AMERICAN PSYCHOLOGICAL ASSOCIATION

Journal of Experimental Psychology: General

ISSN: 0096-3445

2024, Vol. 153, No. 11, 2918–2930
https://doi.org/10.1037/xge0001560

Learning Linguistic Diversity: Listeners Have Race-Based Linguistic Expectations, but Only for Phonological Variation

Alexandra M. Ryken, Emma Tupper, and Drew Weatherhead
Department of Psychology and Neuroscience, Dalhousie University

In three artificial language experiments, we explored the rate at which adults learned associations between linguistic variation and speaker characteristics. Within each of the experiments, we observed that listeners sociolinguistic learning occurred, regardless of whether the speaker characteristic is social (race and sex/gender) or nonsocial (hat wearing), or whether they heard a phonological or morphological variant. However, we found that listener's initial expectations of what social properties were predictive of linguistic variation differed, impacting overall performance. First, participants were much more likely to assume that a phonological variant was predicted by a social property than a nonsocial property (Experiment 1). Most interestingly, participants were more likely to privilege speaker race than sex/gender, but only in the case of a phonological variant (Experiments 2 and 3). The same effect was found in both White and Black participants, though White participants were more likely to correctly articulate which speaker characteristic explained the variation, suggesting that sociolinguistic learning hinges on real-world experiences with language and social diversity.

Public Significance Statement

Sociolinguistic associations are formed when listeners link specific speech properties (e.g., accents and suffix use) with a specific speaker property (e.g., race and sex/gender), which in turn can influence their expectations of how a new person will speak. The present study suggests that some sociolinguistic associations may be learned more easily than others. To be more precise, listeners expect that speaker race is important to how speakers pronounce words, at least when compared to the social property sex/gender. Furthermore, the background and experiences of the listener may influence how they learn sociolinguistic associations. That is, the listener's own race or their experience with race may affect how quickly they form an association between a speaker's race and how they pronounce words. The results of this study highlight the importance of listeners' previous experiences with diversity on their future interactions with new speakers.

Keywords: sociolinguistic learning, social cognition, speech perception

With increasing diversity across western societies, people are interacting with more individuals from different linguistic, racial, and cultural backgrounds than ever before. Of particular importance to our everyday life is linguistic diversity. Within English-Speaking North America there are a multitude of dialects and accents that a language learner may encounter. For example, in the United States and Canada, there are numerous dialects of English (e.g., Labov et al., 2006), though the exact number varies across scholars and publications. While accent information is not collected in the census, recent data in the United States suggest that approximately 22% of Americans speak a language other than English in the home, and 13% of the population are "foreign born" (United States Census Bureau, 2022). In Canada, one in four households speak a language other than English and/or French in the home, and 22% of the population were born in a country other than Canada (Statistics Canada, 2023). While not an exact metric of accent/dialect variation, it gives

This article was published Online First March 21, 2024.
Drew Weatherhead https://orcid.org/0000-0002-3649-2030
This work was funded by the Natural Sciences and Engineering Research Council of Canada (RGPIN-2021-02740) awarded to Drew Weatherhead. The authors thank Melissa Meade for designing the characters used in our stimuli, Abby Green for her contributions to stimuli development for Experiment 1, and all other research assistants in the Cognition and Language Learning Research Group for their assistance. The ideas and data presented in this article have not been presented or published previously. All materials and data can be found at https://osf.io/ygxz8/?view_only=b1d69755c10b4418a7a117a6ca12faae.

Alexandra M. Ryken served as lead for data curation, formal analysis, and visualization and served in a supporting role for project administration, supervision, and writing–review and editing. Emma Tupper served as lead for stimuli creation and served in a supporting role for data curation, writing–original draft, and writing–review and editing. Drew Weatherhead served as lead for funding acquisition, project administration, resources, supervision, and writing–review and editing and served in a supporting role for data curation. Alexandra M. Ryken and Drew Weatherhead contributed equally to conceptualization, methodology, writing–original draft, and investigation.

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

a clear picture of the linguistic diversity experienced by Americans and Canadians. Thus, to become proficient conversation partners, language learners must learn to contend with the variation in their environment and be able to anticipate when it might occur again in future interactions. But how can one predict when someone may speak with an unfamiliar variant of their native language? Work from sociolinguistics suggests that social characteristics of speakers, such as their race, nationality, sex/gender, sexual orientation, and age, can be strongly associated with certain types of linguistic variation (e.g., Campbell-Kibler, 2011; Giles & Billings, 2004; Niedzielski, 1999; Strand & Johnson, 1996). Thus, as listeners, it is beneficial to attend to the social context that co-occurs with language.

Sociolinguistic Learning

Throughout development, we see that listeners attend to the social properties of speakers and attribute these properties to different linguistic patterns. Infants as young as 6-month-olds are sensitive to speaker race as a possible determiner of linguistic information (Uttley et al., 2013; see also May et al., 2019). By 16 months, infants' interpretations of words change depending on the familiarity of the race of the speaker (Weatherhead & White, 2018), and at 22-month social cues influence whether toddlers think words should be generalized to new speakers (Weatherhead & Werker, 2022; Weatherhead & White, 2021). With age and experience, these sociolinguistic associations become stronger and more specific. Adult listeners process the same speech token differently depending on several different social factors such as race, sex/gender, age, nationality, and even stuffed animals associated with social groups (i.e., Kangaroos and Kiwi Birds; Drager, 2011; Hay & Drager, 2010; Niedzielski, 1999; Strand & Johnson, 1996).

Given the increasing amount of linguistic and racial diversity in North America, it is particularly important to understand how sociolinguistic associations are formed and whether certain types of social information are prioritized when we encounter linguistic variation. Of particular interest to the current study is how listeners learn about novel variation at the word level, and how easily they associate different speaker characteristics to linguistic variation. The same words can sound different depending on where the speaker originates from (e.g., Canadian vs. Australian English), or be constructed differently depending on the gender of the referent (e.g., in Spanish, "nina" vs. "nino"). These changes denote phonological and morphological variation, both of which are explored in the current study. In the present study, we ask whether linguistic variation is more strongly associated with one type of speaker characteristic over another.

While there are several different theories that have been proposed to explain sociolinguistic phenomena, these theories in general agree that linguistic information is tied to the social information co-occurring with it. For example, Exemplar Theory claims that properties of speech are stored alongside social information about the speaker and context into mental representations called "exemplars" (e.g., Docherty & Foulkes, 2014; Drager & Kirtley, 2016). During speech perception, exemplars are activated based on their similarity to the input, both in terms of linguistic patterning and dimensions of speaker identity, and the appropriate linguistic representation is selected based on the activated exemplar distribution. Alternatively, in Bayesian approaches (e.g., Clayards et al., 2008; Feldman et al., 2009; Kleinschmidt & Jaeger, 2015; Norris & McQueen, 2008) a linguistic category is represented by a probabilistic distribution, rather than a distribution of stored tokens. Regardless of the framework under which you apply findings in sociolinguistics, one central theme is that a listener's own experience with linguistic and social variation is critical in shaping their linguistic expectations. As such, some information may be encoded with more "weight" because listeners pay closer attention to certain input based on the environment in which they learned language (Docherty & Foulkes, 2014; Drager & Kirtley, 2016).

In the current study, we are interested in whether listeners' linguistic expectations vary based on speaker identity, and whether these expectations are specific to a certain type of linguistic variant (i.e., phonological vs. morphological). Additionally, we consider the social identity of the listener and how it influences perception. Within sociolinguistics and psychology, identity is construed as a multifaceted and complex concept (Ashmore et al., 2004; Bucholtz & Hall, 2005; Noels, 2014). Work from linguistic anthropology demonstrates that speaker's own productions are impacted by the self-prescribed identities (Eckert, 1989; Mendoza-Denton, 2004). It should be noted that in the current study, our conceptualization of identity is far narrower, focusing on the racial background of the listeners. Individuals with different experiences have a different sociolinguistic model to draw from, meaning they may respond differently to social or linguistic primes (Hay et al., 2006). For example, listeners in an area with little linguistic diversity (Gainesville) interpreted the same speech stimuli as being more accented when paired with a South Asian face than a White face; however, listeners from a more multilingual city showed no difference between conditions (Montreal; Kutlu et al., 2022). Though it should be noted that the present study specifically tested nonnative accents which encompass more than phonological variation. Additionally, a study by Drager (2011) examining adult vowel perception found that sensitivity to sociolinguistic variables differed depending on the listener's age and gender, suggesting that the social background of the listener can influence their attention to sociolinguistic markers.

