CAMBRIDGE
UNIVERSITY PRESS

Article

Dialect proficiency and Mandarin rating in dialect identification: The case of Jiangsu province

Dan Jiao[1,*] and Ksenia Gnevsheva[2]

[1]College of International Studies, Yangzhou University, Yangzhou, China and [2]School of Literature, Languages and Linguistics, Australian National University, Canberra, ACT, Australia

Abstract

This paper investigated origin identification in Jiangsu province, China. In total, ten localities were involved, including nine from Jiangsu province and the city of Beijing. Listeners were presented with recordings of forty speakers (four speakers from each locality) speaking local Mandarin and were asked to identify the region of origin of the speaker and score their Mandarin. Results revealed significant effects of speaker dialect and listener dialect in the identification of speaker origin. Firstly, listeners were able to make distinctions between speakers of Jiangsu province dialect and speakers of non-Jiangsu province dialect (Beijing speakers). Secondly, listeners from Jiangsu province were significantly better than non-Jiangsu listeners at correctly identifying the origin of speakers. In addition, we found significant effects of speaker gender, speaker dialect proficiency, listener dialect proficiency, and speaker Mandarin rating on the identification accuracy of speaker origin in individual analyses of dialect areas.

Keywords: Dialect identification; Jiangsu province; speaker origin; Mandarin rating

1. Introduction

The existing literature boasts a growing amount of work on perceptual dialectology and, in particular, identification or categorization of dialects or accents. The predominant majority of these studies have concentrated on English and its varieties, such as American English (e.g., Clopper & Bradlow, 2009; Clopper & Pisoni, 2007; Lance, 1999; Preston, 1989), British English (e.g., Coupland & Bishop, 2007; Leach, Watson, & Gnevsheva, 2016; Montgomery, 2007; Williams, Garrett & Coupland, 1999), Australian and New Zealand English (e.g., Bayard, Weatherall, Gallois & Pittam, 2001; Weatherall, Gallois & Pittam, 1998), and Scottish English (e.g., Kingstone, 2015). Although the geographical and/or attitudinal perceptions of some Asian languages such as Japanese (e.g., Inoue, 1999; Long, 1999a, 1999b) and Korean (e.g., Long & Yim, 2000) have been on the research agenda, much less attention has been given to other languages and the perception of Chinese dialects in particular.

Identification accuracy has been shown to differ widely across the different varieties of English. In one American English study, overall accuracy was as low as 31% for native listeners categorizing regional dialects in a forced-choice task (Clopper & Pisoni, 2004a). The overall identification accuracy of the provenance of the speakers was similarly low (30%) in the performance of young listeners for Welsh English; however, listeners were more accurate at identifying speakers from their own region (45%) than from a different region (24%), and teachers were more accurate (52%) than

adolescents (Williams et al., 1999). The overall accuracy was a bit higher for dialect identification in northern England (37.6%), but the accuracy ranged from 21.4% to 66.6%, depending on the exact locality (Leach et al., 2016). The accuracy became much higher when it came to the identification of country of origin: when identifying speakers of their own accent, New Zealand English listeners did it correctly at least 79% of the time, Australian listeners at least 80%, and American listeners at least 83% (Bayard et al., 2001).

A number of familiarity-related factors have been revealed to impact the identification accuracy of English dialects, such as the participant region of origin and geographic mobility (Clopper & Pisoni, 2004a, 2004b, 2006), amount of experience (Baker, Eddington, & Nay, 2009), geographical proximity and cultural prominence (Leach et al., 2016; Montgomery, 2012). First, where the listeners come from may affect how well they can identify a variety. For example, local listeners perform better in the identification of local varieties than nonlocal listeners (Baker et al., 2009; Clopper & Pisoni, 2004a). Such a local advantage is also shown in listeners distinguishing local varieties more accurately than nonlocal varieties (Clopper & Pisoni, 2004b; Williams et al., 1999). Next, geographic mobility also leads to a greater differentiation of dialects: mobile listeners who had lived in at least three different states in the US were found to be more accurate in categorizing the regional dialect of unfamiliar speakers than nonmobile listeners (Clopper & Pisoni, 2004b).

