener age group and talker dialect as fixed effects to assess whether general age-related processes of cognitive development and language exposure, or direct evaluations of talkers’ region of origin, better account for listeners’ status and solidarity ratings.

Results

Figure 1 shows the mean locality, “smart,” “rich,” “honest,” and “friendly” ratings for each age group for each of the four dialects. In the omnibus models, a significant fixed effect of age group was observed for “smart,” \( \chi^2(1) = 11.64, p < .001 \), and “honest,” \( \chi^2(1) = 4.25, p < .05 \). In both cases, older listeners assigned higher ratings than younger listeners. The fixed effect of age group was not significant in the models including only the subset of children, except for locality, \( \chi^2(1) = 5.30, p < .05 \), for which younger children assigned higher ratings than older children. Thus, the overall effect of age for “smart” and “honest” most likely reflects differences between children’s and adults’ use of the rating scale, whereas the effect of age for locality within the child group most likely reflects their

Figure 1. Ratings on each measure by talker dialect region (M = Midland, N = North, S = South, and NE = New England) and age group for all listeners. Error bars represent 95% confidence intervals.
developing awareness of how speech variation indexes where a talker is from.

The omnibus models also revealed significant fixed effects of talker dialect for locality, \( \chi^2(3) = 26.43, p < .001 \), “smart,” \( \chi^2(3) = 19.45, p < .001 \), “rich,” \( \chi^2(3) = 10.63, p < .05 \), and “honest,” \( \chi^2(3) = 17.54, p < .001 \). For locality, all pairwise talker dialect differences were significant (\( p < .001 \)), except for Midland and North, revealing a three-way locality distinction in which Midland and Northern talkers were rated as most likely to be from Ohio, New England talkers were rated as least likely to be from Ohio, and Southern talkers were rated in between. For “smart,” “rich,” and “honest,” all pairwise talker dialect differences were significant (\( p < .05 \)), except for Midland and North, on the one hand, and New England and South, on the other, revealing a two-way distinction between talkers in a series of post hoc models for each age group. Significant pairwise differences between talker dialects in these post hoc models are shown in Tables 2–6 for locality, “smart,” “rich,” “honest,” and “friendly,” respectively. Differences between Midland and Northern talkers were not significant in any analysis, so these comparisons are not shown in the tables. As shown in Table 2, for 6- and 7-year-olds, New England talkers were rated as less likely than Midland, Northern, and Southern talkers to be from Ohio. For listeners in all other age groups, both Southern and New England talkers were rated as less likely than Midland and Northern talkers to be from Ohio, and New England talkers received lower “from Ohio” ratings than Southern talkers. That is, when listeners were able to accurately report that Southern talkers were less likely to be from Ohio than Midland talkers, they also rated New England talkers as even less likely than Southern talkers to be from Ohio, confirming the three-way locality distinction between Midwestern (Midland and Northern), Southern, and New England talkers observed in the omnibus analysis. The two exceptions to this pattern were in the 12- and 13-year-old group, who rated Midland and Southern talkers as equally likely to be from Ohio, and in the oldest adult group, who rated Northern and Southern talkers as equally likely to be from Ohio.

As shown in Table 3, for 4- and 5-year-olds, only New England talkers were rated as less smart than

---

**Table 2**

Pairwise Talker Regional Dialect Comparisons of Mean “From Ohio” Ratings for Each Age Group

<table>
<thead>
<tr>
<th>Age group (years)</th>
<th>Midland versus New England</th>
<th>North versus New England</th>
<th>South versus New England</th>
<th>Midland versus South</th>
<th>North versus South</th>
</tr>
</thead>
<tbody>
<tr>
<td>4-5</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>6-7</td>
<td>***</td>
<td>***</td>
<td>**</td>
<td>ns</td>
<td>ns</td>
</tr>
<tr>
<td>8-9</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>ns</td>
<td>****</td>
</tr>
<tr>
<td>10-11</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>ns</td>
<td>***</td>
</tr>
<tr>
<td>12-13</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>ns</td>
<td>**</td>
</tr>
<tr>
<td>14-15</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>ns</td>
<td>***</td>
</tr>
<tr>
<td>16-17</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>ns</td>
<td>***</td>
</tr>
<tr>
<td>18-34</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>ns</td>
<td>***</td>
</tr>
<tr>
<td>35-49</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>ns</td>
<td>***</td>
</tr>
<tr>
<td>50-75</td>
<td>***</td>
<td>***</td>
<td>***</td>
<td>ns</td>
<td>***</td>
</tr>
</tbody>
</table>

