

## On the auditory identifiability of Asian American identity in speech: The role of listener background, sociolinguistic awareness, and language ideologies

Charles B. Chang & Kate Fraser\*

**Abstract.** The current study examined the auditory identifiability of Asian American ethnoracial identity, including the role of listener characteristics and ideologies. Results of an identification experiment showed that the overall accuracy of ethnoracial identification on (East and Southeast) Asian talkers was low, but not the lowest among talker groups and not significantly different from accuracy on Black talkers. There were also significant effects of listeners' ethnoracial identity, gender, and linguistic chauvinism (i.e., disfavoring linguistic diversity in the US). In particular, being Asian or a woman was associated with a higher likelihood of accuracy, whereas greater linguistic chauvinism was, to an extent, associated with a lower likelihood of accuracy. Results of a discrimination experiment additionally showed an effect of listeners' awareness of ethnoracially-based language variation: having this awareness was associated with a higher likelihood of accuracy on discrimination trials with one or more Asian talkers. Taken together, these findings converge with previous results showing an effect of the listener's background on ethnoracial perception and further implicate the listener's sociolinguistic awareness and ideologies.

**Keywords.** racial identification; ethnic identification; ethnolects; race; ethnicity; Boston; Asian Americans; sociophonetics; speech perception; language ideologies

### 1. Introduction.

1.1. ASIAN AMERICAN SPEECH AND ETHNORACIAL PERCEPTION. Part of a long line of sociolinguistic research on ethnolects (i.e., constellations of shared norms for linguistic features in speech communities defined along ethnoracial lines; see Labov 1972a,b; Wolfram 1974; Eckert 2008b), a growing number of sociophonetic studies in recent years have examined speech and language variation in Asian American speech communities, including Chinese Americans and Korean Americans in different regions of the US (Wong 2007; Hall-Lew 2009; Hall-Lew & Starr 2010; Wong & Hall-Lew 2014; Cheng et al. in press). Most of these studies have focused on one ethnicity (for a detailed review of this literature, see Cheng et al. 2022), meaning that there is relatively little research that has considered multiple Asian American ethnicities within the same region simultaneously or, indeed, the possibility of an "Asian American" ethnolect. The preponderance of single-ethnicity sociolinguistic studies of Asian Americans reflects the challenging nature of accounting for the socio-demographic, and especially the linguistic, diversity characterizing the large population of English users who are racialized as Asian in the US. However, given Asian Americans' unique position within American society—one of "forever foreigners" but also "honorary whites" (Lo & Reyes 2009)—as well as the social and political connections that exist among Asian Americans of different ethnicities (see, e.g., Maeda 2012; Bauman 2016), the de-

\* We gratefully acknowledge helpful feedback from Cristina Reguera Gómez, Tyler Olds, Samantha Rigor, and audiences at the 97th LSA Annual Meeting, the UC Berkeley Society of Linguistics Undergraduate Students' 6th Annual Undergraduate Linguistics Symposium, and the 24th Annual Undergraduate Research Symposium at Boston University. This work was funded by the Undergraduate Research Opportunities Program and the Peter Paul Career Development Professorship at Boston University. Authors: Charles B. Chang, Boston University ([cc@bu.edu](mailto:cc@bu.edu)) & Kate Fraser, Boston University ([kfraser@bu.edu](mailto:kfraser@bu.edu)).

gree to which shared linguistic norms have developed, or could develop, among Asian Americans *en masse* remains an important question.

One study that examined multiple Asian American ethnicities within the same region was carried out in Boston, Massachusetts (Chang & Dionne 2022; Dionne & Chang 2022). In this exploratory study, the English speech of a small sample of Asian Americans ( $N = 8$ ) comprising four ethnic groups (Chinese, Filipino, Korean, Vietnamese) was analyzed acoustically and auditorily with respect to four phonetic variables. The results of these analyses showed similarities among the different ethnic groups in that all tended to use broadly generalizable features at high rates, stigmatized features at lower rates, and stereotyped features virtually never; however, there were also significant differences among the ethnic groups in terms of specific use rates for every feature that was actually used. Thus, the findings of this study provided evidence of both linguistic unity and diversity among different Asian American ethnic groups, suggesting that, even if there might not be an easily identifiable pan-Asian American ethnolect, the linguistic norms of different Asian American speech communities may overlap to a sufficient degree for constituting a coherent perceptual category for listeners. In other words, it is plausible that listeners of US English have developed an idea—a long-term memory representation—of what “Asian Americans” sound like; this idea may or may not match Asian Americans’ actual speech production patterns, but could be drawn upon by listeners for the purposes of ethnoracial identification.

Is there evidence that listeners have an auditory perceptual category for Asian Americans? That is, can Asian Americans be reliably identified as such just on the basis of their speech? Empirical research on the perception of ethnoracial identity from speech suggests that there is not a simple answer to this question. Although it is clear that listeners can in fact perform ethnoracial identification from speech, their ability to do so appears to vary by ethnolect or ethnoracial group. For some ethnoracial groups such as African Americans, listeners can reliably perceive talkers’ ethnoracial identity from their speech without visual cues, whereas for other ethnoracial groups such as Asian Americans, they do so less consistently (e.g., 86.5% accuracy of race mentions for Black voices vs. 14.3% accuracy of race mentions for Asian voices; Kushins 2014).

Looking across the results of the few studies that have systematically investigated the auditory perceptibility of Asian American identity, we see mixed findings. Some studies found no reliable perception of Asian American identity, at least for monolingual-like talkers (Lee 2014), whereas other studies, using a variety of methodologies, found that some listeners perceive Asian American identity at above-chance levels (Hanna 1997; Newman & Wu 2011; Cheng & Cho 2021). Crucially, prior work on ethnoracial perception often showed an effect of the listener’s background: listeners who shared aspects of the talker’s background were better able to perceive their ethnoracial identity. For example, compared to non-Asian American listeners, Asian American listeners were 20–30% more accurate at identifying Asian American talkers in a forced-choice task (Newman & Wu 2011) and judged Asian American talkers as significantly more likely to be Asian in a rating task (Cheng & Cho 2021); further, Korean American, but not other Asian American, listeners tended to judge Korean American talkers as likely to be Korean specifically, supporting the view that both ethnoracial background and prior experience with a given ethnolectal variety may influence the accuracy of listeners’ ethnoracial perception.

