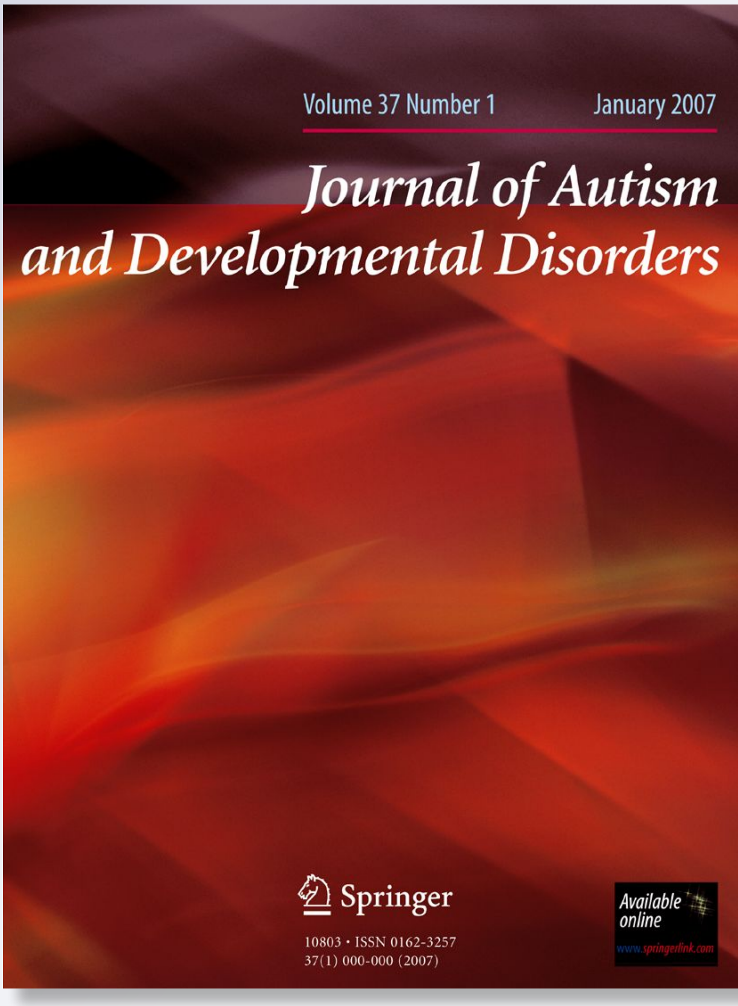
Perception of Dialect Variation by Young Adults with High-Functioning Autism

Cynthia G. Clopper, Kristin L. Rohrbeck & Laura Wagner

Journal of Autism and Developmental Disorders

ISSN 0162-3257
Volume 42
Number 5

J Autism Dev Disord (2012) 42:740-754
DOI 10.1007/s10803-011-1305-y



Springer

J Autism Dev Disord (2012) 42:740–754
DOI 10.1007/s10803-011-1305-y

ORIGINAL PAPER

Perception of Dialect Variation by Young Adults with High-Functioning Autism

Cynthia G. Clopper · Kristin L. Rohrbeck · Laura Wagner

Published online: 14 June 2011

Abstract The linguistic profile of people with Autism spectrum disorders typically involves intact perceptual processing, accompanied by deficits in the social functions of language. In a series of three experiments, the impact of this profile on the perception of regional dialect was examined. Young adults with High-Functioning Autism exhibited similar performance to a typically developing comparison group in regional dialect classification and localness rating tasks, suggesting that they can use indexical information in speech to make judgments about the regional background of unfamiliar talkers. However, the participants with High-Functioning Autism were less able to differentiate among the dialects in a language attitudes task, suggesting that they do not share social stereotypes related to dialect variation with the typically developing comparison group.

Keywords Dialect · Speech perception · Language attitudes · Social language · High-Functioning Autism

One of the hallmark characteristics of people with Autism spectrum disorders (ASD) is a deficit in language and communication. Even adults with High-Functioning Autism (HFA), who have good general language abilities, continue to have difficulties with the more socio-communicative aspects of language interaction, or pragmatics (Baron-Cohen et al. 2000; Rice et al. 2005). The current studies consider a dimension of language competence that has received very little attention in the literature: knowledge of indexical information, in particular, regional dialect information. Indexical information indexes specific features of the speech event, such as the identity of the talker, where the talker is from, and the relationship among interlocutors. The detection of indexical information requires a detailed acoustic analysis of the speech signal, but its content is largely social in nature. It provides, therefore, an interesting domain in which to examine the interplay between perceptual processing (a skill that adults with HFA are typically good at) and social processing (a skill that adults with HFA typically have difficulty with).

C. G. Clopper (✉)
Department of Linguistics, Ohio State University, 1712 Neil Avenue, Columbus, OH 43210, USA
e-mail: clopper.1@osu.edu

K. L. Rohrbeck · L. Wagner
Department of Psychology, Ohio State University, Columbus, OH, USA

Indexical Information in the Speech Signal

Variability in the speech signal has many sources, including linguistic properties of the utterance, indexical properties of the talker, and reverberation qualities of the environment (Klatt 1989). Linguistic variability results from factors such as coarticulation between segments within and across word boundaries, and phonetic reduction due to lexical and semantic properties of the utterance. Indexical variability results from factors such as a talker's age, gender, ethnicity, socioeconomic status, regional background, and current emotional state, as well as individual voice differences across talkers. Listeners extract information from the variable speech signal to make explicit judgments about the linguistic and indexical content of speech. For example, Ladefoged and Broadbent (1957) found that listeners judged the same acoustic signal differently depending on the phonetic properties of the voice preceding the target word. When the voice was consistent with a relatively large man, listeners perceived the ambiguous stimulus as *bit*, interpreting the mid-high vowel as being relatively high in the large man's vowel space, but when the voice was consistent with a smaller man, listeners perceived the same stimulus as *bet*, interpreting the same mid-high vowel as being lower in the smaller man's vowel space. Thus, listeners use indexical information about the talker to make linguistic judgments about target words. Listeners can also make explicit judgments about indexical properties, including identifying familiar talkers (Pollack et al. 1954; Van Lancker et al. 1985) and classifying unfamiliar talkers by age (Ptacek and Sander 1966), gender (Lass et al. 1976), and ethnicity (Lass et al. 1979; Purnell et al. 1999). The focus of the current study was the perception of one particular source of indexical variability: regional dialect.

Dialect Perception by Typically Developing Adults

Previous research has examined the perception of regional dialect by typically developing (TD) adults in a range of explicit classification and implicit language attitudes tasks. The most common classification paradigm involves a forced-choice task, in which listeners hear talkers one at a time and are asked to indicate which region from a specified set of choices they believe each talker is from. Forced-choice dialect categorization tasks have been conducted for regional varieties of English in Wales (Williams et al. 1999), the United Kingdom (Van Bezooijen and Gooskens 1999), and the United States (Baker et al. 2009; Clopper and Pisoni 2004a, b, 2006; Preston 1993), as well as for regional varieties of Dutch in the Netherlands and Belgium (Van Bezooijen and Gooskens 1999). The results of these studies reveal that TD adolescents and adults can identify the regional background of unfamiliar talkers with a reasonable degree of accuracy. In addition, performance is typically more accurate for local varieties than non-local varieties (Baker et al. 2009; Clopper and Pisoni 2004a; Williams et al. 1999), although listeners also exhibit a bias to identify talkers from geographically close regions as local (Clopper and Pisoni 2006; Preston 1993).

The results of the forced-choice dialect classification tasks also suggest that listeners' perceptual dialect categories do not correspond exactly to the dialect categories proposed by sociolinguists based on production data (Clopper and Pisoni 2004b; Preston 1993). To examine the perceptual structure of dialect variation for naïve listeners in more detail, Clopper and Pisoni (2007) developed a free classification task, in which listeners are asked to group unfamiliar talkers by dialect. In this task, dialect categories are not provided by the experimenter and participants are allowed to make as many groups with as many talkers in each group as they want. Clopper and Pisoni (2007) found that listeners make more fine-grained dialect categories than expected. When presented with unfamiliar talkers from six regional dialects, individual listeners make 8-10 groups of talkers on average. However, the aggregate data are similar to forced-choice tasks in revealing modest accuracy with a bias for classifying talkers from geographically close regions with talkers from the local region. Taken together, the results of the forced-choice categorization and free classification tasks suggest that TD listeners can use variation in the speech signal to make explicit judgments about where an unfamiliar talker is from.