In addition, Baker et al. (2009) demonstrated that higher accuracy was achieved by listeners with more linguistic experience such that listeners with more experience of residence in Utah were better at separating Utah speakers from non-Utah speakers than those with less experience. In William et al. (1999), the higher accuracy

Author for Correspondence: Dan Jiao, Email: dan.jiao@yzu.edu.cn

CrossMark

of the teachers over the adolescents in identifying the speakers' origin was also attributed to a greater amount of dialect experience.

Finally, high identification accuracy is also predicted by geographical proximity ("closeness" to an area) and cultural prominence (substantial media exposure of a locality and how established it is "in the national consciousness"). For instance, listeners recognised "home" or "near to" dialect areas more easily than other areas in British English (Montgomery, 2012). However, the effect of proximity was affected by cultural prominence such that culturally salient localities like Newcastle, Liverpool, or Manchester were regularly recognized by participants who live nearby or far away from the dialect areas (Montgomery, 2012). Furthermore, the effects of geographical proximity and cultural prominence were shown to be mediated by the presence, absence, or combination of particularly salient linguistic features in the stimuli (Leach et al., 2016). Thus, we know of several variables that affect identification accuracy, but most of this work is based on English-speaking, western societies.

When Chinese dialects were investigated, most studies focused on the description of dialects, that is, their region and population distribution, phonetic and lexical characteristics, and attitudes toward the dialect. Experimental study of Chinese dialects is gradually increasing with a strong emphasis on the perception of tones (see Gao, Hallé, & Draxler, 2019 for Shanghai dialect; Jin, 2010 for Cantonese and Donghai dialect from Guangdong province; Jin, 2011 for Dianjiang dialect from Chongqing municipality; Jin, 2015a for Hangzhou dialect from Zhejiang province; Jin, 2015b for a comparison between Huilai dialect from Guangdong province and Chaling dialect from Hunan province; Jin & Shi, 2010 for Shantou dialect from Guangdong province; Tang & Li, 2018 for Yangzhou dialect from Jiangsu province; and Zhang & Kong, 2014 for Yuzhou dialect from Henan province), the perception of stops in a dialect (see Wang & Chen, 2016 for Shanghai dialect), or mutual intelligibility across or within dialects (see Tang, 2009 for fifteen Chinese dialects; see Inoue, 2018 for dialect varieties in Fujian province). Very few studies have directly addressed identification of Chinese dialects.

One such example is Blum (2004) who conducted a dialect identification study of Chinese dialects and asked the listeners, most of whom were university students in Kunming, to identify Kunming dialect (speakers speaking local dialect) and Mandarin (speakers speaking Standard Mandarin with different accents) from several other dialects. Listeners could successfully identify the older speakers of Kunming dialect at 63% and the younger speakers at 86–94%. In addition, the accuracy rate for listeners to identify Mandarin was between 55–86%. This study suggests that listeners were able to identify the local dialect from Mandarin and other dialects relatively accurately. This is similar to the findings of English-based studies by Clopper & Pisoni (2004b) and Williams et al. (1999) in that listeners are better at identifying their own dialect, but the accuracy rate is much higher in comparison, probably due to methodological differences and a relative higher degree of similarity between the English dialects involved.

Yan's (2015) work was the first systematic attempt to study the perception of Chinese dialects in Enshi prefecture of Hubei province. Six regional varieties (Enshi, Jianshi, Badong, Hefeng, Xuanen, and Laifeng) were involved in the investigation, and three individual tasks were employed in the study. In a dialect classification task, the participants were presented with real speech samples and asked to identify the county of origin of twelve speakers (two speakers from each variety with one rural and one urban). The

average accuracy rate was 56%, ranging from 92% (correctly identified 11/12 speakers) to 8% (correctly identified only 1/12 speakers). Participants had a positive response bias for the Enshi dialect, in that speakers from other dialects were more frequently classified as Enshi dialect speakers than vice versa. Such an image of being the center of "correctness" was also reported in Inoue (2018) where varieties of Fuzhou and Xiamen were perceived as most central or standard over other varieties for each region in Fujian province. Local participants demonstrated an advantage: both local speakers and listeners outperformed nonlocal ones. This local advantage is similar to that in the identification of English dialects discussed previously (Baker et al., 2009; Clopper & Pisoni, 2004a).