*p < .05. **p < .01. ***p < .001.
Midland talkers. For 6- and 7-year-olds, both New England and Southern talkers were rated as less smart than Northern talkers, but only New England talkers were rated as less smart than Midland talkers. For listeners in all other age groups, both Southern and New England talkers were rated as
less smart than Midland and Northern talkers, but the ratings of Southern and New England talkers did not differ. The only exception to this pattern was in the oldest adult group, who rated Northern and Southern talkers as equally smart, consistent with their locality ratings for these two dialects. Thus, similar to the locality ratings, the youngest children distinguished the New England talkers from the other dialect groups, and beginning at age 8–9 years old, children showed the adult-like two-way distinction: Midland and Northern talkers were rated as smarter than Southern and New England talkers.

As shown in Table 4, beginning with 12- and 13-year-olds, Southern and New England talkers were consistently rated as less rich than Midland and Northern talkers, with the exception of two of the adult groups (18- to 34-year-olds and 50- to 75-year-olds), who did not differentiate the Northern talkers from the Southern and, for 18- to 34-year-olds only, New England talkers. Prior to age 12, regional differences were inconsistent, with New England talkers rated as less rich than Midland, Northern, and Southern talkers by 6- and 7-year-olds; and Southern talkers rated as less rich than Midland talkers by 10- and 11-year-olds. Thus, like the “smart” ratings, the “rich” ratings generally revealed a two-way distinction in which Midland and Northern talkers received higher ratings than Southern and New England talkers, although this pattern took several more years to develop.

As shown in Table 5, beginning with 8- and 9-year-olds, New England talkers were rated as less honest than both Midland and Northern talkers by all age groups, except the 14- through 17-year-olds, who showed that effect only for Midland talkers, and for young adults (18- to 34-year-olds), who did not rate New England talkers as less honest than either Midland or Northern talkers. New England talkers were also rated as less honest than Midland talkers by 6- and 7-year-olds. Southern talkers were rated as less honest than Northern talkers by 8- through 11-year-olds, and as less honest than Midland talkers by 10- and 11-year-olds, 14- and 15-year-olds, and adults over 35 years. Southern and New England talkers were not rated differently from each other by any age group. Overall, like the “smart” and “rich” ratings, Midland and Northern talkers were given higher ratings than Southern and New England talkers, although young adults distinguished less among dialect groups than children and older adults.

As shown in Table 6, the results for the friendliness ratings reveal very few significant pairwise differences and the handful of significant effects follow no consistent developmental pattern. Southern talkers were rated as less friendly than Midland talkers only by 10- and 11-year-olds and 50- to 75-year-olds. Southern talkers were also rated as less friendly than Northern talkers by 10- and 11-year-olds, 14- and 15-year-olds, and adults over 35 years. Southern and New England talkers were not rated differently from each other by any age group. Overall, like the “smart” and “rich” ratings, Midland and Northern talkers were given higher ratings than Southern and New England talkers, although young adults distinguished less among dialect groups than children and older adults.

Predicting Status and Solidarity Ratings From Locality Ratings

The relation between language attitudes and geographic classification was examined by comparing the fits of the omnibus models described above to the fits of parallel models predicting the attitude ratings from the locality ratings. For all four sets of attitude ratings, the omnibus model involving
listener age group and talker dialect had a better model fit ($r^2 = .54$ for “smart”; $r^2 = .48$ for “rich”; $r^2 = .39$ for “honest”; $r^2 = .37$ for “friendly”) than the locality model ($r^2 = .47$ for “smart”; $r^2 = .37$ for “rich”; $r^2 = .31$ for “honest”; $r^2 = .31$ for “friendly”). This pattern of results suggests that locality judgments may be an important precursor to language attitudes but that other developmental factors likely also contribute to this ability.

**Discussion**

Participants from across the life span were asked to rate talkers from four regional dialects of American English along three types of dimensions: geographic locality, status attitudes, and solidarity attitudes. The developmental trajectory for each dimension was somewhat different, but taken together, these results provide evidence that (a) locality judgments exhibit a two-stage developmental pattern, in which the New England dialect is first identified as nonlocal and then the Southern dialect is identified as less local than Midwestern varieties but more local than the New England variety; (b) perceptions of regional dialect locality are fundamentally linked to attitude judgments; but (c) geographic knowledge is not the sole mediator of attitude judgments, and therefore other factors such as language exposure through schooling and travel, and cultural knowledge about stereotypes and language prestige, also likely play a role.