In addition to explicit dialect classification tasks, the perception of regional and social dialect variation has also been examined through implicit language attitudes tasks that tap listeners' evaluations of a talker's personal attributes based on his or her speech. In these tasks, listeners are asked to rate a set of unfamiliar talkers on a range of personal attributes, such as intelligence, level of education, reliability, and friendliness. Correlation and factor analysis techniques have demonstrated that these attributes can be divided into two categories, reflecting dimensions of status (e.g., intelligence and level of education) and solidarity (e.g., friendliness and reliability) (Giles 1970; Luhman 1990). The results of these studies typically reveal that socially prestigious linguistic varieties are rated higher than less prestigious local varieties on attributes reflecting status, but lower than local varieties on attributes reflecting solidarity. For example, Luhman (1990) found that General American English was rated higher for status, but lower for solidarity, than Appalachian American English by listeners from Kentucky. Similar results have been obtained for English and French in Montreal (Lambert et al. 1960), regional varieties of English in the United Kingdom (Ladegaard 1998a), regional varieties of Danish in Denmark (Ladegaard 1998b), and ethnic varieties of English in Australia (Gallois and Callan 1989). These results suggest that TD listeners can also use variation in the speech signal to make evaluative judgments about personal attributes of the talker, and that those judgments typically correspond to community stereotypes about the relative status or prestige associated with different varieties. Socially prestigious linguistic varieties are associated with higher status than less prestigious varieties, but listeners exhibit stronger solidarity with local varieties than with non-local prestige varieties.

Perception of Indexical Categories by People with Autism Spectrum Disorders

The perception of indexical information depends in part on the perceptual processing of phonological and phonetic elements. People with ASD have been found to be generally unimpaired in this domain and often out-perform their

TD peers on tasks involving low-level acoustic processing (see Mottron et al. 2006; Happé and Frith 2006 for reviews). More specifically, with respect to the acoustic processing of language, children and young adults with ASD performed as well as their peers in a phonological production task (Bishop et al. 2004), including in situations with stressful auditory input (Nober and Simmons 1981), and they also perceived native and non-native phoneme contrasts as well as their peers (Constantino et al. 2007). In the related domain of prosody, the results are mixed. When asked to recognize and judge specific prosodic contours, people with ASD performed as well as, and in some conditions better than, their peers (Jarvinen-Pasley et al. 2008a, b). However, other researchers have found persistent problems in the production and perception of various dimensions of prosody in speech (Paul et al. 2005; McCann et al. 2007). Moreover, people with ASD have particular problems perceiving speech when the system is stressed, such as when listening to speech in noise (Alcántara et al. 2004).

The understanding of social information is a well-known deficit of people with ASD, but investigations of the ability of people with ASD to correctly perceive social indexical information from the speech signal have found mixed results. The ability to recognize individual voices and link them to individual talkers is unimpaired in people with ASD (Boucher et al. 2000). Moreover, although people with ASD do worse than their TD peers when asked to adjust their speech for different kinds of interlocutors (e.g., babies and foreigners), they do modify their speech as a function of the social identity of their listener (Volden et al. 2007; Volden and Sorenson 2009). By contrast, people with ASD have been found to have difficulty perceiving the emotional intent of speech (Golan et al. 2007) and whether or not speech was intended to convey sarcasm or praise (Imaizumi et al. 2009). Similarly, they have difficulty in producing target sentences in different emotional colors (Hubbard and Trauner 2007).

Perhaps the most striking finding for the current research comes from an investigation of regional dialect among adolescents with ASD by Baron-Cohen and Staunton (1994). TD adolescents usually adopt the dialect of their peers (Chambers 2002). For example, children with immigrant parents speak with the local native accent of their friends at school as opposed to the non-native accent of their older family members. Baron-Cohen and Staunton (1994) examined the speech of British children with ASD whose mothers were non-native speakers of English. They observed that, unlike TD children, their speech was more similar to the speech of their mothers than to that of their peers. This finding was in particular contrast to the TD siblings of the participants with ASD who showed the more common pattern of producing speech more similar to their peers than to their mothers. Thus, it appears that even though these adolescents with ASD were producing indexical information related to accent, they were not opting to index themselves in the same way as their TD peers. It remains an open question how young adults with ASD socially index regional dialect more generally; they may in fact not realize how social stereotypes are linked to dialect differences even though they can apparently perceive the relevant acoustic differences in the speech signal.

Hypotheses

The current study examined the perception of regional dialect variation by young adults with HFA in a series of three experiments. The first experiment was a free classification task in which participants were asked to group a set of unfamiliar talkers by regional background. The second experiment was a localness judgment task in which participants were asked to rate how likely it was that each talker was from Columbus, Ohio, where the experiment was conducted. The third experiment was a language attitudes task in which participants were asked to rate each talker on two status scales (intelligence, successfulness) and two solidarity scales (friendliness, reliability). All of the tasks were also conducted with a comparison group of TD young adults.

The TD adults were expected to exhibit performance similar to the participants in previous studies examining the perceptual classification of and language attitude judgments towards regional varieties of American English (Baker et al. 2009; Clopper and Pisoni 2004a, b, 2006; Luhman 1990; Preston 1993). In particular, the TD listeners were expected to exhibit modest accuracy and to make a clear distinction between dialects that are geographically closer and dialects that are geographically farther away in the free classification and localness judgment tasks. In the language attitudes task, the TD listeners were expected to rate the talkers of the more prestigious varieties as more intelligent and more successful than the talkers of the less prestigious varieties, but to rate the talkers of the local varieties as more friendly and more reliable than the talkers of the non-local varieties.

Based on the previous research demonstrating the auditory processing skills and social cognition deficits of people with HFA, we expected that the participants with HFA in the current study would be able to perceive the dialect differences between the talkers. The free classification task requires participants to attend to dialect-specific differences in the speech signal and to group talkers with similar dialect characteristics together. If the listeners with HFA can successfully attend to the relevant aspects of the signal (i.e., dialect differences) and ignore irrelevant aspects of the signal (i.e., voice quality and pitch differences), they should be able to produce reasonable dialect groups in the free classification task. The localness judgment task similarly requires the listeners to attend to dialect-specific differences, but also requires listeners to make an explicit judgment about the relationship between the properties of the speech signal and the geographical background of the talker relative to the local area. If the listeners with HFA can attend to the relevant aspects of the signal and distinguish between varieties that are more and less similar to their own variety, they should also be able to produce reasonable responses in the localness judgment task. The language attitudes task further requires the listeners to associate dialect-specific characteristics in the speech signal with social stereotypes about the talker. If the listeners with HFA have not acquired the stereotypes that are shared by the TD community, they may have difficulty differentiating between the dialects in the language attitudes task.

Experiment 1: Free Classification

Methods

Participants

Thirty TD adults (age range 18–28 years) participated in the experiment to receive course credit in an introductory linguistics course. Three were excluded because they were not native English speakers. The 27 remaining participants (11 males, 16 females) were monolingual native speakers of American English with no reported history of hearing or speech disorders. Fourteen TD participants were from the Midland dialect region, 7 from the Northern dialect region, 1 from the New England dialect region, and 5 had moved between dialect regions before age 18.

Twenty-four adults with HFA (age range 18–29 years) were paid to participate in this experiment. One of these adults was excluded because he was not a native English speaker. The 23 remaining adults with HFA (17 males, 6 females) were monolingual native speakers of American English. Sixteen participants with HFA reported being from the Midland dialect region, 1 from the Northern dialect region, and 6 had moved between dialect regions before age 18.

The participants with HFA were recruited from an outreach program specifically designed for young adults who have been diagnosed with HFA. This program is run by a local center for developmental disabilities. Approximately 90% of the participants have graduated from high school and 62% have completed some college. The remaining 10% of the participants are currently enrolled in high school. In addition, most of the participants in this program (~90%) have obtained a formal diagnosis on the Autism spectrum prior to participating in the program, either as a child (33%) or as an adolescent (57%). Moreover, all study participants self-identified as having HFA, and their status was supported using the Autism-Spectrum Quotient (see next section).

Four of the 23 participants with HFA reported having a history of a speech, language, or hearing disorder at the time of the experiment (one with speech apraxia, one who had therapy as a child to help lower his pitch and volume, and two unspecified disorders). The data were analyzed without these 4 participants, but the results did not significantly differ when they were removed and so their data were retained for this and all subsequent experiments to maintain adequate statistical power.

Autism-Spectrum Quotient

Both participant groups completed the Autism-Spectrum Quotient (Baron-Cohen et al. 2001) to independently verify that our two populations (adults with HFA and TD adults) differed. The Autism-Spectrum Quotient is a self-report survey designed to quickly assess the degree to which adults have characteristics associated with ASD. The scores range from 0 to 50, with higher scores being associated with traits more representative of having ASD. Independent two sample t tests showed that, in the current experiment, the adults with HFA scored significantly higher on the Autism-Spectrum Quotient than the TD group (TD $M = 17.67$, SD = 5.54; HFA $M = 25.52$, SD = 8.16; $t(48) = 4.03$, $p < .001$, $d = 1.15$).

Baron-Cohen et al. (2001) posit that there is a "critical minimum" score of 32 that may identify adults with ASD. As a group, the adults with HFA in the current study did not earn a mean score above this critical point and only 7 of the 23 participants with HFA scored above 32. However, other research has similarly found that people with positive ASD diagnoses do not always obtain such high scores on the Autism-Spectrum Quotient, and the scores of the participants in the current study were comparable to scores found in other HFA populations (Ketelaars et al. 2008; Woodbury-Smith et al. 2005). Moreover, the participants with HFA in the current study were recruited from a program specifically designed to help people with HFA improve their communicative abilities and social skills, which are among the specific elements targeted by the Autism-Spectrum Quotient.