Blum (2004) and Yan (2015) are important studies for understanding speaker origin identification in China. However, more studies are needed to paint a fuller picture. First of all, as both studies focused on participant region of origin only, we still do not know what other factors impact dialect identification in China and how it compares to dialect identification in other countries. Second, as the previous studies focused on dialect, the identification accuracy of local Mandarin remains unclear. However, it is important to investigate as speakers are more likely to use local Mandarin over dialect in cross-province communication.

This study aims to fill this gap by investigating how well listeners can identify speaker origin from Mandarin speech samples in a forced-choice task, what variables predict their accuracy, and how this compares with previous studies, especially in the English-speaking countries. To this aim, participants listened to speakers from nine localities from Jiangsu province and the city of Beijing so as to compare provincial Mandarin with more Standard Mandarin. In addition to dialect identification, participants also rated speakers' Mandarin to explore the relationship between origin identification and Mandarin proficiency. The following section details the method of data collection, followed by results and their discussion in light of the previous literature.

2. Method

2.1 Jiangsu province and its dialects

Jiangsu province, with an area of 102.6 thousand km^2 and a population over 80.5 million,[1] is situated between Shandong province and Shanghai on the east coast of China. There are thirteen cities in Jiangsu province. This study covered ten localities, including the cities of Xuzhou (XZ), Suqian (SQ), Nanjing (NJ), Yangzhou (YZ), Nantong (NT), Suzhou (SZ), Wuxi (WX), Changzhou (CZ), and the district of Ganyu (GY) from Jiangsu province, as well as the city of Beijing (BJ) (Map 1).

The dialects in Jiangsu province comprise three major dialect areas: Northern, Jianghuai, and Wu. Northern and Jianghuai dialect areas belong to the Northern Mandarin dialect in China, and Wu dialect area is part of the Wu dialect in the classification of dialects in China.[2] Each dialect area in Jiangsu province contains more than one city: Jianghuai dialect area consists of 42 localities (including cities and counties) in total, Wu dialect area includes 18 localities, and Northern dialect area covers 8 localities. In short, Northern dialect area includes the cities of Xuzhou (and its affiliated counties and county-level cities) and Suqian and the district of Ganyu (belonging to Lianyungang city). Jianghuai dialect area is composed of the cities of Nanjing, Yangzhou, Zhenjiang, Huai'an, Lianyungang, Yancheng, Nantong, Taizhou (and their affiliated counties and county-level cities), and a few counties from Suqian. Wu dialect area consists of the cities of Suzhou, Wuxi, Changzhou (and their affiliated county-level cities), and some
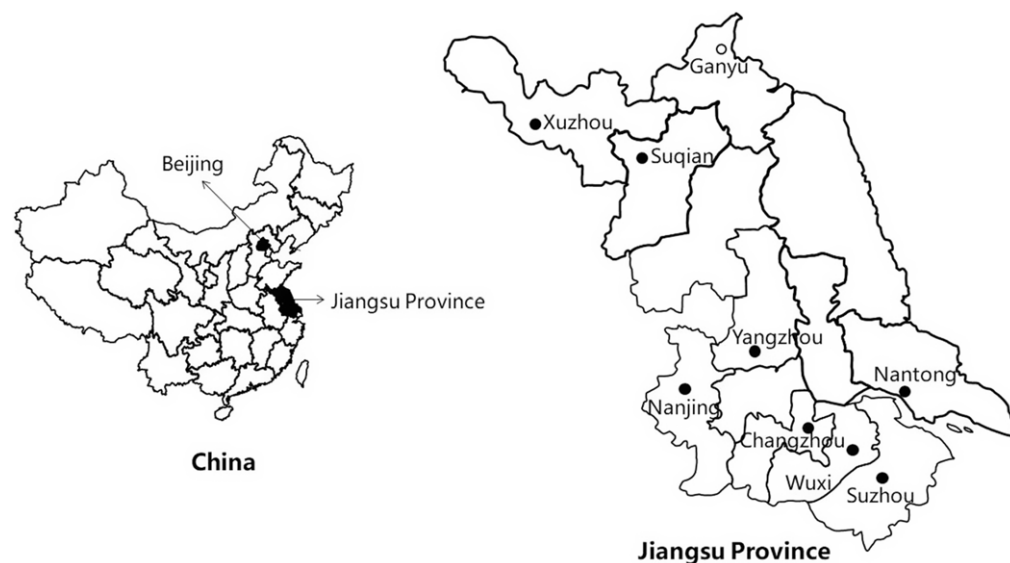
Map 1. The location of Beijing and Jiangsu province in China (left panel); nine targeted localities in Jiangsu province (right panel).