*Development of Locality Judgments*

In the locality rating task, 6- and 7-year-olds were able to correctly rate talkers from the geographically farthest region (New England) as being less local than talkers from the other three regions. This result demonstrates that the basic understanding that speech variation can index geographic information is available by age 6–7 years. This early success is also consistent with Kinzler and DeJesus’s (2013) finding that Northern children around this age were able to correctly make a locality judgment (“from here”) for Northern and Southern talkers (see also Creel, 2017). These successes stand in contrast to the poorer performance described by Girard et al. (2008), and Wagner, Clopper, et al. (2014) who found that 5- and 6-year-olds could not use regional dialect information to put talkers into ad hoc social groups. One factor that may explain these successes versus failures is the explicit use of real geographic labels for the dialects, albeit somewhat general ones (i.e., “from here” and “from Ohio”). This specification may have encouraged children to anchor their judgments to their own dialect or social group, providing important support for their ability to make relevant judgments. That is, young participants’ judgments may be wholly egocentric (i.e., does this talker sound like me?).

Improvements in the locality rating task emerged at age 8–9 years, when children demonstrated adult-like three-way locality distinctions with New England talkers rated as least likely to be from Ohio, Midland and Northern talkers as most likely to be from Ohio, and Southern talkers in between. This intermediate rating for the Southern talkers reflects a veridical judgment about the world because the Southern region, including Kentucky and West Virginia, is geographically adjacent to Ohio (Labov et al., 2006), whereas the New England region is more geographically distant. One possible explanation for this geographically veridical pattern of results is that children were not making a geographically based judgment at all but were simply judging the acoustic–phonetic properties of the dialects. After all, previous work has found that children at this age have a rather incomplete sense of geography (Jahoda, 1963, 1964; Piaget & Weil, 1951). That is, children may perceive the dropping of the postvocalic /r/ in the New England dialect as more salient or more different from the local Midland dialect than the intervocalic voicing of /s/ in the Southern dialect and then simply rate the most perceptually distant dialect as the least likely to be from Ohio. Further evidence supporting the idea that children are basing their judgments on gradient acoustic–phonetic differences comes from similar dialect classification tasks, in which children succeed with foreign-accented and second-language talkers but not with native talkers from a different regional dialect (Girard et al., 2008; Wagner, Clopper, et al., 2014). Presumably the speech of non-native or second-language talkers contains more, or more salient, differences from the listener’s own speech than regional dialects, leading to more accurate classification performance for unfamiliar non-native versus native varieties.

Evidence against the idea that children are relying primarily on a perceptual distance strategy comes from several sources. First, there is no a priori reason to believe that the specific features of the New England dialect are more perceptually salient than those of the Southern dialect nor are there well-established metrics for independently assessing acoustic–phonetic dialect distance. It is therefore
impossible with the current data to disentangle children's perceptual differentiation of the dialects from other associated properties of the dialects, such as their familiarity (which is arguably driven at least in part by geographic distance). Second, the three-way distinction in locality judgments among the four dialects that begins at age 8 years persists throughout adulthood and therefore appears to reflect adult-like knowledge. Although it is possible that the adults and children produced the same rating patterns in this task for very different reasons, it seems equally plausible that there is genuine continuity in the representations of regional dialects from the age of 8 years onward. Given this assumption of continuity, the children must have been making some kind of geographic judgment in the locality ratings task at least by age 8–9 years. Finally, as will be discussed further below, for the language attitude judgments, the New England and Southern dialects patterned together (and not separately, as in the locality task), suggesting that participants interpret the perceptual differences among the dialects differently depending on the task at hand. Thus, what makes an acoustic–phonetic feature perceptually relevant depends on what is being asked.

The observed adult-like performance by age 8–9 years contrasts with the results of related studies (e.g., Jones et al., 2017; Williams et al., 1999), in which adult-like performance was not observed until adolescence. These differences across studies suggest that although basic competence in regional dialect perception is in place early, specific perceptual performance emerges over a period of years as a function of the cognitive demands of the task and language experience (i.e., through schooling, travel, and other kinds of language exposure). Locality-oriented competence should therefore be found, given an appropriate task with appropriate varieties, for all children (e.g., New England children should be able to identify Southern talkers as nonlocal by age 6–7 years), but the endpoint of the trajectory will vary depending on task demands, the specific language varieties being investigated, and children's experience with those varieties.