In the present study, we contributed to this line of research by investigating the role of multiple listener characteristics in Asian American identity perception. In particular, we explored two demographic characteristics (ethnoracial identity, gender) and two metalinguistic characteristics (sociolinguistic awareness, linguistic chauvinism), which we describe in further detail below.

1.2. THE PRESENT STUDY. In the current study, we investigated three questions about the auditory perception of Asian American ethnoraical identity. First, how accurately are Asian Americans perceived as such based on speech only (Q1)? Second, is there an effect of listeners' socio-demographic characteristics—in particular, their own ethnoraical background—on their ability to perceive Asian American identity (Q2)? Third, is there an effect of listeners' metalinguistic characteristics, such as their sociolinguistic awareness and/or language attitudes, on their ability to perceive Asian American identity (Q3)?

On the basis of the existing literature, we sought to explore four hypotheses, H1–H4:

H1: Accuracy of perceiving Asian American identity will be low, especially in comparison to perception of other ethnoraical identities.

H2: Accuracy of perceiving Asian American identity will vary according to listeners' own ethnoraical background.

H3: Accuracy of perceiving Asian American identity will be positively correlated with awareness of sociolinguistic variation along ethnoraical lines.

H4: Accuracy of perceiving Asian American identity will be negatively correlated with linguistic chauvinism (i.e., disfavoring linguistic diversity in the local context).

Each of these hypotheses was based on a specific rationale. In the case of H1, the findings of Newman and Wu (2011) led us to predict that, in a task with a more detailed set of response options, the overall accuracy of perceiving Asian talkers would be relatively low—in particular, lower than the mean identification accuracies on Asian talkers found in their study with fewer response options (41–78%). In the case of H2, we predicted—on the basis of the between-group differences, and in particular the perceptual advantage of Asian listeners, observed in previous studies (Newman & Wu 2011; Cheng & Cho 2021)—that perceptual accuracy would show considerable variation according to listeners' ethnoraical background. To be specific, we hypothesized that listeners racialized as Asian, or more generally as non-white, would show higher perceptual accuracy than white listeners, because the stakes for ethnoraical perception are arguably higher for non-white (i.e., ethnoraical minority) listeners than for the white majority in the US (United States Census Bureau 2021). For example, being able to identify someone over the phone as a member of the same ethnoraical group as oneself may be more beneficial for members of ethnoraical minority groups than for members of an ethnoraical majority. In the case of H3, we predicted that an awareness of ethnoraically-based language variation would be correlated with higher perceptual accuracy, by way of promoting greater attunement to the linguistic features of other language varieties, including those that distinguish the speech of Asian Americans from that of other ethnoraical groups. Finally, in the case of H4, we predicted that being less open to linguistic diversity would be correlated with lower perceptual accuracy, because such linguistic chauvinism might inhibit attunement to the linguistic features of other language varieties.

In regard to the scope of “Asian American”, a term which may mean different things to different people, for the purposes of the present study we limited our scope to the population included in the Asian Americans in Boston Corpus (AAiB; Chang & Dionne 2022)—namely, proficient English speakers who were resident in Boston for at least six months and were of any of the four East or Southeast Asian ethnicities most represented in Boston (i.e., Chinese, Filipino, Korean, Vietnamese). The narrow focus on Boston matched our planned recruitment of local Asian

Listener’s Ethnoracial Background	<i>N</i>
Asian (East/Southeast, including mixed)	17
Black/African American	4
Hispanic/Latino (including mixed)	5
White (non-Hispanic)	16

Table 1. Distribution of listeners across self-reported ethnoracial backgrounds

American listeners within Boston, while the narrow focus on talkers of East/Southeast Asian ethnicities was motivated by the dominant sense of the term “Asian American” in the US, referring to Americans with ethnic roots in East Asia (see, e.g., Oxford Advanced Learner’s Dictionary 2023). On other socio-demographic characteristics, however, we took a more inclusive approach, representing a broad swath of life histories and language backgrounds within our group of Asian American talkers (see §2.2) rather than only US-born individuals or first-language English speakers, for example. The motivation for this approach, which was consistent with the diverse make-up of AAiB, was to represent at least some of the diversity of Asian Americans, diversity that likely feeds into any perceptual representation that English listeners may have for what an Asian American sounds like. That is, we had no reason to believe that a perceptual category for Asian Americans would be limited, for example, to individuals born in the US, so we did not use such demographic dimensions as eligibility criteria for the talkers. This approach has consequences for the interpretation of the results of the study, a point to which we return in the discussion (§4).

## 2. Methods.

2.1. PARTICIPANTS. Participants comprised 42 self-identified adult native listeners of US English who were recruited in Boston, Massachusetts. They represented a range of ages (19–75 yr;  $M = 29.4$ ), genders (31 women, 9 men, 2 non-binary), and sociolinguistic backgrounds. The majority were born in the US ( $N = 34$ ) and did not self-identify as an early bilingual ( $N = 37$ ); nevertheless, the majority reported being regularly exposed to another language besides English ( $N = 26$ ). Participants represented a diverse set of ethnoracial backgrounds (see Table 1), although they were not distributed evenly across them. All of the listeners in the Asian group were of East/Southeast Asian ethnicities.

2.2. MATERIALS. The materials for the perceptual tasks consisted of 43 speech samples from 24 talkers (14 female, 10 male;  $M_{age} = 27.4$  yr, range 19–51) in three corpora: the Santa Barbara Corpus of Spoken American English (SBCSAE; Du Bois et al. 2000–2005), the TIMIT Acoustic-Phonetic Continuous Speech Corpus (TIMIT; Garofolo et al. 1993), and the Asian Americans in Boston Corpus (AAiB; Chang & Dionne 2022).<sup>1</sup> A total of 24 samples were used in the identification task and 20 in the discrimination task (see §2.3), with one sample used in both tasks.

Speech samples were selected in order to include a diversity of talkers representing various ethnoracial groups (see Table 2). Most of the 23 talkers who had birthplace data available for them were born in the US ( $N = 18$ ). The eight Asian talkers (five female, three male) were all from AAiB; three were born in the US and five were born outside the US, with ages of arrival to the US ranging between 5 and 22. All had been living in Boston for at least two years by the

<sup>1</sup> Audio recordings in AAiB are not yet publicly available, but the materials that were used to build this corpus are publicly accessible on the Open Science Framework (OSF) at <https://osf.io/hpqq4>.