county-level cities or districts from Zhenjiang, Nantong, Nanjing, and Taizhou (Jiangsu Local Chronicles Compilation Committee, 1998).

Choosing the nine localities from Jiangsu province took a consideration of political, social, cultural, and linguistic factors. To be more specific, we intended to include the three dialect areas and choose three representative cities for each in the investigation; however, there are more than three cities in each dialect area. The Northern dialect area only includes three major localities. Although Ganyu is not a city and belongs to the city of Lianyungang in the administrative division of Jiangsu province, it was still included here as it displays a unique phonetic characteristic and holds an important place among Jiangsu dialects (Jiangsu Local Chronicles Compilation Committee, 1998). For the Wu dialect area, Suzhou, Wuxi, and Changzhou were chosen because they are more linguistically salient than other Wu dialect area cities in Jiangsu province, and they also hold good economic and social reputations in Jiangsu. For the Jianghuai dialect area, the successful candidates were Nanjing, Yangzhou, and Nantong. Nanjing is the capital of Jiangsu, and Yangzhou and Nantong are both prototypical cities in the division of dialect subareas in Jianghuai dialect area. These nine localities and the city of Beijing were included in stimuli preparation as the origin of speakers.

The dialects in Jiangsu province differ in terms of phonetics, lexicon, and grammar. As we used a read passage as stimulus (see below), the phonetic differences are of utmost importance. We discuss the most salient phonetic differences between the three dialect areas of Jiangsu province and Standard Mandarin here due to limited space (see Jiangsu Local Chronicles Compilation Committee, 1998 for further information).

For Northern dialect area, compared to Standard Mandarin, /a/ is more backed in Xuzhou dialect, and nasalized in Ganyu dialect (when preceding /ŋ/). /ə/ is fronted in Xuzhou dialect and similar to /ɤ/ in Mandarin. /u/ is dropped in triphthongs /uei/ and /uen/ in Suqian and Ganyu dialects, which are pronounced as [e] and [ɔ̃] in Suqian, and [ei] and [ən] in Ganyu respectively. In terms of consonants, only Ganyu dialect has a distinction between word final /n/ and /ŋ/. Xuzhou and Suqian have a /ŋ/, but /n/ is dropped in both dialects in diphthongs /an/ and /en/, becoming [æ̃] and [ɔ̃] respectively. j /tɕ/, q /tɕʰ/, and x /ɕ/ are absent in Ganyu dialect.

Mandarin /ts, tsʰ, s/ are either pronounced as [ts, tsʰ, s] or [tʃ, tʃʰ, ʃ] in Ganyu dialect.