Stimulus Materials

The stimulus materials were produced by 20 white male talkers from the TIMIT corpus (Fisher et al. 1986). Five talkers from each of four dialect regions in the United States (New England, Northern, Midland, and Southern) were selected. The New England dialect is spoken in the northeastern US from Maine to eastern New York state. The Northern dialect is spoken in the upper midwestern US from western New York state to Minnesota. The Midland dialect is spoken in the lower midwestern US from Ohio to Kansas and Nebraska. The Southern dialect is spoken in the American South from Virginia to Texas. The Midland dialect is the local variety where the study was conducted (Columbus, OH). All of the talkers were in their 20s at the time of recording.

The stimulus materials consisted of the sentence "She had your dark suit in greasy wash water all year" produced by each of the 20 talkers. This sentence was recorded for all of the talkers included in the TIMIT corpus, because it was expected to elicit dialect-specific variation between talkers (Fisher et al. 1986). For example, the New England dialect is non-rhotic (or "r-less") and the New England talkers were therefore expected to produce *dark* as "dahk." Similarly, one of the characteristic properties of the Midland and Southern dialects is /u/-fronting and *suit* was therefore expected to be more similar to *seat* for the Midland and Southern talkers than for the Northern and New England talkers. Finally, the Southern dialect exhibits several word-specific pronunciation variants, including "warsh" for *wash* and "greazy" for *greasy*, which were expected to emerge in the Southern talkers' speech. Previous analyses of the sentences produced by the talkers in this study confirmed that r-lessness in *dark* distinguished the New England talkers and that "greazy" for *greasy* distinguished the Southern talkers from the other talkers in the stimulus set (Clopper and Bradlow 2009).

Each sentence was contained in an individual digital audio file sampled at 22,050 Hz with 16-bit resolution and linked to a rectangular digital image of the talker's initials in white font on a black background.

Procedure

Participants were presented with a single Powerpoint slide with the stimulus sentence ("She had your dark suit in greasy wash water all year") printed at the top. There was a 16 × 16 grid in the middle of the slide and the 20 talkers were displayed to the left of the grid as shown in Fig. 1. Participants were instructed to listen to and group the talkers by where they believed the talkers were from. They were told they could make as many groups with as many talkers in each group as they thought was appropriate, the groups did not have to be the same size, and they could listen to and rearrange the groups as many times as they wanted until they were satisfied with their groupings. Participants listened to the talkers by double-clicking on the talker icons and moved the icons onto the grid by dragging them with the mouse.

Results

Classification Strategy

The adults with HFA and the TD adults adopted very similar classification strategies in the free classification task. There were no significant differences between the populations in terms of the number of groups they made ($t(48) = 1.76$, ns), nor in terms of the number of talkers they put into each group ($t(48) = .784$, ns). Table 1 shows the means and standard deviations of the number of groups

Table 1 Means and standard deviations of the number of groups and number of talkers per group for each population

	# Groups	# Talkers/group
Adults with HFA	7.7 (3.9)	3.4 (1.9)
TD Adults	6.1 (2.4)	3.8 (1.4)

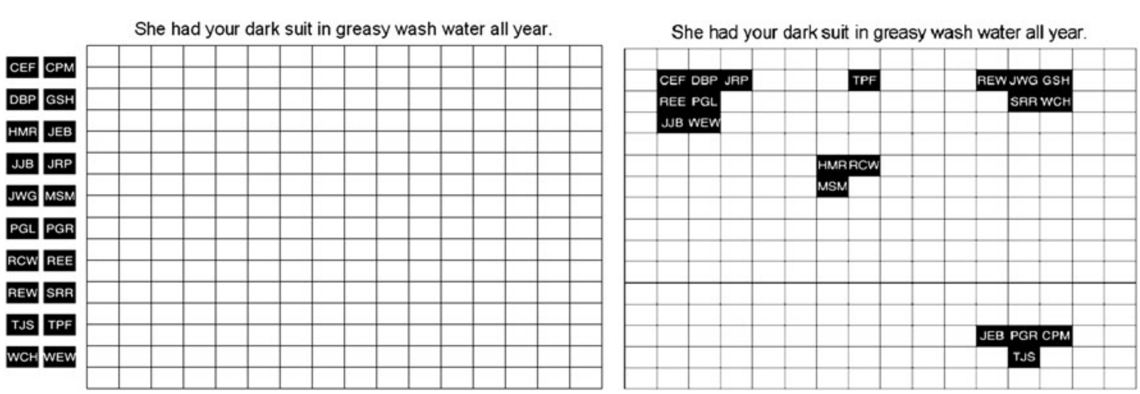


Fig. 1 The *left panel* displays what the participants were presented with at the beginning of Experiment 1. The *right panel* displays an example of one way a participant could group the talkers

and number of talkers per group made by the adults with HFA and the TD adults.

Accuracy

Accuracy was measured as a function of pairs of talkers. Given that there were five talkers in each dialect group, each talker could be correctly paired with (i.e., placed in the same group as) four other talkers and could be incorrectly paired with 15 other talkers. The adults with HFA were less accurate than the TD adults in terms of the percentage of correct pairs they made ($t(48) = 2.81$, $p < .01$, $d = .79$), but they were equally accurate in terms of the percentage of incorrect pairs (errors) they made ($t(48) = 1.93$, ns). Given that participants who made larger groups of talkers were more likely to make both more correct pairings and more incorrect pairings than participants who made smaller groups, a difference score between the percentage of correct pairings and the percentage of errors was calculated to measure overall accuracy. On this overall accuracy score, the adults with HFA were significantly less accurate than the TD adults ($t(48) = 4.61$, $p < .001$, $d = 1.31$). Table 2 shows the means and standard deviations of these accuracy scores.

Talker Similarity

Talker similarity matrices were constructed by counting the number of times each pair of talkers was put in the same group across listeners. Talkers with highly similar dialects were put in the same group by most of the listeners, whereas talkers with different dialects were put in the same group by few of the listeners. Separate matrices were constructed for the TD adults and the adults with HFA. The matrices were submitted to the additive clustering algorithm, ADDTREE (Corter 1982), to obtain a visual representation of the perceptual similarity of the talkers. Figure 2 shows the clustering solutions for the TD adults (left) and the adults with HFA (right). In these figures, the perceptual distance between any two talkers is represented by the length of the horizontal branches connecting them. Vertical distance is irrelevant. The three largest clusters perceived by each group are separated by hand-drawn solid horizontal lines. On the whole, the HFA and TD participants performed similarly. Both populations sorted the New England and the Southern talkers into distinct categories, while the Midland and Northern talkers were less distinguishable. However, the adults with HFA confused the Northern talkers with all of the other dialects, whereas the TD adults confused the Northern talkers primarily with the Midland talkers.

Table 2 Means and standard deviations of % correct, % error, and overall accuracy for each population

	% Correct	% Error	% Correct–% Error (Overall accuracy)
Adults with HFA	28 (17)	14 (12)	14 (12)
TD adults	43 (21)	9 (6)	34 (18)

Phonetic Cues to Dialect Classification

To explore which properties of the stimulus materials the listeners were attending to when making their classification judgments, a series of multiple logistic regression analyses were conducted to predict cluster affiliation from the dialect-specific properties: r-lessness, "warsh" for *wash*, "greazy" for *greasy*, and /u/-fronting. For each analysis, the talkers in a given cluster were assigned a 1 and the talkers in the other clusters were assigned a 0. For example, in the TD analyses, New England 1 was assigned a 1 in the analysis of the New England cluster, but a 0 in the analyses of the Midland and Southern clusters. Given the relatively small number of observations (20), stepwise regression was

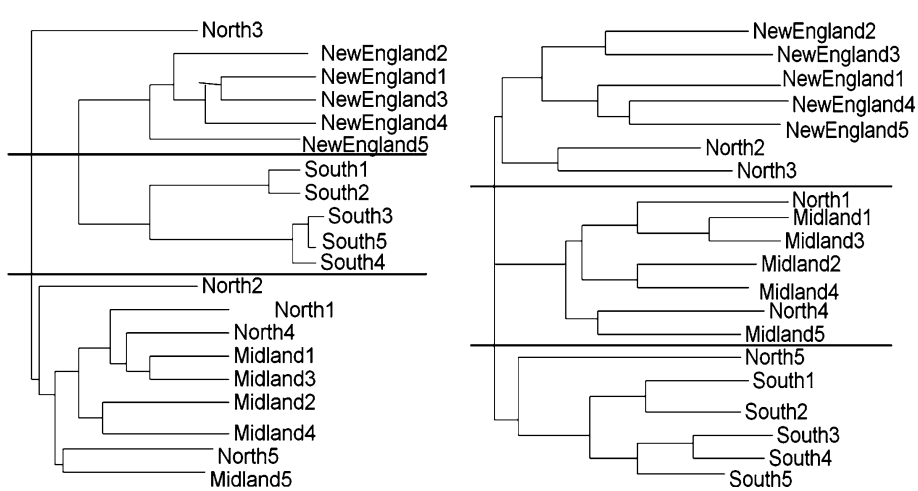


Fig. 2 Clustering solutions for the TD adults (*left*) and the adults with HFA (*right*)

used to identify at most two significant predictor variables in each analysis to avoid overfitting the data. For both populations, membership in the Southern cluster was significantly related to /u/-fronting in *suit* ($p = .03$ for TD, $p = .006$ for HFA) and "greazy" for *greasy* ($p = .003$ for TD, $p = .008$ for HFA), and membership in the Midland cluster was significantly related to "greasy" for *greasy* ($p = .008$ for TD, $p = .04$ for HFA). For the TD adults, membership in the New England cluster was related to r-lessness ($p = .001$), but this effect was only marginal for the adults with HFA ($p = .055$). Thus, the adults with HFA were attending to most of the same phonetic cues as the TD adults in the dialect classification task, but were not as effective at using r-lessness to distinguish the New England talkers from the other talkers.