Talker’s Ethnoracial Background	<i>N</i>
Asian	8
Chinese	2
Filipino	2
Korean	2
Vietnamese	2
Black/African American	2
Hispanic/Latino	3
Native American	1
White (non-Hispanic)	10

Table 2. Distribution of talkers across ethnoracial backgrounds (according to corpus metadata), including a breakdown of the Asian group across four ethnicities

time they were recorded for AAiB. Most (6/8) rated themselves as native-like and/or dominant in English; the others (2/8) rated their English proficiency as “good” (i.e., just below native-like) on a four-point proficiency scale.

With the exception of one talker from SBCSAE, there were two samples for each talker within the set of samples used across the two perceptual tasks. Most samples were of spontaneous speech ( $N = 37$ ); the remainder were scripted speech (read sentences;  $N = 6$ ). The original samples, which had a sampling rate and resolution of 44.1 kHz and 16 bps, respectively, were normalized in Praat (Boersma & Weenink 2022) to an average intensity of 85.0 dB SPL and then converted to MP3 format for faster download. They ranged in duration from 2 to 5 sec.

2.3. PROCEDURE. Participants completed two perceptual tasks via a Qualtrics survey (Qualtrics 2021): an identification task and a discrimination task. The survey began with an informed consent procedure and then proceeded to the identification task, the discrimination task, and a background questionnaire, in that order. Participants completed all tasks online on their own personal device, and were asked to use earbuds or headphones to listen to the audio samples. A copy of the full Qualtrics survey is publicly accessible on OSF at <https://osf.io/hpqq4>. The median time taken by participants to complete the full survey was approximately 33 minutes.

The identification task comprised both forced-choice and free-response questions for each of 24 trials. Each trial played a sample for a unique talker, meaning that there were 24 different talkers heard in this task. The trials were randomized and presented in the same random order for all participants. On each trial, after listening to the audio sample (which could be played multiple times), the participant was asked to identify various characteristics of the talker, including race/ethnicity, birthplace, age, gender, and sexual orientation. The race/ethnicity questions were forced-choice items. The race question included six response options (Black or African American, White or Caucasian, Asian, Native American or Alaska Native, Native Hawaiian or other Pacific Islander, and Mixed race or Mestizo) while the ethnicity question included nine (European American, African American, Hispanic or Latino, West Asian or Middle Eastern, Central Asian, South Asian, East or Southeast Asian, Native American, and Other, where the participant was asked to specify). The birthplace, gender, and sexual orientation questions were also forced-choice, while the age question was free-response. In addition, participants were asked to rate the difficulty of understanding the audio and their confidence in their responses.

The discrimination task was an AX task comprising forced-choice (“same” or “different”) questions for each of 10 trials. The trials were randomized and presented in the same random order for all participants. Each trial played a pair of audio samples of different talkers with no repetition of talkers across trials, meaning that there were 20 different talkers heard in this task. On each trial, after listening to the audio samples (which, as in the identification task, could be played multiple times), the participant was asked to compare the two talkers in terms of gender, race, and ethnicity, as well as to rate their confidence in their responses. The overall distribution of “same” and “different” target answers in this task was 43% vs. 57%, respectively, and in regard to the race and ethnicity discrimination questions specifically, it was 40% vs. 60%.

The background questionnaire following the perceptual tasks consisted of 20 multiple-choice and short-answer questions about the participant’s socio-demographic background, language background, language attitudes, and linguistic awareness. For the purposes of this study, we focused on responses to two questionnaire items probing constructs we refer to as SOCIOLINGUISTIC AWARENESS (i.e., awareness of socio-indexical language variation) and LINGUISTIC CHAUVINISM (i.e., an attitude of valuing one’s own language over others; cf. linguistic openness or “integrativeness”; Gardner 2001):

- (1) SOCIOLINGUISTIC AWARENESS  
“Are there any ethnic or racial groups in the US that you consider to speak a different ‘type’ of English?” (*yes* or *no*)
- (2) LINGUISTIC CHAUVINISM  
“How do you feel about English becoming the official language of Massachusetts/the US?” (1–5 scale; 1 = *strongly against*, 5 = *strongly for*)

2.4. ANALYSIS. Responses for ethnoracial identification in the identification task and for ethnoracial discrimination in the discrimination task were coded as either accurate (1) or inaccurate (0) according to whether or not they matched the target answers from corpus metadata, and then were submitted to statistical analysis using logistic mixed-effects regression modeling in R (R Development Core Team 2022) with `glmer()` in the ‘lmerTest’ package (Kuznetsova et al. 2017). In this paper, we focus on three models built on the binary accuracy variable for ethnoracial identification responses or for ethnoracial discrimination responses. The full dataset is publicly accessible at <https://osf.io/brwfk>. Due to the relatively small size of the dataset, neither random slopes for the fixed effects nor interactions among the fixed effects were included in these models.

Model 1 was meant to explore differences between Asian American talkers and talkers from other backgrounds in terms of ethnoracial identifiability. This model was therefore built on accuracies for all ethnoracial identification responses excluding those on the one Native American talker ( $N = 1932$ ); responses on the Native American talker were not included in the model because, as discussed in §3, accuracy on this talker was at floor (i.e., no responses were accurate), such that there was no variability in responses, causing the model to be nearly unidentifiable. Model 1 included a treatment-coded fixed effect for Talker Race/Ethnicity (Asian, Black, Hispanic/Latino, white; reference level = Asian) as well as random intercepts for Listener and Talker.

Model 2 was meant to test the influence of listener characteristics on the likelihood of identifying Asian American talkers accurately. This model was therefore built on the subset of accuracies for ethnoracial identification responses on Asian American talkers specifically ( $N = 672$ ). Model 2 included four sum-coded fixed effects for characteristics of the listener: Listener

Race/Ethnicity (non-Hispanic white, Asian, Black, Hispanic/Latino), Listener Gender (man, woman, non-binary), Sociolinguistic Awareness (yes/aware, no/unaware), and Linguistic Chauvinism (2, 3, 4, 5/strongly for making English the official language of the US, 1/strongly against). This model also included a random intercept for Talker; however, because the fixed predictors all concerned listener characteristics and the sample size was relatively small, a random intercept for Listener was not included to prevent overfitting. Note that the output of a sum-coded model does not show the last level of each fixed predictor, so Tukey-corrected planned comparisons between levels of the fixed predictors were carried out with the ‘emmeans’ package (Lenth et al. 2021).