Of importance, the associations observed in sociolinguistic experiments discussed thus far are founded on real-world experience. For example, female speakers tend to have a higher acoustic boundary between /s/ and /ʃ/ in their productions (/s/ as in "snake," /ʃ/ as in shoe). As a result, people are more likely to perceive a sibilant on a /s/-/ʃ/ continuum as /ʃ/ when shown a photograph of a woman than a photograph of a man (Strand & Johnson, 1996). These types of effects are found for both stereotypical and nonstereotypical faces (e.g., a feminine female face vs. a masculine female face), with the more stereotypical having an even larger effect (Johnson et al., 1999). Thus, simply being told the talker is a woman influenced participants' perceptual identification of the vowels (Johnson et al., 1999). The results of this study highlight the strength of the sociolinguistic associations people hold. Similarly, associations can be observed between race and dialects/ethnolects of English. For example, in the United States, PIN/PEN mergers occur more among African American speakers than White speakers, even in geographic areas where PIN/PEN mergers are more common (Austen, 2020; Labov, 2006). Together, the work from the past two paragraphs highlights the importance of real-world experience on forming sociolinguistic associations, as well as how these associations may influence our future listening experiences.

Current Study

In the current study, we are interested in understanding how new sociolinguistic associations are formed, and how prior social and linguistic experience may play a role. From previous work, we know that speaker characteristics such as race or sex/gender influence how listeners interpret speech. However, it is an open question what linguistic expectations listeners have when encountering a novel association between linguistic information and social information, and whether these expectations differ depending on the type of linguistic variation, and/or the listeners' own social identity. One way to investigate sociolinguistic learning is by using a novel, or artificial language. Thus, allowing for the careful control of linguistic information, as well as how frequently a speaker characteristic co-occurs with properties of the language. Such artificial rules allow for the understanding of how sociolinguistic associations develop, and what types of social cues may be prioritized in a new learning context. Indeed, prior studies have demonstrated that indexical context learning is possible in an artificial laboratory setting (e.g., Rácz et al., 2017; Samara et al., 2017).

A recent paper demonstrated that certain types of associations may be learned faster than others and generalized more easily to new linguistic and nonlinguistic contexts (Rácz et al., 2020). Specifically, associations with speaker sex/gender were learned more quickly for a morphological variant. We base the current study on the general premise of Rácz et al. (2020) but extend it in many ways. First, we used spoken language rather than written language to be able to test phonological variation as well as morphological variation. Second, unlike previous artificial language studies, we eliminated a training phase, instead moving right into the test phase. Thereby allowing us to see how ready participants were to learn the rule (i.e., their initial expectations) as well as the rate at which they learned the rule (i.e., how quickly they learned to map the linguistic variation onto the speaker characteristic). Finally, we tested both phonological and morphological variations as well as different speaker characteristics and varied the background of our participants to include those from a majority and minority racial background.

In Experiment 1, we assessed whether speaker sex/gender would be associated and learned more quickly with phonological variation than a nonsocial property (hat-wearing). In Experiment 2, we pit speaker race against sex/gender. Importantly, we tested both majority and minority race populations. In Experiment 3, we used a morphological variant to determine if the type of linguistic variation played a role in participants' behavior. On each trial, participants would see a new object and heard two potential variations of the item. For example, "Someone says [variant 1], someone says [variant 2]." Participants were then asked who said one of the variations (e.g., "Who says [variant 1]?"). They then had the option between picking one of the two clipart characters on the screen. In each experiment, two competing cues could be used to make their decision. Thus, participants needed to learn to use one cue to make their decisions while ignoring the other salient cue. At the end of each experiment, we also asked participants to explicitly state what about the speakers they used to make their decisions.

Experiment 1

The goal of Experiment 1 was to determine whether phonological language patterns are easier to learn when the speaker characteristic predicting the pattern is a characteristic with high social relevance, sex/gender, as opposed to a characteristic with less social relevance, hat-wearing.

Method

Transparency and Openness

For all three experiments, we report how we determined our sample size, all data exclusions, all manipulation, and all measures. Data, analysis code, and example materials are available on the Open Science Framework: https://osf.io/ygxz8/?view_only=b1d69755c10b4418a7a117a6ca12faae (Ryken et al., 2024). We used SPSS Version 28.0.1.1 (IBM Corp, 2021) for descriptive statistics, generalized estimating equations (GEE), *t* tests, and chi-square tests. We used R Version 3.6.1 (R Core Team, 2019) for data visualizations (ggplot2 package: Wickham, 2016; gghalves package: Tiedemann, 2020) and generalized additive models (mgcv package: Wood, 2017). Ethical approval was obtained from the authors' institutional Research Ethics Board (REB) prior to data collection (REB: 2022-5984). Design and analyses were not preregistered.

Participants

Ninety-nine participants completed Experiment 1 through the online crowdsourcing platform Prolific. There were 50 participants in the sex condition and 49 participants in the hat condition. We determined our sample size based on a power analysis which used effect sizes from similar previous research. For this analysis, we used $d = 0.609$ (as predicted by Experiment 2, Rácz et al., 2017), power ($1-\beta$ err prob) = .80, and α err prob = .05, finding a suggested sample size of 88. With a more conservative moderate effect, $d = 0.5$, a sample size of 102 was suggested and so we aimed to recruit as close to this number as possible. No significant differences were found between the conditions on reported gender, $\chi^2(3) = 4.79$, $p = .188$; living environment, $\chi^2(2) = 1.16$, $p = .561$; household income, $\chi^2(6) = 3.38$, $p = .761$; highest level of education, $\chi^2(5) = 4.69$, $p = .455$; racial origins, $\chi^2(4) = 5.35$, $p = .254$; or whether participants spoke another language in addition to English, $\chi^2(1) = 0.04$, $p = .836$. Data were collected in March 2022.

Of the 99 participants in Experiment 1, 55 identified as White, 20 as East Asian, seven as South Asian, six as Black or African American, and 11 as "other." Of these participants, eight identified as having Hispanic or Latinx heritage. In terms of gender, 46 identified as men, 49 as women, and four as nonbinary or other. Fifty-nine participants speak another language in addition to English.

Materials

A set of 64 novel words with two variations each were created. Words followed the pattern consonant, vowel, consonant, and vowel. Each word had two phonological variations in pronunciation, one with /ɛ/ as the first vowel and the other with /i/ as the first vowel (/ɛ/ as in dress, /i/ as in fleece). The other three phonemes remained the same. For example, [mɛlu] and [milu]. All of the words were phototactically possible in English but did not correspond to or sound like real English words. Both variations of each novel word were presented as part of a phrase, spoken by an artificial voice generator (ipa-reader.xyz; Kendra voice setting). For example, "Someone says mɛlu. Someone says milu. Who says mɛlu?" A set of 64 novel object images was created. Images depicted one of 10 unique objects which were then rotated and/or color-changed to create additional images. A set of 128 clipart speaker images were created. The use of clipart characters is consistent with previous studies using artificial language learning studies (e.g., Rácz et al., 2020). Speakers varied on two characteristics: sex/gender and whether the speaker was wearing a hat. Of the speakers created, 32 were female and wearing a hat, 32 were female and not wearing a hat, 32 were male and wearing a hat, and 32 were male and not wearing a hat. The haircut of the speaker was varied to communicate sex/gender. Female-presenting speakers had long hair or ponytails and male-presenting speakers had short hair. There were four possible haircuts for each sex/gender. To create 128 unique speaker images, the haircut, hair color, eye color, and skin color were varied. All speakers appeared to be White. Example images of novel objects and speakers can be found in Figure 1. The novel words, novel objects, and speakers were combined into 64 trials with one word, one object, and two speakers in each. The novel words were presented orally, and the novel objects and speakers were presented visually.

Note that when discussing our sex/gender manipulation we are choosing to use the term sex/gender to indicate that sociocultural and biological factors heavily influence each other making them difficult to separate (Hyde et al., 2019; Morgenroth & Ryan, 2021; Morgenroth et al., 2021).