Discussion

The listeners with HFA exhibited similar classification strategies to the TD listeners. Both groups of listeners produced approximately 6–7 groups of talkers with 3–4 talkers per group. This result suggests that the participants with HFA had a similar approach to the task as the TD listeners, although the listeners with HFA were significantly less accurate than the TD listeners. The clustering solutions in Fig. 2 suggest that the perceptual similarity structure of the talkers in the free classification task was similar for the listeners with HFA and the TD listeners. Both groups of listeners clearly distinguished the New England, Southern, and Midland talkers. However, the listeners with HFA were less consistent than the TD listeners in their perception of the Northern talkers. Whereas the TD listeners perceived most of the Northern talkers as highly similar to the Midland talkers, the listeners with HFA perceived two of the Northern talkers as similar to New England talkers, one as similar to Southern talkers, and two as similar to Midland talkers. The analysis of the phonetic cues suggests that the young adults with HFA were also less able to use r-lessness to distinguish the New England talkers from the other talkers than the TD adults.

Taken together, the results of the free classification task suggest that listeners with HFA can perceive the variation in the speech signal and attend to the relevant cues for classifying talkers by dialect. However, their performance differs significantly from TD listeners both in terms of overall accuracy and in the perceptual similarity of the Northern talkers relative to the other talkers. These findings are consistent with the previous research showing that people with ASD do not typically exhibit deficits in auditory processing (Happé and Frith 2006; Mottron et al. 2006), but that they have more difficulty with interpreting social information in the speech signal (Golan et al. 2007; Imaizumi et al. 2009). In addition, the relative success that the listeners with HFA exhibited in this task suggests that they have some knowledge about how social categories, such as regional background, are indexed in the speech signal and that they were able to ignore at least some voice-specific variability and focus on the dialect-specific variation that was crucial for classifying the talkers by dialect.

The free classification task does not require listeners to make explicit judgments about the geographic location associated with each of the dialect groups that they produce. To explore the relationship between dialect variation and geographic location, a localness judgment task was conducted. In the free classification task, the listeners with HFA were able to clearly distinguish the New England and Southern dialects from the local Midland dialect. The results from the localness judgment task will provide additional insight into how these dialect groups are mapped to geographic locations for listeners with HFA in Columbus, Ohio.

Experiment 2: Localness Judgment

Methods

Participants

Thirty TD adults (age range 18–28 years) participated in this experiment and an unrelated experiment to receive course credit in an introductory linguistics course. The order of the two experiments was counterbalanced across participants. Two TD adults were excluded because they were not native English speakers and 1 was excluded for being too much older than the rest of the participants (age 36). The 27 remaining participants (11 males, 16 females) were monolingual native speakers of American English with no reported history of hearing or speech disorders. Ten TD adults were from the Midland dialect region, 1 from the New England dialect region, 7 from the Northern dialect region, and 9 had moved between dialect regions before age 18. None of the TD adults had participated in Experiment 1.

Fourteen adults with HFA (age range 21–30 years) were paid to participate in this experiment and an unrelated experiment. The order of the two experiments was counterbalanced across participants. As in Experiment 1, the participants with HFA were self-identified, recruited from an outreach program specifically designed for people diagnosed with HFA, and scored significantly higher than the TD adults on the Autism-Spectrum Quotient (TD $M = 13.34$, SD = 4.96; HFA $M = 29.07$, SD = 8.37; $t(39) = 7.53$, $p < .001$, $d = 2.29$). All of the adults with HFA (11 males, 3 females) were monolingual native speakers of American English, but four reported a history of speech, language, or hearing disorders (one with speech apraxia, one who had seen a therapist as a child to help with articulation, one who currently sees a therapist to slow his speech and organize his thoughts before talking, and one with an unspecified disorder). Eight of the adults with HFA were from the Midland dialect region, 1 was from the Northern dialect region, and 5 had moved between dialect regions before age 18. Finally, 8 of the participants with HFA also participated in Experiment 1. However, as an average of 11 months had passed between their participation in the two experiments, it is unlikely that their responses were influenced by their previous experience with the stimulus materials.

Stimulus Materials

The same recordings of the 20 talkers reading the sentence "She had your dark suit in greasy wash water all year" were used in Experiment 2 as in Experiment 1. In the current experiment, however, the audio files were not linked to icons on the computer screen.

Procedure

Participants were instructed to listen to each talker and judge "How likely is it that this person comes from Columbus, Ohio?" (the location where the experiment was conducted) on a 5 point scale, where 1 was "not very likely" and 5 was "very likely." Each talker was rated twice. The talkers were presented randomly in two blocks so that each participant rated every talker once before rating all of the talkers a second time.

Results

A repeated measures ANOVA was conducted to examine the influence of population (TD and HFA) and talker dialect (New England, Northern, Midland, and Southern) on the listeners' localness ratings. The results revealed a significant main effect of talker dialect ($F(3,117) = 231.86$, $p < .001$, partial $\eta^2 = .86$). The population factor and the interaction were not significant. Post hoc paired-sample t tests showed that New England talkers were rated less likely to be from Columbus than Northern talkers ($t(40) = 22.37$, $p < .001$, $d = 4.48$), Midland talkers ($t(40) = 21.62$, $p < .001$, $d = 4.64$), and Southern talkers ($t(40) = 2.63$, $p < .05$, $d = .55$), and Southern talkers were rated less likely to be from Columbus than Northern talkers ($t(40) = -15.98$, $p < .001$, $d = 3.19$) and Midland talkers ($t(40) = -15.51$, $p < .001$, $d = 3.34$). The difference between the Midland and Northern talkers was not significant. The mean localness judgment ratings are shown in Fig. 3 for each talker dialect for each population.

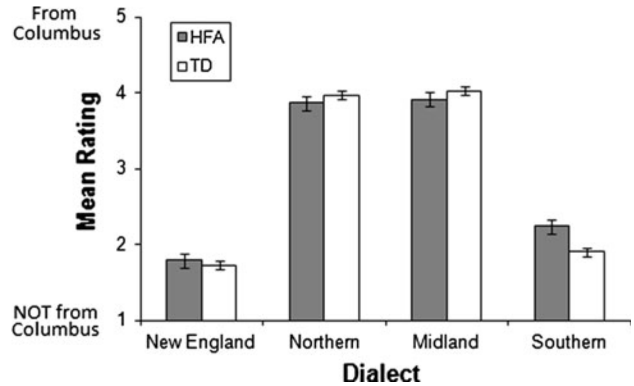


Fig. 3 Mean localness judgment ratings for each talker dialect for each population

To further verify that the observed effects reflected the performance of both listener groups and that the results from the larger TD group did not swamp the results from the smaller HFA group, planned comparisons examined the pattern of ratings separately for the HFA and TD participants. These comparisons revealed that the two groups exhibited very similar patterns of responses. One way ANOVAs showed a significant main effect of talker dialect on the TD group's ratings ($F(3,78) = 201.89$, $p < .001$, partial $\eta^2 = .89$) and on the HFA group's ratings ($F(3, 39) = 71.31$, $p < .001$, partial $\eta^2 = .85$). For the HFA group, the pattern of dialect differences was identical to the overall effects with the two populations combined. For the TD group, the pattern was similar to the overall results, except that the difference between the New England and Southern talkers was not significant.

Phonetic Cues to Localness Judgments

Multiple linear regression analyses were conducted to explore the phonetic cues that the two populations relied on to make their localness rating judgments. As in Experiment 1, stepwise regression was used to identify at most two significant predictor variables in each analysis. The TD adults relied on "greazy" for *greasy* to identify talkers as non-local ($p = .04$), but this effect was only marginal for the adults with HFA ($p = .056$). Thus, as in Experiment 1, the listeners with HFA were attending to appropriate dialect-specific cues in the speech signal to make their explicit judgments about the geographical background of the talkers, but they were less efficient in using those cues than the TD listeners.