Finally, Model 3 examined the influence of listener characteristics on the likelihood of accurately discriminating Asian American talkers from other ethnoracial groups. This model was therefore built on the subset of accuracies for ethnoracial discrimination on trials including at least one ethnically Asian talker ( $N = 504$ ). Model 3 included four sum-coded fixed effects for Listener Race/Ethnicity, Listener Gender, Sociolinguistic Awareness, and Linguistic Chauvinism (all as above in Model 2) as well as a random intercept for Trial. Tukey-corrected planned comparisons between levels of the fixed predictors were carried out with the ‘emmeans’ package.

### 3. Results.

3.1. IDENTIFICATION RESULTS. As an initial step in our analysis of the identification task, we checked accuracies on the multiple-choice gender question (“What do you think this person’s gender is?”), which was expected to be easy (see, e.g., Clopper et al. 2005), to see if listeners could complete the task as instructed. Responses were coded as accurate if the listener’s primary impression of the talker’s gender (as indicated in a choice among “man”, “woman”, “non-binary”, and “other”, where the listener could specify further) matched the talker’s gender as reported in corpus metadata (e.g., responses such as “other: sounded female but might identify as non-binary” were counted as “woman”). Overall accuracy on the gender question was 98.1%, suggesting that listeners were generally able to complete the task successfully.

Turning our attention to accuracy of ethnoracial identification, we found comparatively lower levels of accuracy overall, as well as substantial variation in accuracy across talker backgrounds and listener backgrounds (see Figure 1). In particular, listeners tended to be more accurate on white talkers than on non-white talkers, which may reflect a white bias in ethnoracial identification (i.e., a tendency to choose “white” in the absence of evidence to the contrary); however, Asian talkers did not garner the lowest accuracies among the non-white talkers. Overall accuracy was lowest on the one Native American talker (0%; i.e., no correct identifications), followed by Hispanic/Latino talkers (18%), Asian talkers (30%), Black talkers (34%), and (non-Hispanic) white talkers (77%). Model 1 (see Table 3) confirmed that Asian talkers were significantly more likely to be misidentified than identified correctly as Asian [ $\beta = -0.936, p = 0.002$ ]. Compared to Asian talkers, the likelihood of accurate identification was significantly higher on white talkers [ $\beta = 2.315, p < 0.001$ ] but significantly lower on Hispanic/Latino talkers [ $\beta = -1.108, p = 0.049$ ]. On the other hand, the likelihood of accurate identification was not significantly different between Black and Asian talkers [ $\beta = -0.008, p = 0.990$ ].

Accuracy of ethnoracial identification also varied according to listener characteristics. For example, as shown in Figure 1, Hispanic/Latino listeners tended to be more accurate than Asian, Black, and white listeners on Hispanic/Latino talkers. As for Asian talkers specifically (Q2; see §1.2), the highest accuracies in identification were obtained by Asian listeners and Black listeners. Model 2 (see Table 4) indicated that, compared to the overall average level of accuracy

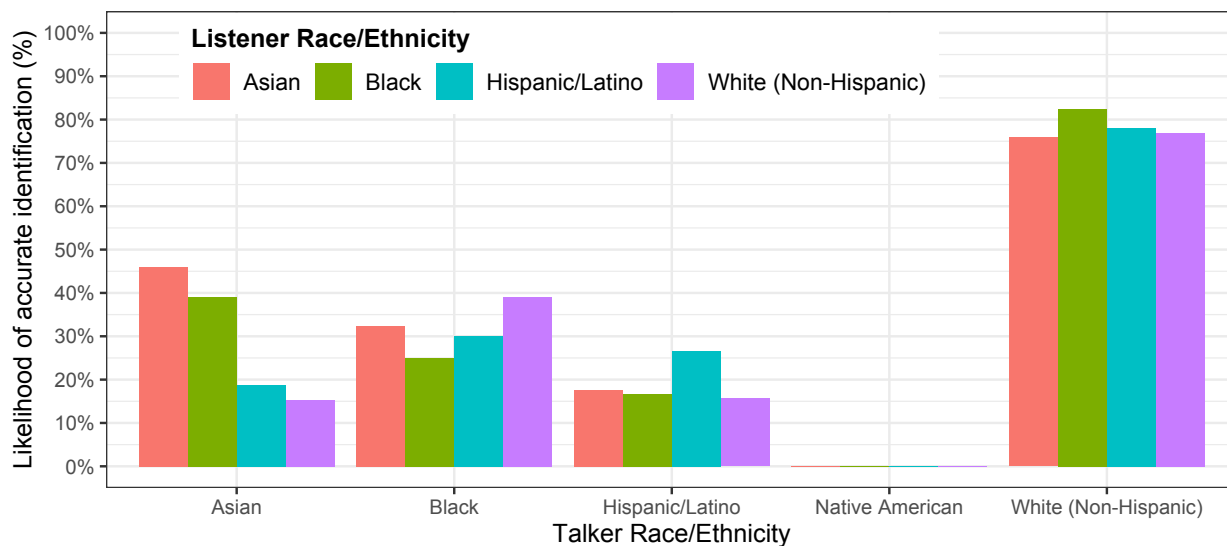


Figure 1. Accuracy of ethnoracial identification, by talker’s and listener’s ethnoracial background

Predictor	$\beta$	$SE$	$z$ -value	$p$ -value	
(Intercept)	-0.936	0.295	-3.171	0.002	**
Talker Race/Ethnicity: Black	-0.008	0.637	-0.012	0.990	
Talker Race/Ethnicity: Hispanic/Latino	-1.108	0.565	-1.962	0.049	*
Talker Race/Ethnicity: white (non-Hispanic)	2.315	0.381	6.078	<0.001	***