Procedure

Participants were randomly assigned to one of two conditions: the sex condition or the hat condition. In the sex condition, the sex/gender of the speaker (whether female or male) determined the correct response. In the hat condition, whether the speaker was wearing a hat (or not) determined the correct response. Each novel word had two variations based on the first vowel, and one variation was assigned to each type of speaker. For example, in the sex condition, female speakers might always use the vowel /ɛ/ while male speakers use /i/. In the hat condition, speakers wearing hats might always use the vowel /ɛ/ while speakers not wearing hats use /i/. Which speaker type used which vowel was counterbalanced. The correct response co-occurred with the target speaker characteristic 100% of the time. However, the correct response co-occurred with the nontarget speaker characteristic 50% of the time. For example, in the sex condition, the target speaker also wore a hat in 50% of trials. To be successful, participants had to learn to pay attention to the target characteristic and ignore the nontarget characteristic.

Each of the 64 test trials began with the image of a novel object on a pedestal in the middle of the screen, with the image of a closed door on either side. Participants heard the first part of a novel word phrase: "Someone says [variant 1]. Someone says [variant 2]." Next, two speakers appeared on the screen in front of images of open doors. Participants heard the second part of the novel word phrase: "Who says [variant 1/variant 2]?" Participants clicked on a speaker to indicate who they believe says the word. Feedback (correct or incorrect) was given on each trial immediately following the participant's choice. An example trial layout can be found in Figure 2. Trials were counterbalanced such that the correct word was presented first 50% of the time and the correct speaker was on the left 50% of the time. Each novel word, unfamiliar object, and speaker was seen once by each participant. The order of trials was randomized. The experiment was presented on Labvanced, a platform for building and running online studies.

Following the experiment, participants were asked an open-ended question to determine whether they were consciously aware of the rule, "how did you decide who said each word?" Participants also filled out a short demographic questionnaire containing both multiple-choice and free-response questions. Participants were asked to select one option for each question to report their gender (man, woman, nonbinary, or other), type of living environment (large urban area with a population larger than 20,000 people, small urban area with a population between 2,500 and 20,000 people, or rural area with less than 2,500 people), approximate household income (less than 20,000, 20,000–34,900, 35,000–49,900, 50,000–74,900, 75,000–99,900, 100,000–124,900, or over 125,000), highest level of education (some high school, high school diploma, some college or university, college diploma, associates' degree, trade certificate or equivalent, bachelor's degree, master's degree, professional degree, or doctorate), racial origins (Black/African Canadian or American, East Asian, Indigenous/First Nations, Pacific Islander, South Asian, White, Multiracial, or other), and whether they identify as having Hispanic, Latino, or Spanish origin (yes or no). If participants reported their racial origins as Multiracial or other, they were given the option to specify in a free-response box. Using free-response boxes,

Figure 1
Examples of Novel Objects (Top) and Speakers in Experiment 1 (Bottom)

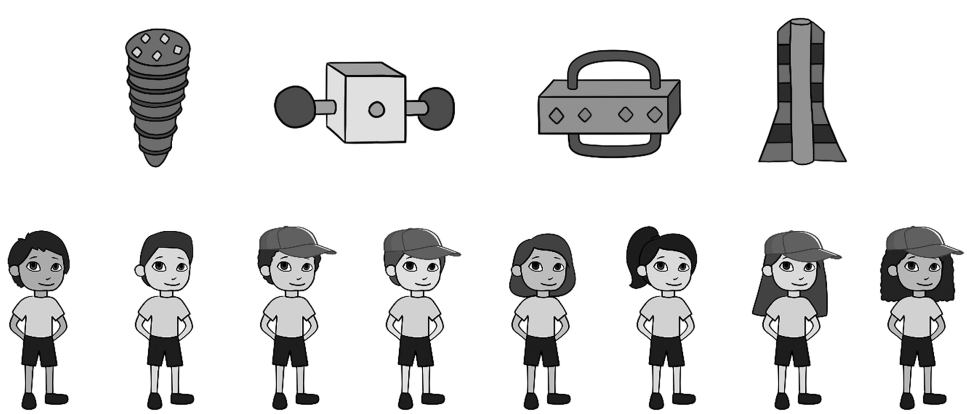
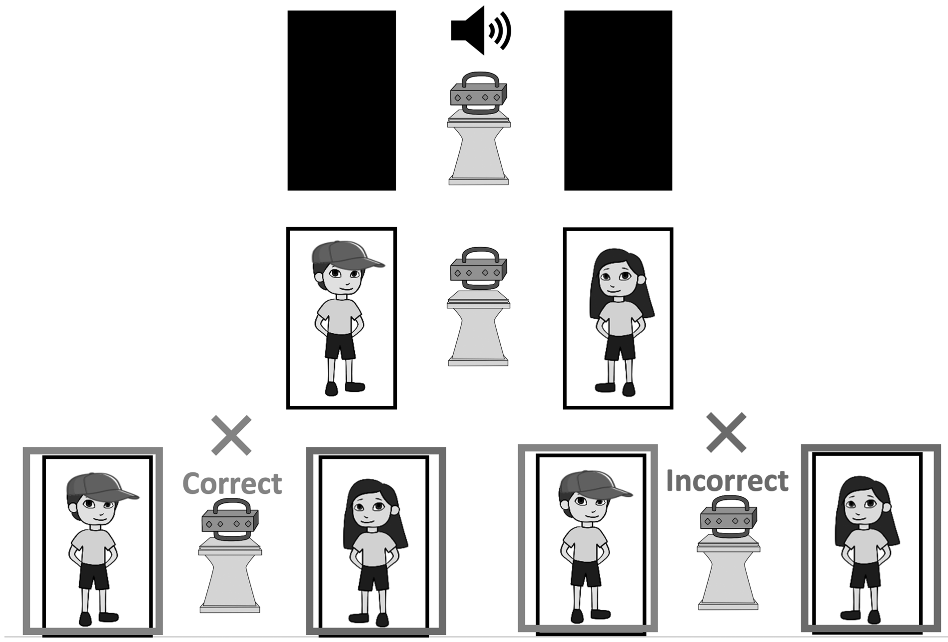


Figure 2

Example Trial Layout Showing the Beginning Screen (Top), the Choice Screen (Middle), and Options for Corrective Feedback (Bottom)



participants were also asked to report their age when they learned English, and any additional languages spoken. Questions and response options can also be found in the Appendix.

Coding

Main Analysis

On each of the 64 trials, participants received a binary score of 0 or 1. If participants chose the correct speaker, they were given a score of 1, otherwise, they received a score of 0.

Exploratory Analysis

For the open-ended question at the end of the study, a coding scheme was developed. For all experiments, responses were coded based on whether the participant correctly or incorrectly identified the cue responsible for the linguistic variation. Participants that mentioned both a correct cue and an incorrect cue were considered incorrect. We further coded participant responses based on which cue (if any) they mentioned as being responsible for the linguistic variation.

Results

Main Analysis

Prior to analyses, the two counterbalanced sex conditions and the two counterbalanced hat conditions were collapsed into single sex and hat conditions. To determine whether phonological language patterns are easier to learn when the speaker characteristic predicting the pattern is sex/gender or whether the speaker is wearing a hat, we ran a GEE binary logistic regression with a logit link and an M-Dependent working correlation matrix. The within-subject factor was trial number (1–64), the between-subjects factor was condition (sex vs. hat), and the dependent variable was accuracy (correct = 1, incorrect = 0). As can be seen in Table 1, an M-Dependent working correlation matrix set to two was the best fit for the data. The analysis revealed main effects of both trial number, Wald $\chi^2(1) = 22.42$, $p < .001$, and condition, Wald $\chi^2(1) = 5.77$, $p = .016$. There was no significant interaction found between trial number and condition, Wald $\chi^2(1) = 1.09$, $p = .298$.

As can be seen in Figure 3, accuracy increases as trial number increases, suggesting that participants are learning their respective rule. Those in the sex condition had higher accuracy scores compared to those in the hat condition. Thus, the sex condition was more readily learned than the hat condition, although the absence of an interaction suggests that participants in both groups learned their respective rules at a similar rate.