Discussion

Performance in the localness judgment task was very similar for the two populations. Both the TD listeners and the listeners with HFA perceived the Midland and Northern talkers as more likely to be from Columbus than the New England and Southern talkers. These results are consistent with the free classification task, which showed that the listeners perceived the New England and Southern dialects as clearly distinct from the Midland dialect. In addition, for the TD listeners, the Midland and Northern talkers were clustered together in the free classification task and were perceived as equally likely to be from Columbus in the localness judgment task. For the listeners with HFA, the Northern talkers were distributed more randomly among the other dialects in the free classification task, but their perceptual similarity to the Midland talkers was revealed in the localness judgment task in which the Northern and Midland talkers were rated as equally likely to be from Columbus. Thus, the difference in the perception of the Northern talkers by the two populations that was observed in the free classification task was not observed in the localness judgment task. However, the listeners with HFA exhibited a perceptual distinction between the New England and Southern talkers in the localness judgment task that the TD listeners did not. Whereas the TD listeners perceived the New England and Southern talkers as similarly unlikely to be from Columbus, the listeners with HFA perceived the Southern talkers as significantly more likely to be from Columbus than the New England talkers. This result is particularly interesting because it reflects the actual geographic relationships between the dialects. For listeners in Ohio, varieties of Southern American English are spoken in neighboring states, including West Virginia and Kentucky, whereas Pennsylvania serves as a buffer between Ohio and the western edge of the New England dialect region in eastern New York State.

The results of the localness judgment task suggest that the listeners with HFA can map the dialect-specific information in the speech signal to geographic distance from the local area. Thus, not only can they attend to dialect-specific information to group talkers by dialect in the free classification task, they also exhibit some knowledge about how dialect variability in the speech signal indexes a talker's regional background. The language attitudes task in Experiment 3 was conducted to examine another kind of knowledge that is associated with dialect variation for TD listeners: social stereotypes. If the listeners with HFA have internalized the local stereotypes associated with regional dialects of American English, their performance in the language attitudes task should be similar to the performance of TD listeners. However, if their knowledge of regional dialect variation is limited to factual geographic information, they may not clearly differentiate between the dialects in the language attitudes task.

Experiment 3: Language Attitudes

Methods

Listeners

Thirty-one TD adults (age range 18–23 years) participated in this experiment and an unrelated experiment to receive course credit in an introductory linguistics course. The order of the two experiments was counterbalanced across participants. Three TD adults were excluded because they were not native English speakers. The remaining 28 participants (6 males, 22 females) were monolingual native speakers of American English with no reported history of hearing or speech disorders. Fourteen TD participants were from the Midland dialect region, 6 from the Northern dialect region, 2 from the Southern dialect region, and 6 had moved between dialect regions before age 18. None of the TD adults had previously participated in either Experiment 1 or Experiment 2.

Fourteen adults with HFA (age range 19–28 years) were paid to participate in this experiment and an unrelated experiment. The order of the two experiments was counterbalanced across participants. As in Experiments 1 and 2, the adults with HFA were self-identified, recruited from an outreach program specifically designed for people diagnosed with HFA, and scored significantly higher than the TD participants on the Autism-Spectrum Quotient (TD $M = 15.39$, SD $= 4.07$; HFA $M = 28.89$, SD $= 7.86$; $t(41) = 6.84$, $p < .001$, $d = 2.08$). All of the participants with HFA (11 males, 3 females) were monolingual native speakers of American English and 4 reported a history of speech, language, or hearing disorders (one who currently sees a therapist for articulation and three with unspecified disorders). Ten adults with HFA were from the Midland dialect region, 1 was from the Northern dialect region, and 3 had moved between dialect regions before age 18. Finally, 6 of the HFA participants had previously participated in both Experiment 1 and Experiment 2, and 5 of the HFA participants had previously participated in Experiment 2 only. However, given that there was a long temporal gap between the studies (13 months on average between Experiments 1 and 3 and 2 months on average between Experiments 2 and 3), it is unlikely that their responses in this experiment were influenced by their previous experience with the stimulus materials.

Stimulus Materials

The same recordings of the 20 talkers reading the sentence "She had your dark suit in greasy wash water all year" were used in Experiment 3 as in Experiment 2. Recordings from an additional set of white male talkers ($N = 15$)

reading the same stimulus sentence were also included in this study. Responses to these additional talkers were analyzed as part of a separate project. The current analysis is limited to the 20 talkers also used in Experiments 1 and 2.

Procedure

There were four blocks of trials in Experiment 3: judging whether each talker was friendly, reliable, intelligent, and successful. In each block, participants were instructed to listen to each talker and judge "How friendly/reliable/intelligent/successful do you think this person is?" on a 5 point scale, where 1 was "not very" and 5 was "very." Each participant rated each of the talkers one time in each of the four blocks. The blocks and the talkers within the blocks were presented in a different random order for each listener. The extent to which a talker is rated as friendly and reliable is considered a measure of solidarity, whereas the extent to which a talker is rated as intelligent and successful is considered a measure of status.

Results

Friendliness Ratings

A repeated measures ANOVA was conducted to examine the influence of population (TD and HFA) and talker dialect (New England, Northern, Midland, and Southern) on the listeners' friendliness ratings. The results showed a significant main effect of population ($F(1, 40) = 4.88$, $p < .05$, partial $\eta^2 = .11$), as the participants with HFA consistently rated all talkers as more friendly than the TD participants did. A significant main effect was also found for talker dialect ($F(3,120) = 4.72$, $p < .01$, partial $\eta^2 = .11$). The interaction was not significant. Post hoc paired-sample t tests were conducted to examine the main effect of talker dialect and revealed that New England talkers were rated as less friendly than Midland talkers ($t(41) = 2.40$, $p < .05$, $d = .37$) and Northern talkers were rated as less friendly than Midland ($t(41) = 3.90$, $p < .001$, $d = .66$) and Southern ($t(41) = 2.21$, $p < .05$, $d = .39$) talkers. None of the other dialect differences were significant. The mean friendliness ratings are shown in the top left panel of Fig. 4 for each talker dialect for each population.

As in Experiment 2, planned comparisons were conducted to verify that the observed main effect of talker dialect reflected the performance of both listener groups and that the results from the larger TD group did not swamp the results from the smaller HFA group. A one-way ANOVA examining only the TD participants' friendliness ratings revealed a significant main effect of talker dialect ($F(3,81) = 4.32$, $p < .01$, partial $\eta^2 = .14$) and the differences between the dialects were qualitatively identical to the overall results with the two populations combined. By contrast, when the HFA participants were considered on their own, they did not differentiate among the talker dialects ($F(3, 39) = 2.72$, ns). The listeners with HFA rated the talkers of all dialects as equally friendly.

Reliability Ratings

The results for the reliability measure were parallel to the results for friendliness, the other solidarity measure. A repeated measures ANOVA was conducted to examine the influence of population (TD and HFA) and talker dialect (New England, Northern, Midland, and Southern) on the listeners' reliability ratings. The results showed a significant main effect of population ($F(1,40) = 5.99$, $p < .05$, partial $\eta^2 = .13$), as the participants with HFA consistently rated all of the talkers as more reliable than the TD participants did. A significant main effect was also found for talker dialect ($F(3,120) = 8.47$, $p < .001$, partial $\eta^2 = .18$). The interaction was not significant. Post hoc paired-sample t tests were conducted to explore the main effect of talker dialect and showed that Midland talkers were rated as more reliable than New England talkers ($t(41) = 6.18$, $p < .001$, $d = .99$), Northern talkers ($t(41) = 3.01$, $p < .01$, $d = .38$), and Southern talkers ($t(41) = 2.91$, $p < .01$, $d = .57$), and Northern talkers were rated as more reliable than New England talkers ($t(41) = 5.00$, $p < .001$, $d = .64$). None of the other dialect differences were significant. The mean reliability ratings are shown in the top right panel of Fig. 4 for each talker dialect for each population.

Planned comparisons examined whether HFA and TD participants showed differences in their patterns of ratings. A one-way ANOVA on the TD participants' reliability ratings revealed a significant main effect of talker dialect ($F(3, 81) = 9.11$, $p < .001$, partial $\eta^2 = .25$) and the differences between the dialects were qualitatively identical to the overall results with the combined populations. By contrast, when the participants with HFA were considered on their own, they did not differentiate among the talker dialects ($F(3, 39) = 2.51$, ns). As was found for the friendliness ratings, the listeners with HFA rated the talkers of all dialects as equally reliable.