Table 3. Fixed-effect coefficients in Model 1 of the log odds of accuracy in ethnoracial identification ( $N = 1932$ , log likelihood =  $-1013.4$ ). Model formula: Accuracy  $\sim$  TalkerEthnicity + (1 | Listener) + (1 | Talker). The intercept represents the predicted log odds for ethnically Asian talkers. Significance codes: \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

on Asian talkers, the likelihood of accuracy was significantly above average for Asian listeners [ $\beta = 0.872$ ,  $p < 0.001$ ] and Black listeners [ $\beta = 0.723$ ,  $p = 0.004$ ], but significantly below average for white listeners [ $\beta = -0.733$ ,  $p < 0.001$ ]. Paired comparisons further indicated that Asian listeners were significantly more likely to be accurate than both white listeners [ $est.$  (Asian – white) = 1.606,  $z$ -ratio = 6.275,  $p < 0.001$ ] and Hispanic/Latino listeners [ $est.$  = 1.735,  $z$ -ratio = 4.916,  $p < 0.001$ ], and that Black listeners were also significantly more likely to be accurate than white listeners [ $est.$  = 1.456,  $z$ -ratio = 4.008,  $p < 0.001$ ] and Hispanic/Latino listeners [ $est.$  = 1.585,  $z$ -ratio = 3.666,  $p = 0.001$ ]; no other between-ethnicity differences were significant. In addition, Model 2 showed a gender effect: women’s likelihood of accuracy was significantly above average [ $\beta = 0.760$ ,  $p < 0.001$ ]. Paired comparisons indicated that women were significantly more likely to be accurate than men [ $est.$  = 0.994,  $z$ -ratio = 3.561,  $p = 0.001$ ]; no other between-gender differences were significant.

There was also an effect of linguistic chauvinism, although not of sociolinguistic awareness. Disregarding the results for a response of 4 (which was represented by only one listener) on the



Predictor	$\beta$	$SE$	$z$ -value	$p$ -value	
(Intercept)	-1.444	0.301	-4.796	<0.001	***
Listener Race/Ethnicity: white (vs. average)	-0.733	0.188	-3.897	<0.001	***
Listener Race/Ethnicity: Asian (vs. average)	0.872	0.174	5.025	<0.001	***
Listener Race/Ethnicity: Black (vs. average)	0.723	0.250	2.894	0.004	**
Listener Gender: man (vs. average)	-0.234	0.253	-0.925	0.355	
Listener Gender: woman (vs. average)	0.760	0.218	3.484	<0.001	***
Sociolinguistic Awareness: yes (vs. average)	-0.015	0.101	-0.153	0.879	
Linguistic Chauvinism: 2 (vs. average)	-0.056	0.227	-0.248	0.804	
Linguistic Chauvinism: 3 (vs. average)	-0.480	0.212	-2.264	0.024	*
Linguistic Chauvinism: 4 (vs. average)	1.061	0.483	2.198	0.028	*
Linguistic Chauvinism: 5 (vs. average)	-0.771	0.428	-1.802	0.072	

Table 4. Fixed-effect coefficients in Model 2 of the log odds of accuracy in ethnoracial identification of Asian American talkers ( $N = 672$ , log likelihood =  $-353.1$ ). Model formula: Accuracy  $\sim$  ListenerEthnicity + ListenerGender + SociolinguisticAwareness + LinguisticChauvinism + (1 | Talker). The intercept represents the average log odds, over all predictors. Significance codes: \*  $p < 0.05$ , \*\*  $p < 0.01$ , \*\*\*  $p < 0.001$

linguistic chauvinism questionnaire item (2), we found evidence in Model 2 that the likelihood of accuracy for listeners who gave a response of 3 (i.e., were unsure of or indifferent to making English the official language of the US) was significantly below average [ $\beta = -0.480, p = 0.024$ ]. In addition, paired comparisons indicated that listeners who gave a response of 3 were significantly less likely to be accurate than listeners who gave a response of 1 (i.e., were strongly against making English the official language of the US) [ $est. = -0.725, z$ -ratio =  $-2.837, p = 0.037$ ]. As for sociolinguistic awareness, Model 2 did not provide any evidence that the likelihood of accuracy for listeners who responded affirmatively to the sociolinguistic awareness questionnaire item (1) (i.e., reported being aware of ethnoracially-based language variation) was different from average, and a follow-up paired comparison further showed no significant difference between listeners who responded affirmatively and those who responded negatively [ $est. = -0.031, z$ -ratio =  $-0.153, p = 0.879$ ].

To provide additional context for understanding how the ethnically Asian talkers were perceived in comparison to other talkers, we inspected the patterns of errors in ethnic identification specifically. These error patterns are shown in the confusion matrix in Table 5. Error patterns supported the view that there was a white bias in ethnoracial identification: for every talker group other than (non-Hispanic) white talkers, the most common type of ethnic misidentification was as European American (i.e., white). By contrast, white talkers were most commonly misidentified as African American. As for Asian talkers specifically, they resembled the other non-white talker groups in terms of being most commonly misidentified as European American (57% of errors); however, they were also misidentified as Hispanic/Latino (15% of errors), African American (8% of errors), and Native American (3% of errors). Most of the “Other” responses on Asian talkers that were coded as errors were “Can’t tell”.

3.2. DISCRIMINATION RESULTS. Similar to the identification task, we began our analysis of the discrimination task by first checking accuracies on the gender question (“Do you perceive these

Talker Race/Ethnicity	Ethnic Identification Response					
	EastAs	AfAm	HispLat	NatAm	EuroAm	Other
Asian (247)	—	8	15	3	<b>57</b>	17
Black (55)	0	—	7	0	<b>87</b>	5
Hispanic/Latino (117)	7	27	—	2	<b>54</b>	10
Native American (42)	10	5	21	—	<b>52</b>	12
White (105)	11	<b>52</b>	13	7	—	16

Table 5. Confusion matrix of errors in ethnic identification (vertical = actual talker race/ethnicity; horizontal = responses, abbreviated). Each cell shows the percentage of all errors on the given talker group represented by the given response (rows may not add to 100% due to rounding); the most common error type for each group is bolded. The total number of errors for each group (across all listeners) is shown in parentheses. Full responses are: “East or Southeast Asian”, “African American”, “Hispanic or Latino”, “Native American or Alaska Native”, “European American”, “Other” (includes all other responses)

Listener Race/Ethnicity	Accuracy (%)		
	Overall	Trials without Asian Talkers	Trials with Asian Talkers
Asian	52	40	59
Black	44	47	42
Hispanic/Latino	55	55	55
White (non-Hispanic)	55	55	55

Table 6. Accuracy of ethnoracial discrimination overall, on trials without Asian talkers, and on trials with one or more Asian talkers, by listener race/ethnicity

speakers to be of the same gender?”) to see if listeners could complete the task as instructed. Responses were coded as accurate or inaccurate according to the talker genders reported in corpus metadata. Overall accuracy on the gender question was 98.4%, suggesting that listeners were generally able to complete the discrimination task successfully.