In all experiments, we further explored the rate of learning by comparing two generalized additive models, which do not assume a linear

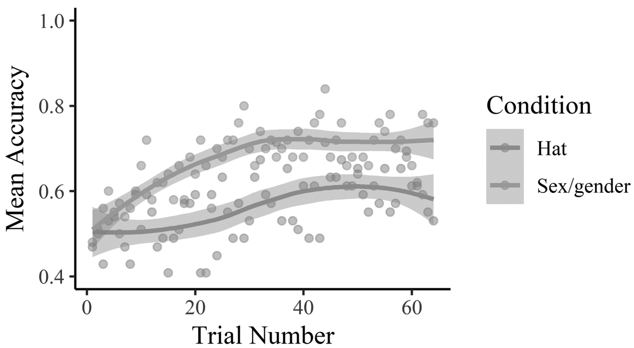
Table 1

Fit of Tested Correlation Structures for Each Experiment

Working correlation matrix	Experiment 1 QICs	Experiment 2 QICs	Experiment 3 QICs
Unstructured	8,376.716	**16,440.956**	N/A[a]
Independent	8,328.758	16,463.355	6,404.278
AR(1)	8,328.678	16,462.881	6,403.829
Exchangeable	8,329.355	16,463.741	6,422.575
M-dependent (set to 2)	**8,328.639**	16,462.692	**6,403.645**

Note. The bolded QIC value for each experiment indicates the best-fitting model. QICs = quasi information criterion; AR (1) = autoregressive model; N/A = not applicable.
[a] Received an error when running the unstructured correlation matrix for Experiment 3 was not appropriate for the data.

Figure 3
Mean Accuracy on Each Trial by Condition in Experiment 1, With Trend Line and 95% Confidence Interval



association. Specifically, we compared a model where the slope of the trend did not vary based on condition to a model where the slope of the trend did vary based on condition. In the first model, accuracy was predicted by condition and smoothed trial number. In the second model, accuracy was predicted by condition and smoothed trial number by condition. The Akaike information criterion (AIC) and Bayesian information criterion (BIC) values for each model can be found in Table 2. Differences of six or more are considered substantive (Raftery, 1995), and a lower AIC or BIC value indicates a better-fitting model. A difference in rate of learning between the conditions would be indicated if the better fitting model is the second model where the slope of the trend did vary based on condition. However, there were no differences in model fit according to the AIC values, and according to the BIC values the model where the slope of the trend did not vary based on condition was a better fit for the data. Consistent with the lack of interaction in the GEE, the results suggest that participants in both groups learned their respective rules at a similar rate. Overall performance was significantly above chance for both the sex condition, $t(49) = 5.88$, $p < .001$, and the hat condition, $t(48) = 2.71$, $p = .009$. The distribution of total trials correct is plotted in Figure 4.

Exploratory Analysis

A chi-square test of independence revealed that condition (sex or hat) does impact how likely participants are to correctly articulate the rule, $\chi^2(1) = 7.56$, $p = .006$. A larger proportion of participants in the sex condition (22/50 or 44.00%) correctly articulated the rule compared to participants in the hat condition (9/49 or 18.37%). A 2×2 analysis of variance (ANOVA) by condition and whether participants correctly articulated the rule revealed that those who did correctly articulate the rule were correct on more trials overall ($M = 53.35$) compared to participants who did not ($M = 33.29$), $F(1, 95) = 109.56$, $p < .001$, although this did not vary by condition.

Experiment 2

Experiment 1 demonstrates that participants more easily develop sociolinguistic associations for social properties than nonsocial properties (consistent with Rácz et al., 2020). The first goal of Experiment 2 was to determine whether phonological language patterns are easier to learn when the speaker characteristic predicting the pattern is sex/gender or race. The second goal of Experiment 2 was to determine whether the racial background of the participant would impact phonological language pattern learning.

Method

Participants

Two hundred three participants completed Experiment 2 through the online crowdsourcing platform Prolific. There were 99 participants in the sex condition and 104 participants in the race condition. To address our second goal, we specifically recruited participants with either a majority (White, $n = 100$) racial background or a minority (Black, $n = 103$) racial background. One additional participant completed the study but reported their race as White and South Asian. Since their experience is likely different than both background groups, they were excluded from analyses. No significant differences were found between the conditions on reported gender, $\chi^2(3) = 3.88$, $p = .274$; living environment, $\chi^2(3) = 4.49$, $p = .213$; household income, $\chi^2(5) = 10.42$, $p = .108$; highest level of education, $\chi^2(8) = 10.07$, $p = .260$; racial origins, $\chi^2(2) = 2.05$, $p = .359$; or whether participants spoke another language in addition to English, $\chi^2(1) = 0.28$, $p = .598$. Data were collected in March and May 2022.

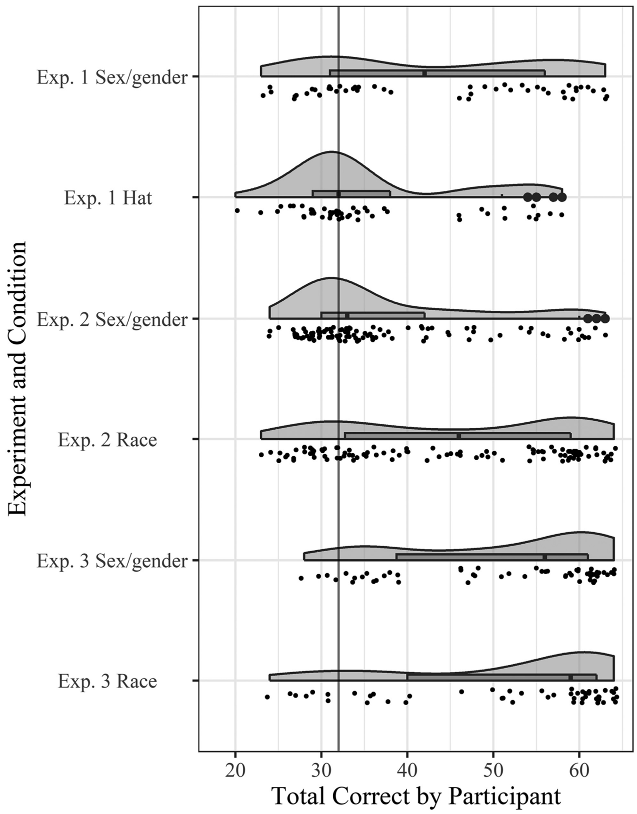
Of the 100 participants who identified as White, 51 identified as a woman, 48 as a man, and one as nonbinary. None identified as having Hispanic or Latinx heritage. Forty-seven reported speaking another

Table 2
Fit Indices for Generalized Additive Models Predicting Accuracy

Model	AIC	ΔAIC	BIC	ΔBIC
Experiment 1				
Condition and trial number	8,688.742	—	**8,724.182**	—
Condition and trial number by condition	8,686.934	1.808	8,733.186	9.004
Experiment 2				
Condition and trial number	**8,688.742**	—	**8,724.182**	—
Condition and trial number by condition	8,772.431	83.689	8,811.308	87.126
Experiment 3				
Condition and trial number	6,290.045	—	**6,362.622**	—
Condition and trial number by condition	6,292.116	2.071	6,400.183	37.561

Note. The bolded values indicate the best-fitting model in cases where differences are larger than 6 and are considered substantive. AIC = Akaike information criterion; BIC = Bayesian information criterion.

Figure 4
Distribution of Total Trials Correct for Each Condition in Each Experiment



Note. The total number of trials was 64. The vertical line indicates chance performance. Exp. = experiment.

language in addition to English. Of the 103 participants who identified as Black, 53 identified as a woman, 49 as a man, and one as nonbinary. One reported having Hispanic or Latinx heritage. Fifty-two reported speaking another language in addition to English.

Materials

The novel word and novel object stimuli were identical to Experiment 1. However, the set of 128 speaker images was slightly different. Speakers varied on two characteristics: sex/gender and race (White or Black). Of the speakers created, 32 were female and White, 32 were female and Black, 32 were male and White, and 32 were male and Black. As in Experiment 1, the haircut of the speaker was varied to communicate sex/gender. Female-presenting speakers had long hair or ponytails and male-presenting speakers had short hair. There were four possible haircuts for each sex/gender. The skin tone and haircut of the speaker were also varied to communicate race. White speakers had lighter skin tones while Black speakers had darker skin tones. White speakers also primarily had straight hair while Black speakers primarily had curly hair. Example images of speakers can be found in Figure 5. To create 128 unique speaker images, the haircut, hair color, eye color, and exact skin color were varied.