Intelligence Ratings

A repeated measures ANOVA was conducted examining the influence of population (TD and HFA) and talker dialect (New England, Northern, Midland, and Southern) on the listeners' intelligence ratings. The results showed a significant main effect of population ($F(1,40) = 5.21$,

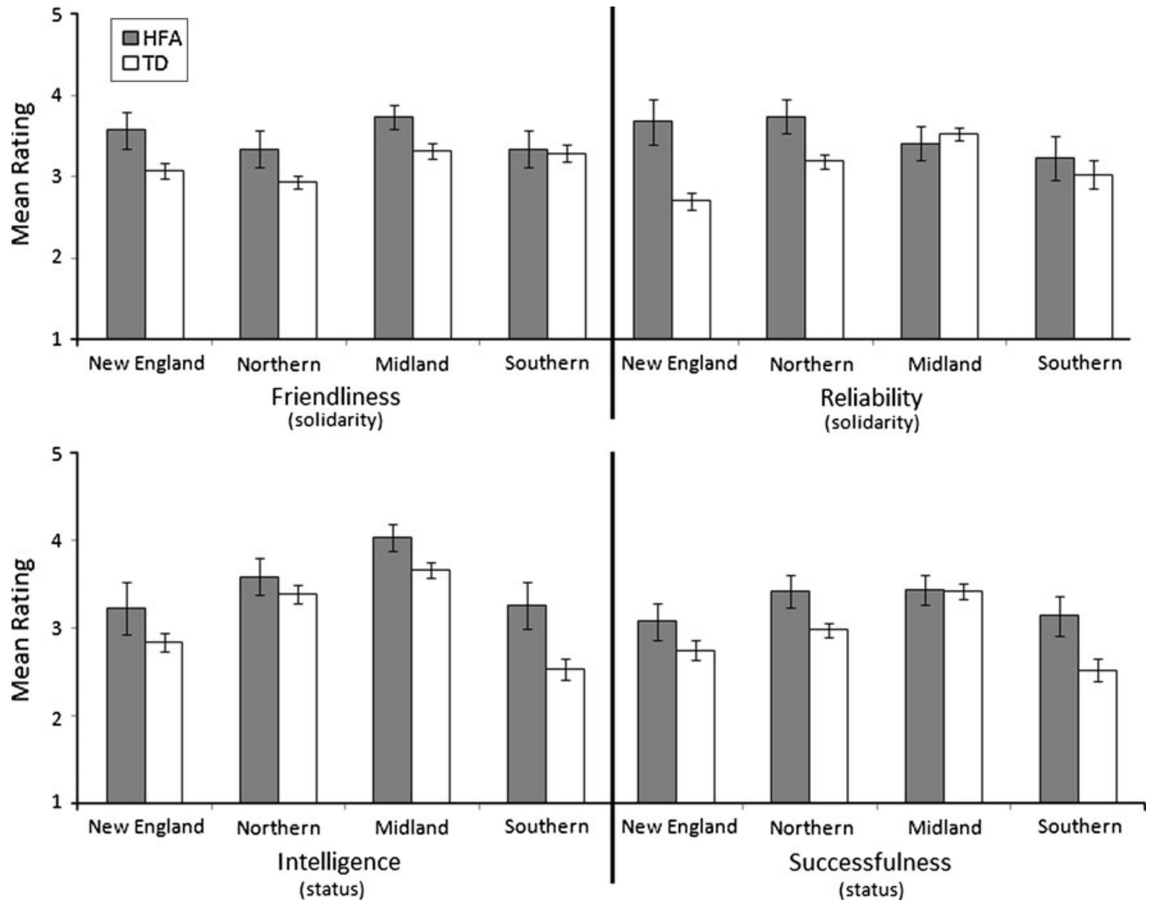


Fig. 4 Mean language attitude ratings for each talker dialect for each population

$p < .05$, partial $\eta^2 = .12$), as the participants with HFA consistently rated all talker dialects as more intelligent than the TD participants did. A significant main effect was also found for talker dialect ($F(3,120) = 30.93$, $p < .001$, partial $\eta^2 = .44$). The interaction was not significant. Post hoc paired-sample *t* tests exploring the main effect of talker dialect showed that New England talkers were rated as less intelligent than Northern talkers ($t(41) = 4.55$, $p < .001$, $d = .67$) and Midland talkers ($t(41) = 7.48$, $p < .001$, $d = 1.17$), but more intelligent than Southern talkers ($t(41) = 2.04$, $p < .05$, $d = .22$). Northern talkers were rated as less intelligent than Midland talkers ($t(41) = 4.23$, $p < .001$, $d = .58$) and more intelligent than Southern talkers ($t(41) = 5.63$, $p < .001$, $d = .90$). Further, Midland talkers were rated as more intelligent than Southern talkers ($t(41) = 8.54$, $p < .001$, $d = 1.41$). The mean intelligence ratings are shown in the bottom left panel of Fig. 4 for each talker dialect for each population.

Planned comparisons were conducted to examine whether HFA and TD participants showed differences in their patterns of ratings. A one-way ANOVA on the TD participants' intelligence ratings revealed a significant main effect of talker dialect ($F(3,81) = 31.16$, $p < .001$, partial $\eta^2 = .54$) and the differences between the dialects were qualitatively identical to the overall analysis with the combined populations. Unlike for the friendliness and reliability ratings, for the intelligence ratings, the participants with HFA also exhibited a significant main effect of talker dialect ($F(3, 39) = 9.13$, $p < .01$, partial $\eta^2 = .41$). More specifically, Midland talkers were rated as more intelligent than Southern talkers ($t(13) = 4.08$, $p < .01$, $d = .89$), New England talkers ($t(13) = 3.84$, $p < .01$, $d = .85$), and Northern talkers ($t(13) = 3.89$, $p < .01$, $d = .62$). None of the other dialect differences were significant. Thus, in contrast to the solidarity attributes of friendliness and reliability, the adults with HFA were able to link the different dialects to stereotypes about the status feature of intelligence.

Successfulness Ratings

A repeated measures ANOVA was conducted to examine the influence of population (TD and HFA) and talker dialect (New England, Northern, Midland, and Southern) on the listeners' successfulness ratings. The results showed a significant main effect of population ($F(1,40) = 6.14$,

$p < .05$, partial $\eta^2 = .13$), as the participants with HFA consistently rated all talker dialects as more successful than the TD participants did. A significant main effect was also found for talker dialect ($F(3,120) = 10.90$, $p < .001$, partial $\eta^2 = .22$). The interaction was not significant. Post hoc paired-sample *t* tests were conducted to explore the main effect of talker dialect and showed that New England talkers were rated as less successful than Northern talkers ($t(41) = 2.70$, $p < .05$, $d = .45$) and Midland talkers ($t(41) = 4.83$, $p < .001$, $d = .99$), and as successful as Southern talkers ($t(41) = 1.25$, ns). Northern talkers were rated as less successful than Midland talkers ($t(41) = -3.48$, $p < .01$, $d = .56$) and more successful than Southern talkers ($t(41) = 3.40$, $p < .01$ $d = .59$). Further, Midland talkers were rated as more successful than Southern talkers ($t(41) = 4.93$, $p < .001$ $d = 1.04$). The mean successfulness ratings are shown in the bottom right panel of Fig. 4 for each talker dialect for each population.

Planned comparisons examined whether HFA and TD participants showed differences in their patterns of ratings. A one-way ANOVA on the TD participants' successfulness ratings revealed a significant main effect of talker dialect ($F(3,81) = 14.20$, $p < .001$, partial $\eta^2 = .35$). The dialect differences were similar to the overall results with both populations combined, except that the difference between the New England and Northern talkers was not significant. By contrast, when the HFA participants were considered on their own, they did not differentiate among the talker dialects ($F(3, 39) = 2.60$, ns). The listeners with HFA rated the talkers of all dialects as equally successful.

Phonetic Cues to Language Attitude Ratings

Multiple linear regression analyses were conducted to explore the phonetic cues that the two populations relied on to make their language attitude judgments. As in Experiment 2, stepwise regression was used to identify at most two significant predictor variables in each analysis. No significant predictors emerged in the analysis of the solidarity measures, friendliness and reliability, for either the TD listeners or the listeners with HFA. However, in the analysis of the status measures, intelligence and successfulness, "greasy" for *greasy* emerged as a significant predictor of higher ratings for the TD listeners ($p = .01$ for intelligence, $p = .03$ for successfulness). No significant predictors were observed for the listeners with HFA for the two status measures.

Discussion

The language attitudes task revealed more differences between the two populations than the free classification and localness judgment tasks. The TD listeners made more distinctions between the dialects than the listeners with HFA. For the friendliness, reliability, and successfulness ratings, the listeners with HFA rated all four of the talker dialects similarly, whereas the TD listeners distinguished between the dialects in their ratings. For the intelligence ratings, both populations' ratings distinguished between the dialects, but the TD listeners made more distinctions than the listeners with HFA. The listeners with HFA also produced higher mean ratings than the TD listeners on all four scales. This difference in overall rating may reflect differences across the two groups in the use of the scales that are not particularly interesting or meaningful. However, given that the participants with HFA were recruited from a group designed to help with developing appropriate social skills, the listeners with HFA may have been less willing than the TD listeners to rate any of the talkers negatively in an experimental laboratory setting. This interpretation is independent of the finding that the listeners with HFA did not distinguish between the talker dialects in their language attitude ratings. Although the listeners with HFA used the higher end of the ratings scale, their ratings were not at ceiling and they used a sufficient range within the scale to allow us to observe significant differences between the talker dialects in the intelligence rating task.