As for accuracy in ethnoracial discrimination, results converged with those of the identification experiment in suggesting that Asian listeners tended to be more sensitive than other listener groups to Asian talkers (see Table 6) although no between-group differences reached significance. Overall mean accuracy in ethnoracial discrimination was not very different from chance performance (i.e., 50% accuracy) in the binary forced-choice task, ranging from 44% for Black listeners to 55% for Hispanic/Latino and (non-Hispanic) white listeners. However, whereas Black, Hispanic/Latino, and white listeners showed similar accuracies between trials without Asian talkers and trials with Asian talkers, Asian listeners showed substantially higher accuracy on the latter (by 19%) and the highest accuracy on these trials of all listener groups. Nevertheless, the results of Model 3 (see Table 7) indicated that the likelihood of accuracy for Asian listeners on trials with one or more Asian talkers was not significantly different from average [ $\beta = 0.227, p = 0.187$ ]; this was also the case for the other listener groups [all  $p$ 's > 0.1]. Paired comparisons additionally showed no significant between-ethnicity differences. There was also no significant effect of gender evident in the coefficients of Model 3 or in paired comparisons between genders [all  $p$ 's > 0.1].

Predictor	$\beta$	$SE$	$z$ -value	$p$ -value
(Intercept)	-0.031	0.269	-0.115	0.908
Listener Race/Ethnicity: white (vs. average)	-0.041	0.171	-0.241	0.810
Listener Race/Ethnicity: Asian (vs. average)	0.227	0.172	1.320	0.187
Listener Race/Ethnicity: Black (vs. average)	-0.339	0.257	-1.321	0.187
Listener Gender: man (vs. average)	0.095	0.219	0.433	0.665
Listener Gender: woman (vs. average)	0.228	0.190	1.200	0.230
Sociolinguistic Awareness: yes (vs. average)	0.197	0.097	2.019	0.044 *
Linguistic Chauvinism: 2 (vs. average)	0.177	0.232	0.762	0.446
Linguistic Chauvinism: 3 (vs. average)	-0.010	0.202	-0.051	0.960
Linguistic Chauvinism: 4 (vs. average)	-0.446	0.525	-0.850	0.395
Linguistic Chauvinism: 5 (vs. average)	0.444	0.311	1.430	0.153

Table 7. Fixed-effect coefficients in Model 3 of the log odds of accuracy in ethnoracial discrimination on trials with one or more Asian talkers ( $N = 504$ , log likelihood =  $-334.2$ ). Model formula: Accuracy  $\sim$  ListenerEthnicity + ListenerGender + SociolinguisticAwareness + LinguisticChauvinism + (1 | Trial). The intercept represents the average log odds, over all predictors. Significance code: \*  $p < 0.05$

In regard to the metalinguistic predictors, we found no evidence for an effect of linguistic chauvinism, but did find evidence of an effect of sociolinguistic awareness. Neither the coefficients of Model 3 nor paired comparisons between the different levels of linguistic chauvinism showed any significant effects. On the other hand, Model 3 indicated that having awareness of ethnoracially-based language variation was associated with a likelihood of accuracy on ethnoracial discrimination trials with Asian talkers that was significantly higher than average [ $\beta = 0.197, p = 0.044$ ]. Further, a follow-up paired comparison confirmed that listeners with this type of sociolinguistic awareness were significantly more likely to be accurate than those without it [ $est. = 0.393, z$ -ratio =  $2.019, p = 0.044$ ].

**4. Discussion.** The findings of this study provided partial support for our hypotheses H1–H4 concerning the auditory perceptibility of Asian American identity from speech. Results of the identification experiment indicated that listeners’ overall ethnoracial identification accuracy on Asian talkers (30%) was low, and much lower than overall accuracy on (non-Hispanic) white talkers (77%), supporting H1. However, in contrast to previous findings, Asian Americans were not the most challenging to identify of the ethnoracial groups examined; in particular, accuracy was significantly higher on Asian talkers than on Hispanic/Latino talkers, and not significantly different between Asian and Black talkers. Consistent with H2, we also found effects of listeners’ ethnoracial background, as well as of listeners’ gender. As in prior research, ethnically Asian listeners showed an advantage in identifying Asian talkers, significantly outperforming both Hispanic/Latino and white listeners; interestingly, Black listeners showed this advantage as well. There was also a perceptual advantage for women, who significantly outperformed men on Asian talkers; this gender effect was not predicted by any of our hypotheses, but falls in line with previous findings suggesting that, compared to men, women tend to have better hearing and auditory processing (McFadden 1998; Krizman et al. 2021) and greater sensitivity to sociolinguistic variation (Wolfram 1969; Labov 1972a). Finally, consistent with H3 and H4, we found effects of lis-

teners' sociolinguistic awareness and linguistic chauvinism: listeners who reported being aware of ethn racially-based language variation were more likely to be accurate in ethn racial discrimination of Asian talkers from other talker groups, while listeners who were relatively indifferent to linguistic diversity in the US were less likely to be accurate in ethn racial identification of Asian talkers than those who favored linguistic diversity.

Before discussing the implications of the findings further, we would like to point out that the effects of talker race/ethnicity observed here are unlikely to be due to differences in audio quality. Although we roughly matched the durations of speech samples across talker groups and pre-processed the samples in the same way, ultimately the samples came from different corpora recorded under different conditions, and we did not know in advance whether they would be equivalent in terms of comprehensibility. Therefore, to check if there were systematic differences between talker groups in comprehensibility, we conducted a post hoc analysis of listeners' ratings of the difficulty of understanding the audio (see §2.3) via a linear mixed-effects model including a fixed effect of Talker Race/Ethnicity (as in Model 2) and random intercepts for Listener and Talker and Tukey-corrected paired comparisons with 'emmeans'. The results of this analysis indicated that the audio samples for Black talkers were rated as significantly more difficult to understand than those for white talkers [*est.* = 0.823, *z*-ratio = 2.746, *p* = 0.044] and Asian talkers [*est.* = 0.911, *z*-ratio = 3.057, *p* = 0.020], but no other between-ethnicity differences were significant. Crucially, this pattern of differences in comprehensibility does not quite align with the pattern observed in ethn racial identification. Although the Black–white difference in comprehensibility could help account for the much lower ethn racial identification accuracy observed on Black talkers compared to white talkers, the Black–Asian difference is not reflected in the identification results, where there was no statistical difference between Asian talkers and Black talkers and the numerical difference actually favored Black talkers. We take this as evidence that the between-ethnicity differences observed in Model 1—in particular, the differences between Asian talkers and talkers from other ethn racial groups—cannot be due to differences in audio quality.