Procedure

The procedure was identical to Experiment 1 with the exception of the conditions and the wording of the open-ended question. Participants were randomly assigned to one of two conditions: the sex condition or the race condition. As in Experiment 1, in the sex condition, the sex/gender of the speaker (whether female or male) determined the correct response. In the race condition, the race of the speaker (White or Black) determined the correct response. The wording of the open-ended question was shifted following Experiment 1 in an attempt to focus responses on speaker characteristics. The wording of the question was changed to, “what about the speakers (if anything) helped you decide who said each word?”

Coding

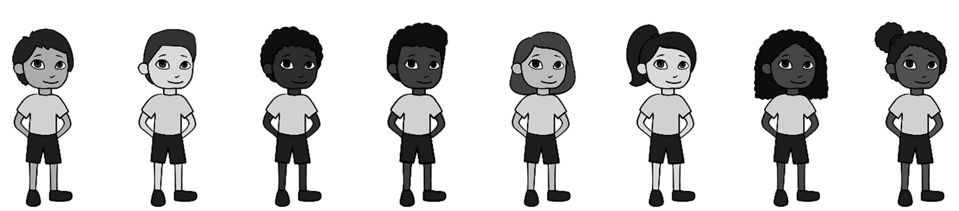
See Experiment 1.

Results

Main Analysis

Prior to analyses, the two counterbalanced sex conditions and the two counterbalanced race conditions were collapsed into single race and sex conditions. To determine whether phonological language patterns are easier to learn when the speaker characteristic predicting the pattern is sex/gender or race and to determine whether the racial background of participants (White or Black) impacts the learning of phonological language patterns, we ran a GEE binary logistic regression with a logit link and an unstructured working correlation matrix. The within-subject factor was trial number (1–64), the between-subject factors were condition (race vs. sex) and background (White or Black), and the dependent

Figure 5
Examples of Speakers in Experiments 2 and 3



variable was accuracy (correct = 1, incorrect = 0). As can be seen in Table 1, an unstructured working correlation matrix was the best fit for the data. The analysis revealed main effects of trial number, Wald $\chi^2(1) = 28.25$, $p < .001$, and condition, Wald $\chi^2(1) = 27.65$, $p < .001$, but no main effect of background, Wald $\chi^2(1) = 3.66$, $p = .056$. There were no significant interactions found between trial number and condition, Wald $\chi^2(1) = 1.12$, $p = .290$, trial number and background, Wald $\chi^2(1) = 0.21$, $p = .651$, or trial number, condition, and background, Wald $\chi^2(1) = 3.30$, $p = .069$.

As can be seen in Figure 6, accuracy increases as trial number increases, suggesting that participants are learning their respective rule. Those in the race condition had higher accuracy scores compared to those in the sex condition. Thus, the race condition was more readily learned than the sex condition, although the absence of interaction suggests that participants in both groups learned their respective rules at a similar rate. As in Experiment 1, we further explored the rate of learning by comparing two generalized additive models (see Experiment 1 for details). According to both the AIC and BIC values, the model where the slope of the trend did not vary based on condition was a better fit for the data. Consistent with the lack of interaction in the GEE, suggesting that participants in both groups learned their respective rules at a similar rate. Overall performance was significantly above chance for both the sex condition, $t(98) = 5.12$, $p < .001$, and the race condition, $t(103) = 10.32$, $p < .001$. The distribution of total trials correct is plotted in Figure 4.

Exploratory Analysis

A chi-square test of independence revealed that condition (sex or race) does impact how likely participants are to correctly articulate the rule, $\chi^2(1) = 11.92$, $p < .001$. A larger proportion of participants in the race condition (53/104 or 50.96%) correctly articulated the rule compared to participants in the sex condition (27/97 or 27.84%). A 2 × 2 ANOVA by condition and whether participants correctly articulated the rule revealed that those who did correctly articulate the rule were correct on more trials ($M = 50.67$) compared to participants who did not ($M = 32.51$), $F(1, 199) = 204.21$, $p < .001$, with performance being higher overall in the race condition. Across both conditions, a chi-square test of independence revealed that a participants' racial background (White or Black) also impacts how likely they are to correctly articulate the rule, $\chi^2(1) = 6.09$, $p = .014$. A larger proportion of White participants (48/100 or 48.00%) correctly articulated the rule compared to Black participants (32/103 or 31.07%).

Experiment 3

The results of Experiment 2 suggest that phonological associations with sex/gender and race are learned at relatively equal rates. However, participants' initial responses and overall performance suggest that they are more likely to expect a phonological variant is related to race. The goal of Experiment 3 was to determine whether a similar phenomenon is observed for morphological variations. For English speakers, neither characteristic is commonly associated with real-world morphological language patterns. However, previous work has suggested that sex/gender may be more readily associated with a morphological variant than other speaker cues (Rácz et al., 2020).

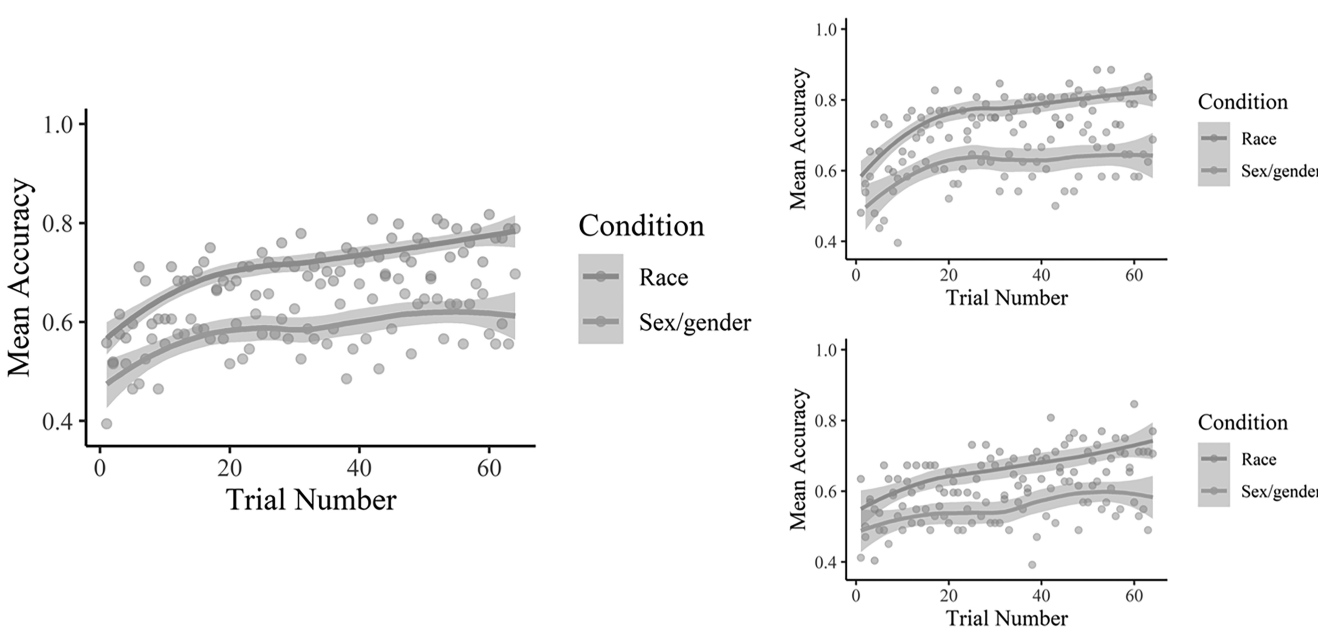
Method

Participants

One hundred two participants completed Experiment 3 through the online crowdsourcing platform Prolific. There were 52

Figure 6

Mean Accuracy on Each Trial by Condition in Experiment 2 for All Participants (Left), White Participants (Top Right), and Black Participants (Bottom Right), With Trend Lines and 95% Confidence Intervals



participants in the sex condition and 50 participants in the race condition. No significant differences were found between the conditions on reported gender, $\chi^2(2) = 0.98$, $p = .612$; living environment, $\chi^2(3) = 6.35$, $p = .096$; household income, $\chi^2(7) = 3.96$, $p = .784$; highest level of education, $\chi^2(7) = 5.29$, $p = .625$; racial origins, $\chi^2(6) = 5.86$, $p = .440$; or whether participants spoke another language in addition to English, $\chi^2(1) = 1.38$, $p = .239$. Data were collected in August 2022.