**Development of Language Attitude Judgments**

The adult ratings for status attitudes (“smart” and “rich”) were consistent with the previous literature (Cheyne, 1970; Ladegaard, 1998; Levin et al., 1994; Luhman, 1990): the Southern and New England dialects are stereotyped as relatively nonstandard (Niedzielski & Preston, 2003), and the talkers from these regions were accordingly downgraded on both measures of status. Developmentally, the youngest children agreed with the adults about Midland and New England talkers for the “smart” status dimension, giving relatively high “smart” ratings to Midland talkers and lower ratings to New England talkers. Starting at age 6–7 years, children also downgraded the Southern talkers on the “smart” dimension and showed the adult-like pattern by age 8–9 years. For the “rich” ratings, children did not display the adult-like pattern favoring Midland and Northern talkers over Southern and New England talkers until age 12–13 years, although they showed some inconsistent downgrading of the less local dialects beginning at age 6–7 years. It is unclear why adult-like performance on the “rich” ratings was delayed relative to the “smart” ratings, but the young children may have had a less fully developed understanding of wealth compared to intelligence, especially given that the socioeconomic status (SES) of the average science center visitor is high. Although discussions of children's intelligence are not uncommon in American households, discussions of relative wealth in these high SES households may be. Anecdotally, some of the children explicitly mentioned princesses and bags of gold when they were performing the “rich” ratings, suggesting that their understanding of this concept may be supported more by fantasy and fictional accounts than by real-world examples. If this interpretation is correct, it points to the general difficulty of aligning children's conceptual worlds with those of adults, and suggests that children's conceptual development may also play a critical role in the development of language attitudes.

The ratings for the two measures of solidarity (“honest” and “friendly”) were consistent with the literature in that they appear to be influenced by a wider range of factors than status ratings (Fuertes et al., 2012). For the “honest” ratings, the overall pattern was similar to the pattern found for both measures of status, with Midland and Northern talkers receiving higher ratings than Southern and, especially, New England talkers. However, as shown in Figure 1, the differences between dialect groups were less extreme for the “honest” ratings than for either type of status rating, and inspection of the results for each age group suggests a somewhat inconsistent developmental trajectory. The “friendly” ratings patterned differently from both the status ratings and the “honest” ratings. The developmental trajectory was highly inconsistent, and in most of the age groups, none of the dialects
were differentiated. It is unclear why participants failed to use dialect information to direct their “friendly” ratings. One possibility is that they were being influenced by the prominent customer-friendly atmosphere at the testing site: notably friendly people are the norm at the science center where the participants were recruited. This local environment may have been a more compelling cue for the ratings than participants’ general knowledge about dialect.

The Relation Between Language Attitudes and Locality Judgments

The results of the current study provide some evidence for the assumption that locality judgments mediate language attitudes: The models examining locality as a predictor of language attitude ratings revealed that locality was a significant predictor of all four language attitudes. This positive finding stands in contrast to the work by Creel (2017), who found no correlation between locality ratings and a friendship judgment. However, the models containing only locality as a predictor did not fit as well as the more complex models containing both listener age group and talker dialect as predictors, and more generally, the model fits in the regression analyses were consistently only moderate in magnitude ($0.3 < r^2 < 0.6$), suggesting that factors beyond locality, listener age, and talker dialect must contribute to language attitude ratings.

Together with the previous literature, the current results implicate a wide range of additional factors that are likely to be relevant for the development of language attitudes: cognitive abilities, lexical knowledge, metalinguistic awareness, dialect familiarity, geographic knowledge, and sociocultural knowledge. The first three of these factors relate to skills that undergo a great deal of development over childhood but are expected to remain relatively stable once they have been established. Thus, as noted previously, young children may not have fully articulated lexical entries to go with the specific attitude descriptors used in these tasks and an underspecified representation of the meaning of a word such as “rich” would necessarily impact ratings for that attitude. At some point in development, however, children’s linguistic knowledge solidifies and can be drawn on for the remainder of the life span. For example, the current data suggest that knowledge of the word “rich” solidifies at age 12–13 years.

In contrast to these first three factors, other influences have more potential for ongoing change throughout the life span. For example, as noted above, the locality ratings reflect geographic facts about the world, but those facts also align with the overall familiarity of listeners in Ohio with each of the four dialects because geographically close varieties are likely to be more familiar than geographically more distant varieties. Indeed, there is a hint within these data that familiarity is important and further, that more extensive life experience can change what is familiar. Specifically, the oldest age group (50- to 75-year-olds), unlike all other participants over the age of 8 years, rated Southern talkers as equally likely to be from Ohio as Northern talkers. This age group included the fewest number of lifelong Midwesterners (only 25%), and it is thus possible that having less familiarity with the Midwest made these listeners less sensitive to the local conceptualization of regional dialect boundaries (cf. Preston, 1993).