Taken together, the findings of this study have several implications for our understanding of ethn racial identification from speech and the auditory perceptibility of Asian American identity. First, the results of the identification experiment, which suggest that Asian American talkers were not especially difficult to identify compared to other ethn racial minority groups, support the view that there are indeed socio-indexical features in Asian American speech that allow for Asian American identity to be perceived at rates above chance, similar to Black or African American identity. In other words, at least in terms of speech variation, it may not be appropriate to think of Asian Americans as “honorary whites” (cf. Lo & Reyes 2009), as Asian American identity is clearly being marked, and perceived, in Asian American speech (cf. Kushins 2014). An important caveat, however, is that we purposefully included a wide range of Asian American life histories in our group of ethn ically Asian talkers; consequently, results could differ for a more narrowly-defined sample of Asian talkers. Second, the results of the identification experiment converge with previous findings of an effect of listeners' ethn racial background (Newman & Wu 2011; Cheng & Cho 2021), replicating the perceptual advantage for Asian listeners and additionally showing a perceptual advantage for Black listeners (which occurred in spite of the fact that only half of the Black listeners reported current social connections with Asian Americans; cf. Wong & Babel 2017). Third, the results of both experiments argue in favor of adding sociolinguistic awareness and linguistic chauvinism/openness to the list of factors that may influence listeners' sensitivity to the socio-indexical marking of Asian American identity.

That all said, there are a number of limitations of this study that provide reason to be cautious about the results and that highlight clear avenues for future research. First, our sample size of listeners ( $N = 42$ ) was small, considerably smaller than those of previous studies (e.g., more than 100 listeners in Newman & Wu 2011 and Cheng & Cho 2021), and unevenly distributed across ethnoracial groups, meaning that the current results are almost certainly under-powered and effects related to ethnoracial groups represented by fewer listeners (e.g., Black listeners) in particular should be considered tentative. Second, given the size of our dataset, we could not analyze interactions among predictors, even though such interactions are likely to help account for variation in accuracy in the perceptual tasks (e.g., the interaction between talker race/ethnicity and listener race/ethnicity; see Figure 1). Thus, replicating this study with a larger, more balanced listener sample would allow for greater insight into the effects of listener characteristics on ethnoracial perception of Asian Americans. Third, accuracies in ethnoracial discrimination were quite low overall (in fact, not very different from chance), which leads us to be wary of overinterpreting the patterns observed in ethnoracial discrimination of Asian talkers from other talker groups. The explanation for the relatively poor level of ethnoracial discrimination observed in this study is not entirely clear, but we suspect it may be related to variation in how listeners linked different ethnoracial identities to the concepts of “race” and “ethnicity” in the discrimination task. For example, if listeners were to construe the category of “Hispanic/Latino” as a “race” instead of an “ethnicity”, this could lead them to (incorrectly, in the context of the current study) indicate that a non-Hispanic white talker and a Hispanic/Latino white talker were different races. Thus, it would be useful in future work to probe listeners’ understanding of race and ethnicity directly (e.g., in a study debriefing, which was not done in the current study) and to experiment with different methodologies for testing ethnoracial discrimination.

In closing, we would like to comment on perhaps the most important limitation of this study: its categorical approach to race and ethnicity, both for talkers and for listeners. To simplify analyses, we specifically selected talkers that were classified in terms of one race/ethnicity, grouped listeners in a similar manner, and assumed that any effect of ethnoracial background would be consistent across listeners and across perceptual tasks. In reality, however, the social constructs of race and ethnicity are not so clear-cut, and this categorical approach glosses over potentially interesting effects that may occur in the space between the most widely used ethnoracial categories in the US. For example, might listeners who identify as “Blasian” (Black and Asian) pattern differently from both Black listeners and Asian listeners in terms of ethnoracial perception of Asian talkers? What about Black listeners who also identify as Hispanic? Does the perceptual advantage on Asian talkers associated with a listener’s Asian American identity represent merely a potential behavior, which depends on context (see Eckert 2008a on the related concept of “indexical field” in production)? These are the types of questions that cannot be addressed easily using a categorical approach to race and ethnicity, but that are essential to ask to better understand how ethnoracial identity, including Asian American identity, is expressed and perceived by language users across a range of backgrounds. Future research on the perception of ethnoracial identity in speech would benefit from considering race and ethnicity in more nuanced ways, using experimental methodologies (e.g., visual analogue scale assessment; free classification) that are better-suited for capturing gradient and probabilistic classification behavior.

**Author contributions.** Conceptualization: Charles B. Chang. Data curation: Charles B. Chang, Kate Fraser. Formal analysis: Charles B. Chang. Funding acquisition: Charles B. Chang, Kate

Fraser. Investigation: Charles B. Chang, Kate Fraser. Methodology: Charles B. Chang, Kate Fraser. Project administration: Charles B. Chang. Resources: Charles B. Chang. Supervision: Charles B. Chang. Validation: Charles B. Chang. Visualization: Charles B. Chang. Writing – original draft: Charles B. Chang. Writing – review & editing: Charles B. Chang, Kate Fraser.