Of the 102 participants, 60 identified as White, 16 as East Asian, eight as South Asian, seven as Black, two as Indigenous, and seven as other. Zero reported Hispanic or Latinx heritage. Fifty-five participants reported speaking another language in addition to English. In terms of gender, 50 participants identified as a man, 51 as a woman, and one as nonbinary.

Materials

The novel objects were identical to Experiments 1 and 2, and the speaker images were identical to Experiment 2. However, Experiment 3 used a different set of 64 novel words with one root word and two morphological variations each. Root words followed the pattern consonant, vowel, consonant. Each word had two variations in pronunciation, one with /fæl/ added as a suffix and the other with /pos/ added as a suffix. For example, the root word [mɛl] would have two morphological variations: [mɛlfæl] and [mɛlpos]. All the words were phototactically possible in English but did not correspond to or sound similar to real English words. The root and both variations of each novel word were presented as part of a phrase, spoken by an artificial voice generator (ipa-reader.xyz; Kendra voice setting). For example, "Both say [mɛl]. Someone says [mɛlfæl]. Someone says [mɛlpos]. Who says [mɛlpos]?"

Procedure

The procedure was identical to Experiment 2 except for the language pattern. Each novel word had two morphological variations based on the first vowel, and one variation was assigned to each type of speaker. For example, in the Sex condition, female speakers might always use the suffix /fæl/ while male speakers use /pos/ (see Rácz et al., 2020). In the Race condition, White speakers might always use the suffix /fæl/ while Black speakers use /pos/. Which speaker type used which suffix was counterbalanced. As in Experiments 1 and 2, each of the 64 test trials began with the image of a novel object on a pedestal in the middle of the screen, with the image of a closed door on either side. However, to communicate that the addition of the suffix had meaning and was, therefore, a morphological change, participants heard a slightly different first part of the novel word phrase: "Both say [root word]" accompanied by a large image of the novel object. Next, a smaller image of the novel object was shown while participants heard "Someone says [variant 1]. Someone says [variant 2]." As in Experiments 1 and 2, the two speakers then appeared on the screen in front of images of open doors while participants heard the final part of the novel word phrase: "Who says [variant 1/variant 2]?"

Coding

See Experiment 1.

Results

Main Analysis

Prior to analyses, the two counterbalanced Race conditions and the two counterbalanced Sex conditions were collapsed into single Race and Sex conditions. To determine whether morphological language patterns are easier to learn when the speaker characteristic predicting the pattern is sex/gender or race, we ran a GEE binary logistic regression with a logit link and an M-Dependent working correlation matrix. The within-subject factor was trial number (1–64), the between-subject factor was condition (race vs. sex), and the dependent variable was accuracy (correct = 1, incorrect = 0). As can be seen in Table 1, an M-Dependent working correlation matrix set to two was the best fit for the data. The analysis revealed a main effect of trial number, Wald $\chi^2(1) = 46.24$, $p < .001$, but no main effect of condition, Wald $\chi^2(1) = 0.35$, $p = .557$. There was no significant interaction found between trial number and condition, Wald $\chi^2(1) = 0.07$, $p = .798$.

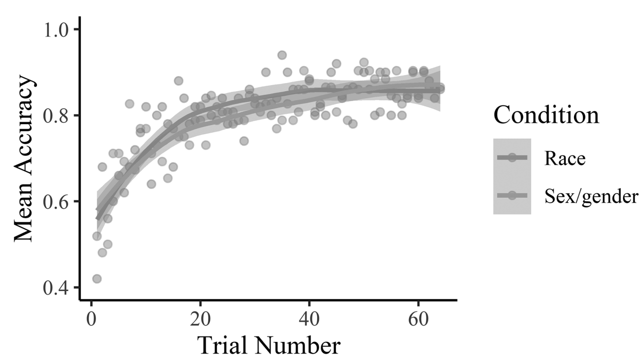
As can be seen in Figure 7, accuracy increases as trial number increases, suggesting that participants are learning their respective rule. However, neither condition is easier than the other. The absence of an interaction also suggests that participants in both groups learned their respective rules at a similar rate. We further explored the rate of learning by comparing two generalized additive models (see Experiment 1 for details). There were no differences in model fit according to the AIC values, and according to the BIC values the model where the slope of the trend did not vary based on condition was a better fit for the data. Consistent with the lack of interaction in the GEE, suggesting that participants in both groups learned their respective rules at a similar rate. Overall performance was significantly above chance for both the sex condition, $t(52) = 11.65$, $p < .001$, and the race condition, $t(49) = 10.69$, $p < .001$. The distribution of total trials correct is plotted in Figure 4.

Exploratory Analysis

A chi-square test of independence revealed that condition (sex or race) did not impact how likely participants are to correctly articulate the rule, $\chi^2(1) = 0.025$, $p = .874$, with similar proportions of participants correctly articulating the rule in the sex condition (32/52 or 61.54%) and race condition (30/50 or 60.00%). A 2×2 ANOVA

Figure 7

Mean Accuracy on Each Trial by Condition in Experiment 3, With Trend Line and 95% Confidence Interval



by condition and whether participants correctly articulated the rule revealed that those who did correctly articulate the rule were correct on more trials ($M = 57.27$) compared to participants who did not ($M = 41.97$), $F(1, 98) = 60.23$, $p < .001$, although this did not vary by condition.

Discussion

Across three experiments, we demonstrated that adult listeners are able to learn the association between a nonsocial feature of a speaker (hat wearing) and phonological variation (Experiment 1), between a social feature (sex/gender and race) and phonological variation (Experiments 1 and 2), and between a social feature (sex/gender and race) and morphological variation (Experiment 3; consistent with Rácz et al., 2017, 2020). Across all studies, participants showed stronger evidence of learning when they were explicitly aware of the association. Within each experiment, participants in both conditions showed evidence of learning the rule. What did vary between conditions within an experiment was participants' "readiness" to learn a specific association. That is, participants brought assumptions in about what social cue is more important to pay attention to initially. First, like Rácz et al. (2017), we found that participants more readily learn associations between speaker characteristics and linguistic variation, and are better able to articulate that association, when the speaker characteristic is socially relevant. Participants in Experiment 1 had higher accuracy scores and were better able to articulate the language patterns for speaker sex/gender (i.e., a cue with high social relevance), compared to hat-wearing (i.e., a cue with low social relevance). Our exclusion of a training phase allows us to more confidently conclude that accuracy rates reflect participants' real-world knowledge of sociolinguistic variation as we were able to track the trajectory of their performance from the very first instance of exposure. In all three experiments, performance initially increases at a higher rate before leveling off somewhat. We included all trials in our analyses including these early trials for the same reasons we excluded a training phrase; the initial learning period is important for examining initial expectations and rate of learning.

Of particular importance, participants also more readily learned an association between speaker race and a phonological pattern when compared to speaker sex/gender and a phonological pattern in Experiment 2, but no differences in readiness were observed for associations between speaker race or speaker sex/gender and a morphological pattern in Experiment 3. Recall, readiness in the context of the current study refers to participants' initial expectations of what cues are relevant to linguistic variation. Participants in Experiment 2 had higher accuracy scores and were better able to articulate the language patterns for speaker race compared to speaker sex/gender, while participants in Experiment 3 showed no differences in accuracy or articulating the rule when it was based on speaker race compared to speaker sex/gender. The difference in accuracy scores in Experiment 2 appears to be largely driven by listeners' initial trials, or a readiness to learn the rule. If listeners have strong existing associations with a social characteristic (i.e., race) and a type of linguistic variation (i.e., phonological), their future interactions with novel speakers will be influenced (Drager, 2010). Put plainly, if they learn that speaker race is important in one context (e.g., a real-world PIN/PEN merger more commonly found in African American speakers), they may then search for subsequent associations between race and accent that may or may not exist. Samara et al. (2017) suggest that the type of linguistic variation is equally important. Thus, we presume that listeners have experienced language input reflecting these real-world sociolinguistic differences. Thus, although the phonological variation was novel to the participants, they had stronger expectations phonological variation was associated with speaker race based on their prior experiences. Based on the pattern of results in Experiment 3, in which accuracy rates between the sex/gender and race condition did not differ significantly, we suggest that participants who learned the morphological variant did not necessarily have stronger expectations that morphological variation is associated with speaker race or sex/gender. Based on these points, it appears that the type of linguistic variation, and social relevance of the conditioning cue (which can be shaped by listener experience and real-world language use) impacts performance. This finding supports a larger body of research citing that people associate language and race in various ways (Babel & Russell, 2015; Kutlu, 2023; McGowan, 2015; Rubin, 1992).