Alternatively, or in addition, the geographically close varieties in the current study are more prestigious than the geographically more distant varieties (Campbell-Kibler, 2012; Niedzielski & Preston, 2003). Prestige is a complex and culturally embedded notion requiring an understanding of how dialects are socially perceived across a variety of contexts, and this kind of knowledge takes years to accumulate. The observed relation between locality and attitudes could therefore reflect prestige because both of the Midwestern dialects (Midland and Northern) have higher social standing than the Southern or New England dialects (Niedzielski & Preston, 2003). Unfortunately, prestige and familiarity were heavily confounded in this study and it would require data from participants in the South or New England to disentangle the effects of exposure from cultural knowledge about language prestige.

One very strong reason to believe that there is not a simple, direct link between the locality judgments and language attitudes is the fact that the locality ratings showed a three-way differentiation among the dialects by age 8 years, whereas by contrast, language attitudes were divided into a simpler two-way split in which the more local dialects received more positive ratings than the less local dialects, regardless of how much less local they were. Thus, despite participants’ more nuanced judgments about where the talkers were from, their attitudes reflected a binary judgment. One possible reason for the simpler two-way distinction for attitudes is that they depend on a more complex characterization of the dialects. In particular, as noted above, Midland and Northern are not only the local Ohio dialects, they are also higher in prestige than the two non-Ohio dialects. If prestige plays an important role in language attitudes and is
more binarily construed than geography, fewer distinctions may emerge in language attitudes than in locality judgments. As noted above, data from participants from other areas of the country would allow the relative contributions of geography, dialect familiarity, and prestige to language attitudes to be disentangled.

Additional support for the lack of a direct link between locality and attitude judgments comes from the subtle pattern of differences observed for the Midland and Northern dialects. Although these two dialects were never rated as statistically different from one another, they showed statistically different patterns relative to the other two dialects. Specifically, as shown in Tables 3–6, there were many cases where the Midland dialect was rated as significantly different from the Southern or New England dialect, but the Northern dialect was not. Inspection of the mean ratings in Figure 1 reveals that in these instances, the Northern talkers received slightly lower ratings than the Midland talkers. These differences were never large enough to statistically differentiate the Northern and Midland talkers from each other, but they were sufficiently large to change the relative comparisons between these two dialects and the Southern and New England dialects. As discussed earlier, the Midland and Northern dialects are spoken in neighboring geographic regions and both are spoken in the state of Ohio. It is thus geographically accurate to rate them both as local, despite their acoustic-phonetic differences. However, those acoustic-phonetic differences can be leveraged by listeners of various ages when making language attitude judgments, revealing that the same linguistic cues can be interpreted differently depending on the task.

Conclusions

Taken together, the results of this study provide novel evidence about the time course of the development of regional dialect classification, regionally differentiated language attitudes, and their relation. Of necessity, this study was conducted with children living in a specific location, and the four talker dialects were chosen with respect to that location. However, the specificity of the materials does not undermine the generality of the findings or what they reveal about how dialect information is represented and linked to language attitudes. The results demonstrate that the fundamental aspects of dialect perception are understood by children as young as 6–7 years: these children classified some dialects as being more local than others, and also showed some adult-like status and solidarity judgments. Over the course of development, locality ratings were a significant, but modest, predictor of attitude ratings, suggesting that geographic knowledge is one contributor to language attitudes throughout development, although other factors must also play a role. These data further suggest that one key factor is familiarity or exposure, as listeners had earlier and more consistent ratings about the least familiar dialect relative to the most familiar one. Another key factor is cultural knowledge about language stereotypes and prestige, as listeners’ judgments were generally consistent with the cultural status of the dialects and with cultural geographic boundaries relevant to the area. Yet another important factor is general cognitive and conceptual development, as the trajectories for the attitudes were different: status judgments, which are about more verifiable and stable properties, were more consistent and learned earlier, whereas solidarity judgments, which depend more on local interpersonal interactions, were less consistent across the life span. Replicating this study in a different geographic region would require careful consideration of which talker dialects to include, but children’s language attitudes are expected to be influenced by the same factors.

The results of this study also demonstrate that in this domain of dialect perception, the development of adult-like performance is protracted, and only by examining abilities across an extended age range can the nature of the changes that take place and their connections to one another be observed. A great deal of language development has been accomplished by the time children enter school, but not everything is learned early. Dialect perception is a domain that requires experience with a range of linguistic varieties, an appreciation of social constructs such as status and solidarity, and general geographic and cultural knowledge. These elements need to be integrated not only with each other but with the linguistic system more generally to allow for the rapid interpretation of social-indexical information in the speech signal.

References