In Jianghuai dialect area, /u/ is less rounded, and /y/ is more open (similar to /ø/) in Nantong dialect compared to Mandarin. In terms of consonants, both sets of z, c, s and zh, ch, sh are pronounced as z, c, s in Yangzhou (e.g., 'zhen 针, zheng 征' [tsən, tsən]) and Nantong dialects (e.g., 'zhen 针, zheng 征' [tsɛ̃, tsɛ̃]). /n/ and /ŋ/ in Nanjing and Yangzhou dialects are different from Mandarin. When /n/ occurs after vowels in Nanjing and Yangzhou dialects, it is not pronounced. Both /n/ and /ŋ/ are dropped when they occur after vowels in Nantong dialect. Nanjing and Yangzhou do not distinguish /n/ from /l/, and both are pronounced as [l]. The mixing of /n/ and /l/ is more common in the middle and older ages of Nanjing speakers. Some words of j /tɕ/, q /tɕʰ/, x /ɕ/ are pronounced as g /k/, k /kʰ/, h /x/ in Yangzhou dialect (e.g., j in 'jia 家' is pronounced as [k]).

For Wu dialect area, in terms of vowels, there is a distinction between /a/ and /ɑ/ in Suzhou and Wuxi dialects (e.g., [pã] ≠ [pɑ̃]). /ɑ/ is fronted as a monophthong and in diphthongs /iɑ/ and /uɑ/ in Wuxi dialect. Another distinguishing feature of Wuxi dialect is a fronted /o/ as a monophthong and in diphthong /io/. /e/ in /ei/ is backed in Changzhou dialect and is lower and backer in Wuxi dialect. Monophthong /ɤ/ and diphthongs /iɤ/ and /uɤ/ are fronted in Changzhou dialect compared to Mandarin. Apparent fricative sound is found in /i/ in Changzhou dialect. In terms of consonants, Suzhou and Changzhou have only z, c, s, but not zh, ch, sh. /n/ and /ŋ/ are not distinguished in the three localities of Wu dialect area, especially after the vowels /ə/ and /i/ (e.g., 'yin 音' = 'ying 英' [in]). Some words with j /tɕ/, q /tɕʰ/, x /ɕ/ are pronounced with z /ts/, c /tsʰ/, s /s/ in Suzhou dialect (e.g., 'xi 西' [si], 'jian 尖' [tsiɪ]). In addition to the aspirated and unaspirated consonant pairs (e.g., /p, pʰ/) in Wu dialect area, there is a third voiced consonant (/b/) at the same place of articulation, creating three-consonant triplets /p, pʰ, b/, /t, tʰ, d/, /k, kʰ, g/, /tɕ, tɕʰ, dʑ/, and /tʂ, tʂʰ, dʐ/.

Apart from vowels and consonants, rhotic -r /ɚ/ words appear in Xuzhou, Ganyu, and Nanjing dialects and some county-level city dialects of Nantong (e.g., Rugao city), but not in Suzhou, Wuxi, or Changzhou dialects (Cai, 2015). Xuzhou, Suqian, and

Ganyu have four tones as does Mandarin. The remaining six localities in Jiangsu province all have the entering tone, which is a final glottal stop /ʔ/ and forms its independent vowel system. Some of these dialect features make their way into speakers' local Mandarin.

Standard Mandarin is the official spoken language in China. In people's daily communication, they often speak Mandarin that exhibits characteristics of the local dialect, thus it is called "local Mandarin" (Chen, 1990; Chen, 1991; Yao, 1989). Local Mandarin is considered as neither Standard Mandarin nor local dialect, because Standard Mandarin is based on Beijing dialect. Local Mandarin is an interlanguage transiting from local dialect to Mandarin (Chen, 1991; Li, 2010). Local Mandarin is actually the Mandarin that people speak every day and everywhere in China.

2.2 Speakers and speech stimuli

In order to have an even distribution of age and gender for each locality, we recruited one younger female, one younger male, one older female, and one older male from each locality. Therefore, 40 speakers were selected (2 genders * 2 ages * 10 localities). All younger speakers (age range 19–21 years; mean age 20 years) were first- or second-year undergraduates at Yangzhou University. Older speakers (age range 43–50 years; mean age 46 years) were the father or mother of the younger speakers. Both younger and older speakers were locally born and raised. Most speakers reported being able to speak their hometown dialect to a certain proficiency, except two younger speakers. Self-reported dialect proficiency included five levels: Cannot speak at all (n = 2; both younger), Beginner (2; both younger), Intermediate (5; 3 younger), Advanced (3; all younger), and Native (28; 10 younger). Most speakers (35/40) reported their parents to be native speakers of the local dialect. None of the speakers had any speaking disorders, and none participated in the dialect perception experiment online as listeners.