Finally, one discrepancy from Rácz et al. (2020) is that they found that participants prioritized speaker sex/gender above other speaker characteristics when learning a morphological variant. Recall, our findings suggest that participants learned race-based and sex/gender-based rules and at equivalent rates. One thing to note is there are large methodological differences between these two studies. The key differences being that: (a) Rácz et al. (2020) used written language as opposed to spoken language, (b) they had a training phase before testing participants' accuracy, and (c) they had more than two cues in competition in some conditions of the study. Because of these methodological differences, it is difficult to ascertain what may account for the observed discrepancy. However, it does leave open an interesting question as to whether sociolinguistic expectations differ for written and spoken language.

Finally, examining the distribution of total trials correct in Figure 4 reveals additional patterns, especially that learning was somewhat bimodal. The number of participants whose performance is near chance compared to the number of participants with high accuracy varies by condition and experiment, but these patterns tend to reinforce what we see in Figures 3, 6, and 7 and in the statistical analyses. In Experiment 1, many participants in the Hat condition perform near chance. A similar pattern appears in Experiment 2 for the sex/gender condition. That these conditions are so difficult suggest that the distractor cue in each may be difficult to ignore. In Experiment 1, those in the Hat condition may have difficulty ignoring sex/gender. In Experiment 2, those in the Sex/gender condition may have difficulty ignoring race. Although our analyses did not compare performance between different experiments, the difference in distributions is very striking when looking at the sex/gender condition in Experiment 1 (where sex/gender is difficult to ignore in favor of hat wearing) and the sex/gender condition in Experiment 2 (where sex/gender seems easier to ignore in favor of race). The conditions with the highest accuracy are the two conditions in Experiment 3, suggesting that the morphological variation was perhaps more salient than the phonological variation, or that the distractor cues were less distracting.

Limitations and Future Directions

A limitation of our experiments is that they reflect only the sociolinguistic associations held by the listener at the time of participating. From the results of the current study, we are unable to speak to the development of associations. A tenet of exemplar theory states that representations are built on experience and can also be updated and altered as an individual gains more listening experience (Docherty & Foulkes, 2014). It would be worthwhile to examine how race–language associations are learned earlier in life from a developmental perspective. Based on the influence of speaker race on language comprehension (Singh et al., 2020), word processing (Weatherhead & White, 2018), and language-oriented face recognition (Clerc et al., 2022), it is possible that a younger population may show a similar pattern of learning in a study like the present.

Another limitation of these studies stems from the speaker's frequency of particle usage. That is, across experiments and conditions, the speaker characteristic co-occurred with the same particle (either the same vowel shift or the same suffix use) in 100% of trials, which is not reflective of real-word sociolinguistic variation. For example, while in the United States, you may be likely to encounter White speakers with Northern or Southern U.S. accents, or dialects from other countries (e.g., Australian English), or even nonnative accent (e.g., Russian-accented English). Thus, speaker race never occurs with a particular phonological variation 100% of the time. Even within individuals, rates of various phonological or morphological features may differ depending on contextual factors. For example, speakers may shift their accent or dialect while in a formal versus informal context (Giles, 1973; Giles & Ogay, 2007). Previous studies have altered speakers' frequency of particle use to reflect that real-world linguistic variation is not absolute, but probabilistic (Samara et al., 2017). Our methodological choice was intentional, to reduce task demands considering that the absence of a training phase may reduce evidence of learning. However, in future studies, it will be more ecologically valid to vary the frequency with which the speaker characteristic co-occurs with a linguistic property.

Finally, sociolinguistic research must consider speaker race within a broader societal context. Discourse on race and language proposes that the associations people hold about race and language may go beyond experience (e.g., Rosa & Flores, 2017). Our task relied upon listeners detecting a linguistic difference and then attributing it to a speaker property. Thus, if a listener has stronger racial bias, they may enter our task expecting to find differences based on race. However, participants' racial bias was not measured in the current study. Future work may want to further explore the role that racial bias may play in participants' race-based linguistic expectations, as well as the rate of learning new sociolinguistic associations. In general, sociolinguistic learning research should strive to diversify their study samples to capture any differences in perception and learning, especially where the listener identity has yielded conflicting impacts on sociolinguistic perception (Drager, 2011; Huntley et al., 1987).

In summary, across three experiments we observed that listeners form novel associations between speaker characteristics and linguistic variation. However, due to previous experience, associations between certain speaker properties and types of linguistic variation are more salient, and thus these associations are more readily learned in a novel context. We observed that both White and Black participants more readily associated speaker race with a phonological variant. Though, in general, White participants were more conscious of why they were making their decision. Thus, it is critical moving forward to use more representative samples when examining sociolinguistic learning.

Constraints of Generality

These findings suggest that certain types of sociolinguistic associations may be more readily learned than others. However, the specific findings in the current study, for example, that race was privileged over sex/gender when learning a phonological variant may not hold true for listeners outside of the United States/Canada. Additionally, the observed results may not be generalizable for different speaker characteristics (e.g., characters from different racial backgrounds than White or Black, sex/gender that includes other gender identities, nor other social properties such as age). Our results highlight the importance of prior experience on listener's linguistic expectations. Listeners from other countries may experience linguistic variation in an entirely different manner, and thus bring different expectations into an artificial language task.

References

Ashmore, R. D., Deaux, K., & McLaughlin-Volpe, T. (2004). An organizing framework for collective identity: Articulation and significance of multidimensionality. *Psychological Bulletin*, *130*(1), 80–114. https://doi.org/10.1037/0033-2909.130.1.80

Austen, M. (2020). Production and perception of the PIN-PEN merger. *Journal of Linguistic Geography*, *8*(2), 115–126. https://doi.org/10.1017/jlg.2020.9

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

Bucholtz, M., & Hall, K. (2005). Identity and interaction: A sociocultural linguistic approach. *Discourse Studies*, *7*(4–5), 585–614. https://doi.org/10.1177/1461445605054407

Campbell-Kibler, K. (2011). Intersecting variables and perceived sexual orientation in men. *American Speech*, *86*(1), 52–68. https://doi.org/10.1215/00031283-1277510

Clayards, M., Tanenhaus, M. K., Aslin, R. N., & Jacobs, R. A. (2008). Perception of speech reflects optimal use of probabilistic speech cues. *Cognition*, *108*(3), 804–809. https://doi.org/10.1016/j.cognition.2008.04.004

Clerc, O., Fort, M., Schwarzer, G., Krasotkina, A., Vilain, A., Méary, D., Lœvenbruck, H., & Pascalis, O. (2022). Can language modulate perceptual narrowing for faces? Other-race face recognition in infants is modulated by language experience. *International Journal of Behavioral Development*, *46*(2), 83–90. https://doi.org/10.1177/01650254211053054

Docherty, G. J., & Foulkes, P. (2014). An evaluation of usage-based approaches to the modelling of sociophonetic variability. *Lingua*, *142*, 42–56. https://doi.org/10.1016/j.lingua.2013.01.011

Drager, K. (2010). Sociophonetic variation in speech perception. *Language and Linguistics Compass*, *4*(7), 473–480. https://doi.org/10.1111/j.1749-818x.2010.00210.x

Drager, K. (2011). Speaker age and vowel perception. *Language and Speech*, *54*(1), 99–121. https://doi.org/10.1177/0023830910388017

Drager, K., & Kirtley, M. J. (2016). Awareness, salience, and stereotypes in exemplar-based models of speech production and perception. In A. M. Babel (Ed.), *Awareness and control in sociolinguistic research* (pp. 1–24). Cambridge University Press.

Eckert, P. (1989). *Jocks and burnouts: Social categories and identity in the high school*. Teachers College Press.

Feldman, N. H., Griffiths, T. L., & Morgan, J. L. (2009). The influence of categories on perception: Explaining the perceptual magnet effect as optimal statistical inference. *Psychological Review*, *116*(4), 752–782. https://doi.org/10.1037/a0017196

Giles, H. (1973). Accent mobility: A model and some data. *Anthropological Linguistics*, *15*(2), 87–105. https://www.jstor.org/stable/30029508

Giles, H., & Billings, A. C. (2004). Assessing language attitudes: Speaker evaluation studies. In A. Davies & C. Elder (Eds.), *The handbook of applied linguistics* (pp. 187–209). Blackwell.