For the solidarity ratings, the Midland talkers were rated highest by the TD listeners in terms of both friendliness and reliability, consistent with previous research demonstrating higher solidarity ratings for local dialects than non-local dialects (e.g., Luhman 1990). However, the ratings for the other three dialects differed for the two solidarity rating scales. The TD listeners rated the Southern talkers as more friendly than the Northern and New England talkers, but the Northern talkers as more reliable than the Southern and New England talkers. This difference may reflect social stereotypes about the northern and southern regions of the United States that extend beyond local solidarity. The rating of Southern talkers as friendly may reflect general stereotypes of southern hospitality, whereas the rating of Northern talkers as reliable may reflect general stereotypes of midwestern dependability. For the listeners with HFA, no significant differences between the dialects were observed for either of the solidarity measures, suggesting that they do not have access to either solidarity with local talkers or the stereotypes associated with other regions that are shared by the TD listeners. None of the phonetic cues were significantly associated with the solidarity ratings for either listener group, suggesting that their responses may have been based on other aspects of the signal that were not included as predictor variables in the regression models.

For the status ratings, the Midland talkers were rated highest by the TD listeners in terms of both intelligence and successfulness, consistent with previous research demonstrating higher status ratings for more socially prestigious varieties than less prestigious varieties (e.g., Luhman 1990). However, the TD listeners also produced more distinctions among the dialects for the intelligence ratings than the successfulness ratings. After the Midland talkers, the Northern talkers were rated as most intelligent, followed by the New England talkers, with the Southern talkers rated as least intelligent. For the successfulness ratings, the only non-Midland dialects that received significantly different ratings were the North and South, with the Northern talkers rated as more successful than the Southern talkers. For the listeners with HFA, the Midland talkers were also rated as significantly more intelligent than the other three dialect groups, but no differences were observed across dialects for the successfulness ratings. Thus, the listeners with HFA appear to have some representation of the Midland dialect as a prestigious dialect that is spoken by more intelligent people.

For all four of the rating scales, the TD listeners rated the Midland talkers significantly higher than the Northern talkers, suggesting that the TD listeners were able to perceptually distinguish between the two talker dialects, despite classifying them together in the free classification task and rating them equally in the localness judgment task. Similarly, the listeners with HFA rated the Midland talkers as significantly more intelligent than the Northern talkers, despite also rating them equally in the localness judgment task. These results suggest that the explicit classification tasks are fundamentally different from the language attitudes task and that listeners use the information in the signal differently in the two kinds of tasks. Whereas in the language attitudes task, the listeners in both groups demonstrated the ability to distinguish between the Midland and Northern dialects for the purpose of rating the personal attributes of the talkers, they rated the Midland and Northern talkers as equally likely to be from Columbus in the localness judgment task. Thus, the Midland and Northern dialects may map to different social varieties, rather than different regional varieties, for listeners in Columbus. Both dialects are classified as local, but the more socially prestigious Midland dialect is perceived as higher in status and more socially attractive than the Northern dialect.

General Discussion

These studies examined the ability of young adults with HFA to perceive indexical information related to regional dialect. Experiment 1 used a free classification paradigm and found that participants with HFA were able to group talkers of different regional dialects according to where they were from using the same phonetic properties as the TD participants. Participants with HFA were somewhat less accurate than their TD peers in this task, but still succeeded overall in creating appropriate dialect groups. In Experiment 2, participants with HFA showed further that they can explicitly judge talkers of these dialects as being from or not from the local area. In this task, participants with HFA were virtually identical to their TD peers. Finally, in Experiment 3, participants with HFA differed from their TD peers by not associating appropriate social stereotypes to the different dialect groups. While the TD participants rated the local dialect as being more favorable in terms of solidarity and status, the participants with HFA exhibited this bias for the intelligence ratings only. Moreover, the TD participants showed more general knowledge of regional stereotypes (e.g., people with southern dialects are more friendly), but the adults with HFA failed to differentiate the dialects in these terms.

Overall, these results confirm the proposed hypotheses, and show that while young adults with HFA can perceive the acoustic differences among dialects, and can even make a basic factual judgment about them (i.e., is this the way people around here talk?), they have not linked this aspect of the speech signal to all of the relevant social information it conveys. These findings are consistent with the previous literature demonstrating that listeners with HFA can make accurate factual judgments about talkers, such as identifying familiar talkers based on their speech (Boucher et al. 2000). This interpretation is also consistent with the finding that adolescents with ASD adopt the variety of their mothers instead of their peers. Baron-Cohen and Staunton (1994) argued that the adolescents with ASD in their study did not identify socially with their peers, and therefore were not motivated to use their peers' dialect. The current results suggest that the problem may be somewhat deeper than Baron-Cohen and Staunton suggested. The process of using speech to index one's own social identity pre-supposes the knowledge that speech can be used to index social meaning. The participants with HFA in the current study showed very little understanding of how dialect variation in speech is linked to social stereotypes among the TD population. If they do not appreciate the social indexical information that others convey through their speech, there is little reason to expect them to use it to convey their own social identity.

The difficulties that the participants with HFA exhibited in these tasks were largely restricted to the social domain. When participants with HFA were asked to link dialect information in the speech signal to indexical facts about geographical area (Experiments 1 and 2), their performance was similar to the performance of TD adults. The largest differences between the two populations emerged when the task required them to link the indexical information in the speech signal to social stereotypes (Experiment 3). The narrow focus of the difficulties exhibited by the participants with HFA in these studies strongly suggests that it is this particular type of information that poses special difficulties for people with HFA. These results therefore support the general view that specifically social deficits are central to the ASD profile.

The current results, however, should be interpreted with some degree of caution. The participants with HFA were recruited through a program catering to that population (see the methods in Experiment 1 for details) and were self-identified as being on the Autism spectrum. However, no additional formal diagnosis was obtained. Moreover, the participants' scores on the Autism-Spectrum Quotient—while significantly higher than those of our TD participants—were not firmly within the range associated with ASD, particularly in Experiment 1 (but see Ketelaars et al. 2008). A more conservative interpretation of our results would therefore be that people with higher scores on the Autism-Spectrum Quotient show a profile of dialect perception that is consistent with the general profile of ASD.

One additional potential point of concern with the current study is that although each experiment was conducted with a separate set of TD participants, some of the participants with HFA participated in more than one of the experiments, and 6 participated in all three of them (see the methods sections for Experiments 2 and 3 for details). Thus, the differences we observed between the two populations may reflect differences in their experience with the stimulus materials, which were the same in all three experiments. However, the time between the experiments ranged from 1 month, 29 days (between Experiments 2 and 3) to 12 months, 4 days (between Experiments 1 and 2), making it rather unlikely that the participants were able to connect their responses in one task to their responses in a different task. In addition, the task requirements were different in the three experiments (classification, localness ratings, and language attitudes), making practice effects with the stimulus materials less likely. Finally, if our participants with HFA were learning something about the talkers or the dialects through engagement with the stimulus materials, we would expect to observe more similarities between the two populations in the later studies than in the earlier studies. Our data reveal the opposite pattern: performance by the participants with HFA was substantially less like the performance by the TD adults in Experiment 3 than in Experiments 1 and 2.

A related concern is that performance may have been affected by the repeated exposure to the same sentence over the course of each experiment, either due to practice or fatigue effects. However, given that the trials were randomized separately for each participant, any effects of practice or fatigue should wash out across participants. In addition, Clopper and Pisoni (2004b) found comparable performance in a forced-choice dialect classification task when the stimulus materials consisted of many talkers reading the target sentence in the current study and when they consisted of the same talkers each reading a unique sentence. We therefore would not expect to observe substantially different results in the current experiments if a different sentence were used for each talker.

In summary, these experiments demonstrated that young adults with HFA are able to perceive the phonetic correlates of dialect variation and are able to group and judge talkers of various dialects with respect to their geographical background. However, they are less able to judge talkers of various dialects with respect to the social stereotypes associated with their dialect. These results are consistent with established results of the broader phenotype of ASD: the participants had access to perceptual information, but failed to use it to support social reasoning.

Acknowledgments This research was partially funded by a seed grant from the Ohio State University Center for Cognitive Science. We thank Jeff Siegel, Bridget Smith, and Renee Devlin for their help with this research.

References

Alcántara, J. I., Weisblatt, E. J. L., Moore, B. C. J., & Bolton, P. F. (2004). Speech-in-noise perception in high-functioning individuals with autism or Asperger's syndrome. *Journal of Child Psychology and Psychiatry, 45*, 1107–1114.

Baker, W., Eddington, D., & Nay, L. (2009). Dialect identification: The effects of region of origin and amount of experience. *American Speech, 84*, 48–71.