## References

- Bauman, Carina. 2016. *Speaking of sisterhood: A sociolinguistic study of an Asian American sorority*. New York: NYU dissertation.
- Boersma, Paul & David Weenink. 2022. Praat: Doing phonetics by computer. Computer program. Version 6.2.23. <http://www.praat.org>.
- Chang, Charles B. & Danielle Dionne. 2022. Unity and diversity in Asian American language variation: Data from Chinese, Filipino, Korean, and Vietnamese Americans. *Proceedings of Meetings on Acoustics* 46(1). 060002.
- Cheng, Andrew, Laretta S. P. Cheng, Wilkinson Daniel Wong Gonzales & Pocholo Umbal. 2022. Variation in Asian and Pacific Islander North American English: A structured literature review. Unpublished manuscript. Available from <https://osf.io/ctuhp/>.
- Cheng, Andrew & Steve Cho. 2021. The effect of ethnicity on identification of Korean American speech. *Languages* 6. 186.
- Cheng, Andrew, Lisa Jeon & Dot-Eum Kim. in press. A comparative study of English vowel shifts and vowel space area among Korean Americans in three dialect regions. *Journal of Linguistic Geography*.
- Clopper, Cynthia G., Brianna Conrey & David B. Pisoni. 2005. Effects of talker gender on dialect categorization. *Journal of Language and Social Psychology* 24(2). 182–206. <https://doi.org/10.1177/0261927X05275741>.
- Dionne, Danielle & Charles B. Chang. 2022. Sociophonetic variation among Asian Americans: The role of ethnicity and style. *Journal of the Acoustical Society of America* 151(4). A63. <https://doi.org/10.1121/10.0010663>.
- Du Bois, John W., Wallace L. Chafe, Charles Meyer, Sandra A. Thompson, Robert Englebretson & Nii Martey. 2000–2005. Santa Barbara corpus of spoken American English, parts 1–4. Web Download. Philadelphia: Linguistic Data Consortium.
- Eckert, Penelope. 2008a. Variation and the indexical field. *Journal of Sociolinguistics* 12(4). 453–476. <https://doi.org/10.1111/j.1467-9841.2008.00374.x>.
- Eckert, Penelope. 2008b. Where do ethnolects stop? *International Journal of Bilingualism* 12(1–2). 25–42. <https://doi.org/10.1177/13670069080120010301>.
- Gardner, Robert C. 2001. Integrative motivation and second language acquisition. In Zoltán Dörnyei & Richard Schmidt (eds.), *Motivation and second language acquisition*, 1–19. Honolulu: Second Language Teaching & Curriculum Center, University of Hawai'i.
- Garofolo, John S., Lori F. Lamel, William M. Fisher, Jonathan G. Fiscus, David S. Pallett, Nancy L. Dahlgren & Victor Zue. 1993. TIMIT acoustic-phonetic continuous speech corpus LDC93S1. Web Download. Philadelphia: Linguistic Data Consortium.
- Hall-Lew, Lauren. 2009. *Ethnicity and phonetic variation in a San Francisco neighborhood*. Stanford, CA: Stanford University dissertation.
- Hall-Lew, Lauren & Rebecca L. Starr. 2010. Beyond the 2nd generation: English use among Chinese Americans in the San Francisco Bay Area. *English Today* 103(26). 12–19. <https://doi.org/10.1017/S0266078410000155>.

- Hanna, David B. 1997. Do I sound “Asian” to you?: Linguistic markers of Asian American identity. *University of Pennsylvania Working Papers in Linguistics* 4(2). 141–153.
- Krizman, Jennifer, Elena K. Rotondo, Trent Nicol, Nina Kraus & Kasia M. Bieszczad. 2021. Sex differences in auditory processing vary across estrous cycle. *Scientific Reports* 11. 22898. <https://doi.org/10.1038/s41598-021-02272-5>.
- Kushins, Eric R. 2014. Sounding like your race in the employment process: An experiment on speaker voice, race identification, and stereotyping. *Race and Social Problems* 6. 237–248.
- Kuznetsova, Alexandra, Per Bruun Brockhoff & Rune Haubo Bojesen Christensen. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82(13). 1–26. <https://doi.org/10.18637/jss.v082.i13>.
- Labov, William. 1972a. *Sociolinguistic patterns*. Philadelphia, PA: University of Pennsylvania Press.
- Labov, William. 1972b. Some principles of linguistic methodology. *Language in Society* 1(1). 97–120. <https://doi.org/10.1017/S0047404500006576>.
- Lee, Eun Hwa. 2014. *Speech production and perception of heritage speakers of Korean*. Los Angeles: UCLA MA thesis.
- Lenth, Russell V., Paul Buerkner, Maxime Herve, Jonathon Love, Hannes Riebl & Henrik Singmann. 2021. emmeans: Estimated marginal means, aka least-squares means [R package]. Version 1.7.0. Available online: <https://cran.r-project.org/web/packages/emmeans/>.
- Lo, Adrienne & Angela Reyes. 2009. Introduction: On Yellow English and other perilous terms. In Angela Reyes & Adrienne Lo (eds.), *Beyond Yellow English: Toward a linguistic anthropology of Asian Pacific America*, 3–17. Oxford, UK: Oxford University Press.
- Maeda, Daryl. 2012. *Rethinking the Asian American movement*. New York: Routledge.
- McFadden, Dennis. 1998. Sex differences in the auditory system. *Developmental Neuropsychology* 14(2–3). 261–298. <https://doi.org/10.1080/87565649809540712>.
- Newman, Michael & Angela Wu. 2011. “Do you sound Asian when you speak English?” Racial identification and voice in Chinese and Korean Americans’ English. *American Speech* 86(2). 152–178. <https://doi.org/10.1215/00031283-1336992>.
- Oxford Advanced Learner’s Dictionary. 2023. Asian American (noun). Available from [https://www.oxfordlearnersdictionaries.com/us/definition/english/asian-american\\_1](https://www.oxfordlearnersdictionaries.com/us/definition/english/asian-american_1).
- Qualtrics. 2021. Qualtrics research suite. Provo, Utah, USA. <http://www.qualtrics.com>.
- R Development Core Team. 2022. R: A language and environment for statistical computing. Version 4.2.1. <http://www.r-project.org>.
- United States Census Bureau. 2021. QuickFacts: United States. Available from <https://www.census.gov/quickfacts/fact/table/US/PST045221>.
- Wolfram, Walt. 1974. *Sociolinguistic aspects of assimilation: Puerto Rican English in New York City*. Arlington, VA: Center for Applied Linguistics.
- Wolfram, Walter. 1969. *A sociolinguistic description of Detroit Negro speech*. Washington, DC: Center for Applied Linguistics.
- Wong, Amy. 2007. Two vernacular features in the English of four American-born Chinese in New York City. *University of Pennsylvania Working Papers in Linguistics* 13(2). 217–230.
- Wong, Amy Wing-mei & Lauren Hall-Lew. 2014. Regional variability and ethnic identity: Chinese Americans in New York City and San Francisco. *Language and Communication* 35(1). 27–42. <https://doi.org/10.1016/j.langcom.2013.11.003>.
- Wong, Phoebe & Molly Babel. 2017. Perceptual identification of talker ethnicity in Vancouver English. *Journal of Sociolinguistics* 21(5). 603–628. <https://doi.org/10.1111/josl.12264>.