Giles, H., & Ogay, T. (2007). Communication accommodation theory. In B. B. Whaley & W. Samter (Eds.), *Explaining communication: Contemporary theories and exemplars* (pp. 293–310). Lawrence Erlbaum.

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

Hay, J., Warren, P., & Drager, K. (2006). Factors influencing speech perception in the context of a merger-in-progress. *Journal of Phonetics*, *34*(4), 458–484. https://doi.org/10.1016/j.wocn.2005.10.001

Huntley, R., Hollien, H., & Shipp, T. (1987). Influences of listener characteristics on perceived age estimations. *Journal of Voice*, *1*(1), 49–52. https://doi.org/10.1016/s0892-1997(87)80024-3

Hyde, J. S., Bigler, R. S., Joel, D., Tate, C. C., & van Anders, S. M. (2019). The future of sex and gender in psychology: Five challenges to the gender binary. *American Psychologist*, *74*(2), 171–193. https://doi.org/10.1037/amp0000307

IBM Corp. (2021). *IBM SPSS Statistics for Windows* (Version 28.0) [Computer software]. https://www.ibm.com/spss

Johnson, K., Strand, E. A., & D'Imperio, M. (1999). Auditory–visual integration of talker gender in vowel perception. *Journal of Phonetics*, *27*(4), 359–384. https://doi.org/10.1006/jpho.1999.0100

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

Kutlu, E. (2023). Now you see me, now you mishear me: Raciolinguistic accounts of speech perception in different English varieties. *Journal of Multilingual and Multicultural Development*, *44*(6), 511–525. https://doi.org/10.1080/01434632.2020.1835929

Kutlu, E., Tiv, M., Wulff, S., & Titone, D. (2022). Does race impact speech perception? An account of accented speech in two different multilingual locales. *Cognitive Research: Principles and Implications*, *7*(1), Article 7. https://doi.org/10.1186/s41235-022-00354-0

Labov, W. (2006). *The social stratification of English in New York city*. Cambridge University Press.

Labov, W., Ash, S., & Boberg, C. (2006). *The atlas of North American English: Phonetics, phonology and sound change*. Mouton de Gruyter.

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

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

Mendoza-Denton, N. (2004). Language and identity. In J. K. Chambers, P. Trudgill, & N. Schilling-Estes (Eds.), *The handbook of language variation and change* (pp. 475–499). Wiley.

Morgenroth, T., & Ryan, M. K. (2021). The effects of gender trouble: An integrative theoretical framework of the perpetuation and disruption of the gender/sex binary. *Perspectives on Psychological Science*, *16*(6), 1113–1142. https://doi.org/10.1177/1745691620902442

Morgenroth, T., Sendén, M. G., Lindqvist, A., Renström, E. A., Ryan, M. K., & Morton, T. A. (2021). Defending the sex/gender binary: The role of gender identification and need for closure. *Social Psychological & Personality Science*, *12*(5), 731–740. https://doi.org/10.1177/1948550620937188

Niedzielski, N. (1999). The effect of social information on the perception of sociolinguistic variables. *Journal of Language and Social Psychology*, *18*(1), 62–85. https://doi.org/10.1177/0261927X99018001005

Noels, K. A. (2014). Language variation and ethnic identity: A social psychological perspective. *Language & Communication*, *35*, 88–96. https://doi.org/10.1016/j.langcom.2013.12.001

Norris, D., & McQueen, J. M. (2008). Shortlist B: A Bayesian model of continuous speech recognition. *Psychological Review*, *115*(2), 357–395. https://doi.org/10.1037/0033-295X.115.2.357

Rácz, P., Hay, J. B., & Pierrehumbert, J. B. (2017). Social salience discriminates learnability of contextual cues in an artificial language. *Frontiers in Psychology*, *8*, Article 51. https://doi.org/10.3389/fpsyg.2017.00051

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

Raftery, A. E. (1995). Bayesian model selection in social research. *Sociological Methodology*, *25*, 111–163. https://doi.org/10.2307/271063

R Core Team. (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.

Rosa, J., & Flores, N. (2017). Unsettling race and language: Toward a raciolinguistic perspective. *Language in Society*, *46*(5), 621–647. https://doi.org/10.1017/S0047404517000562

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

Ryken, A., Tupper, E., & Weatherhead, D. (2024, January 16). *Learning linguistic diversity: Listeners have race based linguistic expectations, but only for phonological variation*. https://osf.io/ygxz8/?view_only=b1d69755c10b4418a7a117a6ca12faae

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

Singh, L., Tan, A. R. Y., Lee, K., & Quinn, P. C. (2020). Sensitivity to race in language comprehension in monolingual and bilingual infants. *Journal of Experimental Child Psychology*, *199*, Article 104933. https://doi.org/10.1016/j.jecp.2020.104933

Statistics Canada. (2023). *Multilingualism of Canadian households*. https://www12.statcan.gc.ca/census-recensement/2021/as-sa/98-200-X/2021014/98-200-X2021014-eng.cfm

Strand, E. A., & Johnson, K. (1996, October). Gradient and visual speaker normalization in the perception of fricatives. In D. Gibbon (Ed.), *Natural language processing and speech technology, results of the 3rd KONVENS Conference*, Bielefeld, Germany. de Gruyter 1996. https://dblp.org/db/conf/konvens/konvens1996.html

Tiedemann, F. (2020). *gghalves: Compose half-half plots using your favourite geoms*. https://CRAN.R-project.org/package=gghalves

United States Census Bureau. (2022). *Language use in the United States: 2019*. https://www.census.gov/library/publications/2022/acs/acs-50.html

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

Weatherhead, D., & Werker, J. F. (2022). 20-Month-olds selectively generalize newly learned word meanings based on cues to linguistic community membership. *Developmental Science*, *25*(4), Article e13234. https://doi.org/10.1111/desc.13234

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

Weatherhead, D., & White, K. S. (2021). Toddlers link social and speech variation during word learning. *Developmental Psychology*, *57*(8), 1195–1209. https://doi.org/10.1037/dev0001032

Wickham, H. (2016). *Ggplot2: Elegant graphics for data analysis*. Springer-Verlag.

Wood, S. (2017). *Generalized additive models: An introduction with R* (2nd ed.). Chapman & Hall/CRC.

(*Appendix follows*)

Appendix

Demographic Questionnaire

Before we begin, we would like to collect some demographic information from you. This information is completely anonymous.

There are no right or wrong answers, and none of your answers will prevent you from participating. This information helps us understand and describe our sample, which is important to understanding the generalizability of our study's findings. You may choose not to answer any questions you do not wish to answer.

Gender: How do you identify?

- ○ Man
- ○ Woman
- ○ Nonbinary
- ○ Other

In what type of environment do you live?

- ○ Large urban area (population larger than 20,000 people)
- ○ Small urban area (population between 2,500 and 20,000 people)
- ○ Rural area (less than 2,500 people)

What is your family's household income?

- ○ Less than 20,000
- ○ 20,000–34,900
- ○ 35,000–49,900
- ○ 50,000–74,900
- ○ 75,000–99,900
- ○ 100,000–124,900
- ○ Over 125,000

What is your highest level of education achieved?

- ○ Some high school
- ○ Highschool diploma
- ○ Some college or university, no degrees
- ○ College diploma, associates' degree, trade certificate, or equivalent
- ○ Bachelor's degree (e.g., BA, BSc)
- ○ Master's degree (e.g., MA, MSc)
- ○ Professional degree (e.g., MD, DDS)
- ○ Doctorate (e.g., PhD, EdD)

At what age did you learn English?

Do you speak any additional languages? If so, please list below.

Describe your racial origins

- ○ Black/African Canadian or American
- ○ East Asian (e.g., China, Korea)
- ○ Indigenous/First Nations
- ○ Pacific Islander (e.g., Hawaii, Samoa)
- ○ South Asian (e.g., India, Pakistan)
- ○ White
- ○ Multiracial or other (explain below)

Do you identify as having Hispanic, Latino, or Spanish origin?

- ○ Yes
- ○ No

Received February 15, 2023
Revision received January 16, 2024
Accepted January 22, 2024 ■