Baron-Cohen, S., & Staunton, R. (1994). Do children with autism acquire the phonology of their peers? An examination of group identification through the window of bilingualism. *First Language, 14*(42, pt 3), 241–248.

Baron-Cohen, S., Tager-Flusberg, H., & Cohen, D. J. (2000). *Understanding other minds: Perspectives from developmental cognitive neuroscience*. Oxford: Oxford University Press.

Baron-Cohen, S., Wheelwright, S., Skinner, R., Martin, J., & Clubley, E. (2001). The autism-spectrum quotient (AQ): Evidence from Asperger syndrome/high-functioning autism, males and females, scientists and mathematicians. *Journal of Autism and Developmental Disorders, 31*(1), 5–17.

Bishop, D. V. M., Maybery, M., Wong, D., Maley, A., Hill, W., & Hallmayer, J. (2004). Are phonological processing deficits part of the broad autism phenotype? *American Journal of Medical Genetics Part B (Neuropsychiatric Genetics), 128B*, 54–60.

Boucher, J., Lewis, V., & Collis, G. M. (2000). Voice processing abilities in children with autism, children with specific language impairments, and young typically developing children. *Journal of Child Psychology and Psychiatry, 41*, 847–857.

Chambers, J. (2002). Dynamics of dialect convergence. *Sociolinguistics, 6*, 117–130.

Clopper, C. G., & Bradlow, A. R. (2009). Free classification of American English dialects by native and non-native listeners. *Journal of Phonetics, 37*, 436–451.

Clopper, C. G., & Pisoni, D. B. (2004a). Homebodies and army brats: Some effects of early linguistic experience and residential history on dialect categorization. *Language Variation and Change, 16*, 31–48.

Clopper, C. G., & Pisoni, D. B. (2004b). Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics, 32*, 111–140.

Clopper, C. G., & Pisoni, D. B. (2006). Effects of region of origin and geographic mobility on perceptual dialect categorization. *Language Variation and Change, 18*, 193–221.

Clopper, C. G., & Pisoni, D. B. (2007). Free classification of regional dialects of American English. *Journal of Phonetics, 35*, 421–438.

Constantino, J. N., Yang, D., Gray, T. L., Gross, M. M., Abbacchi, A. M., Smith, S. C., et al. (2007). Clarifying the associations between language and social development in autism: A study of non-native phoneme recognition. *Journal of Autism and Developmental Disorders, 37*(7), 1256–1263.

Corter, J. E. (1982). ADDTREE/P: A PASCAL program for fitting additive trees based on Sattath and Tversky's ADDTREE algorithm. *Behavior Research Methods and Instrumentation, 14*, 353–354.

Fisher, W. M., Doddington, G. R., & Goudie-Marshall, K. M. (1986). The DARPA speech recognition research database: Specifications and status. *Proceedings of the DARPA speech recognition workshop* (pp. 93–99).

Gallois, C., & Callan, V. (1989). Attitudes to spoken Australian English: Judgments of ingroup and ethnic outgroup speakers. *Australian Journal of Linguistics, 9*, 149–160.

Giles, H. (1970). Evaluative reactions to accents. *Educational Review, 22*, 211–227.

Golan, O., Baron-Cohen, S., Hill, J. J., & Rutherford, M. D. (2007). The 'reading the mind in the voice' test-revised: A study of complex emotion recognition in adults with and without Autism Spectrum conditions. *Journal of Autism and Developmental Disorders, 37*, 1096–1106.

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders, 36*(1), 5–25.

Hubbard, K., & Trauner, D. A. (2007). Intonation and emotion in autistic spectrum disorders. *Journal of Psycholinguistic Research, 36*(2), 159–173.

Imaizumi, S., Furuya, I., & Yamasaki, K. (2009). Voice as a tool communicating intentions. *Logopedics Phoniatrics Vocology, 34*(4), 196–199.

Jarvinen-Pasley, A., Pasley, J., & Heaton, P. (2008a). Is the linguistic content of speech less salient than its perceptual features in autism? *Journal of Autism and Developmental Disorders, 38*(2), 239–248.

Jarvinen-Pasley, A., Wallace, G. L., Ramus, F., Happé, F., & Heaton, P. (2008b). Enhanced perceptual processing of speech in autism. *Developmental Science, 11*(1), 109–121.

Ketelaars, C., Horowitz, E., Sytema, S., Bos, J., Wiersma, D., Minderaa, R., et al. (2008). Brief report: Adults with mild autism spectrum disroders (ASD): Scores on the autism spectrum quotient (AQ) and comorbid psychopathology. *Journal of Autism and Developmental Disorders, 38*, 176–180.

Klatt, D. H. (1989). Review of selected models of speech perception. In W. Marslen-Wilson (Ed.), *Lexical representation and process* (pp. 169–226). Cambridge, MA: MIT Press.

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

Ladegaard, H. J. (1998a). National stereotypes and language attitudes: The perception of British, American and Australian language and culture in Denmark. *Language & Communication, 18*, 251–274.

Ladegaard, H. J. (1998b). Assessing national stereotypes in language attitude studies: The case of class-consciousness in Denmark. *Journal of Multilingual and Multicultural Development, 19*, 182–197.

Lambert, W. E., Hodgson, R. C., Gardner, R. C., & Fillenbaum, S. (1960). Evaluational reactions to spoken language. *Journal of Abnormal and Social Psychology, 60*, 44–51.

Lass, N. J., Hughes, K. R., Bowyer, M. D., Waters, L. T., & Bourne, V. (1976). Speaker sex identification from voiced, whispered, and filtered isolated vowels. *Journal of the Acoustical Society of America, 59*, 675–678.

Lass, N. J., Tecca, J. E., Mancuso, R. A., & Black, W. I. (1979). The effect of phonetic complexity on speaker race and sex identification. *Journal of Phonetics, 7*, 105–118.

Luhman, R. (1990). Appalachian English stereotypes: Language attitudes in Kentucky. *Language in Society, 19*, 331–348.

McCann, J., Peppe, S., Gibbon, F. E., O'Hare, A., & Rutherford, M. (2007). Prosody and its relationship to language in school-aged children with high-functioning autism. *International Journal of Language & Communication Disorders, 42*(6), 682–702.

Mottron, L., Dawson, M., Soulières, I., Hubert, B., & Burack, J. (2006). Enhanced perceptual functioning in autism: An update, and eight principles of autistic perception. *Journal of Autism and Developmental Disorders, 36*(1), 27–43.

Nober, E. H., & Simmons, J. Q. (1981). Comparison of auditory stimulus processing in normal and autistic adolescents. *Journal of Autism and Developmental Disorders, 11*, 175–189.

Paul, R., Augustyn, A., Klin, A., & Volkmar, F. R. (2005). Perception and production of prosody by speakers with autism spectrum disorders. *Journal of Autism and Developmental Disorders, 35*(2), 205–220.

Pollack, I., Pickett, J. M., & Sumby, W. H. (1954). On the identification of speakers by voice. *Journal of the Acoustical Society of America, 26*, 403–406.

Preston, D. R. (1993). Folk dialectology. In D. R. Preston (Ed.), *American dialect research* (pp. 333–378). Philadelphia: John Benjamins.

Ptacek, P. H., & Sander, E. K. (1966). Age recognition from voice. *Journal of Speech and Hearing Research, 9*, 273–277.

Purnell, T., Idsardi, W., & Baugh, J. (1999). Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology, 18*, 10–30.

Rice, M. L., Warren, S. F., & Betz, S. K. (2005). Language symptoms of developmental language disorders: An overview of autism, Down syndrome, fragile X, specific language impairment, and Williams syndrome. *Applied Psycholinguistics, 26*(1), 7–27.

Van Bezooijen, R., & Gooskens, C. (1999). Identification of language varieties: The contribution of different linguistic levels. *Journal of Language and Social Psychology, 18*, 31–48.

Van Lancker, D., Kreiman, J., & Emmorey, K. (1985). Familiar voice recognition: Patterns and parameters part I: Recognition of backward voices. *Journal of Phonetics, 13*, 19–38.

Volden, J., Magili-Evans, J., Goulden, K., & Clarke, M. (2007). Varying language register according to listener needs in speakers with autism spectrum disorder. *Journal of Autism and Developmental Disorders, 37*, 1139–1154.

Volden, J., & Sorenson, A. (2009). Bossy and nice requests: Varying language register in speakers with autism spectrum disorder (ASD). *Journal of Communication Disorders, 42*, 58–73.

Williams, A., Garrett, P., & Coupland, N. (1999). Dialect recognition. In D. R. Preston (Ed.), *Handbook of perceptual dialectology* (pp. 345–358). Philadelphia: John Benjamins.

Woodbury-Smith, M. R., Robinson, J., Wheelwright, S., & Baron-Cohen, S. (2005). Screening adults for Asperger Syndrome using the AQ: A preliminary study of its diagnostic validity in clinical practice. *Journal of Autism and Developmental Disorders, 35*, 331–335.