The younger speakers were recorded in a soundproof booth at Phonetics, Hearing and Cognition Lab at Yangzhou University using professional recording facilities with the sampling rate of 44,100 kHz and sampling size of 32 bit. The recording was saved in.wav format. The older participants were asked to record themselves in a quiet space over a recording app on their mobile phone. The sampling rate was set at 44,100 kHz, and the recording format was.wav. All speakers were given 20 yuan for their contribution.

The speakers read a Chinese version of "The North Wind and the Sun" (see Appendix) in Standard Mandarin (Putonghua). There were eight sentences in the passage, and we extracted the first five sentences for each speaker as the stimuli, which were between 18–36 seconds long, depending on the speech rate of the speaker.

2.3 Listeners and online experiment

Sixty-three participants were recruited as listeners for an online perception experiment. Three participants' data were excluded because two were nonnative speakers of Mandarin and one reported recognizing a speaker in the stimuli. Therefore, 60 participants' (57F, 3M; age range 19–25 years; mean age 22 years) data were used for analysis in the study. All participants were studying toward a Bachelor's or Master's degree at Yangzhou University. Most participants (41/60) could speak a dialect of Jiangsu province: Northern (n = 3), Jianghuai (23), Wu (15), and non-Jiangsu (19; dialects that were not from Jiangsu province and were reported as Mandarin). The dialect proficiency for the listeners had the same

five levels as for the speakers, and there was a relatively balanced distribution across different self-reported levels of proficiency: Cannot speak at all (n = 1), Beginner (11), Intermediate (20), Advanced (15), and Native (13). All listeners reported normal hearing and were offered a small gift for their participation.

The audio stimuli were presented on an online questionnaire platform "Wen Juan Wang" (www.wenjuan.com) in Chinese. The participants were asked to listen to the audio clip first and then answer four questions one after another: 1) 您觉得说话者来自哪个城市 (Where do you think the speaker comes from)? A fixed choice of ten localities was offered to the listeners in the same order for all audio clips (see Figure 1a for a screenshot of this task): 北京 Beijing, 徐州 Xuzhou, 宿迁 Suqian, 赣榆 Ganyu, 南京 Nanjing, 扬州 Yangzhou, 南通 Nantong, 苏州 Suzhou, 无锡 Wuxi, and 常州 Changzhou. 2) 请简要说明您选择这个城市的理由 (Briefly explain the reason for your choice of the city). 3) 请给说话者的普通话打分，"1" 代表非常不地道，"7" 代表非常地道 (Rate the speaker's Mandarin on a Likert scale from "1" - very inauthentic to "7" - very authentic) (see Figure 1b for a screenshot of this task). 4) 请简要说明您给出的分数的理由 (Briefly explain the reason for your rating). The second and fourth were open-ended questions, and participants typed their responses in a text box. The order of the speakers in the online experiment was randomized beforehand, but the presentation order of the audio clips was the same for all participants due to the limitations of the online platform. Participants could listen to the audio stimuli as many times as they wished, and there was no time limit for the whole perception task. In total we received 2,400 valid judgment responses (60 listeners * 40 clips) from the online experiment after excluding the data of the three participants.

In this paper, we first calculated the identification accuracy for the first question. Second, we summarized the result of the Mandarin rating for the third question. Third, we fitted several statistical models to investigate the factors that influence identification accuracy. Listeners' comments in the second question for identification and the fourth question for Mandarin rating are not the focus of this paper.

3. Results

3.1 Identification accuracy

The overall accuracy of dialect identification was 16.2% (388 correct identifications out of 2,400 responses). However, there were some accuracy differences across the localities, as shown in Figure 2. Beijing was by far the most correctly identified by the listeners of all localities (44.6%), next were the cities of Suzhou (18.3%) and Yangzhou (17.9%). The identification accuracies of Suqian (15.4%), Ganyu (13.3%), Nanjing (12.9%), Wuxi (12.1%), and Xuzhou (11.3%) were all below the average accuracy rate. Nantong (8.8%) and Changzhou (7.1%) were the least correctly identified by the listeners.

If we collapse these results by dialect area, we can see that, although Beijing was still most correctly identified of all (44.6%), the difference between Beijing and the other dialect areas diminishes: Northern dialect area speakers were correctly identified at 40.4%, Jianghuai dialect area at 36.9%, and Wu dialect area at 35.3% (Figure 3).

3.2 Mandarin rating

Figure 4 visualizes the distribution of Mandarin ratings of all speakers from the ten localities received from the 60 listeners in

(a)

1.您觉得说话者来自哪个城市?

○ 北京

○ 徐州

○ 宿迁

○ 赣榆

○ 南京

○ 扬州

○ 南通

○ 苏州

○ 无锡

○ 常州

(b)

3.请给说话者的普通话打分，"1"代表非常不地道，"7"代表非常地道。 *

○ 1

○ 2

○ 3

○ 4

○ 5

○ 6

○ 7

Figure 1. Mobile phone screenshots of the dialect identification task (a) and Mandarin rating task (b) for listeners in the online experiment.
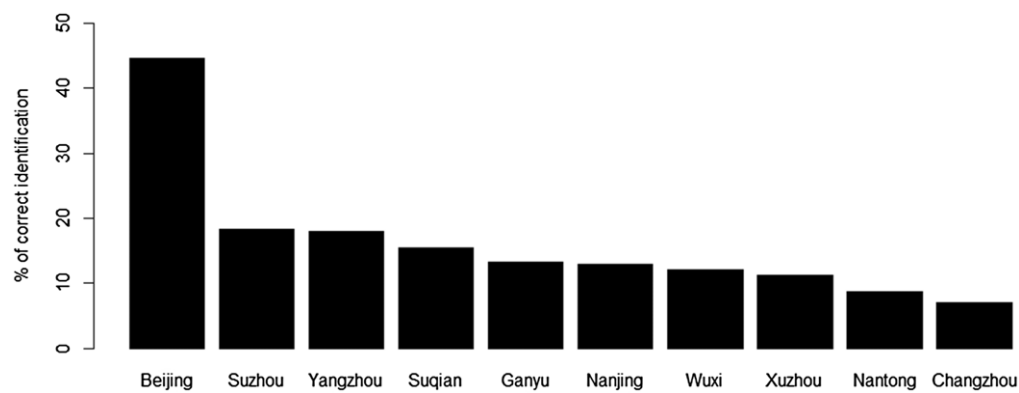

Figure 2. Identification accuracy of speaker origin for each locality.
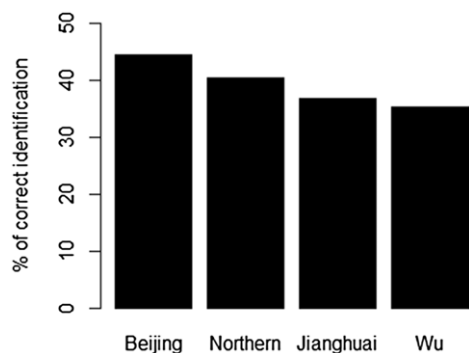

Figure 3. Identification accuracy of speaker origin for each dialect area.

question 3. Most of the mean rating scores fell between 4 and 6, except for one Nanjing speaker and one Wuxi speaker. The mean rating score for Beijing Mandarin speakers was between 5 and 6, for Northern dialect area speakers it ranged from 4 to 6, for Jianghuai and Wu dialect area speakers 3 to 6.

Beijing Mandarin speakers had the least variation in the Mandarin ratings, which suggests that participants had more agreement on the scoring of their Mandarin. For the remaining localities, at least one speaker's mean score was rated lower than 5, and the boxes had longer whiskers compared to the Beijing speakers, suggesting that there was more variation in the rating scores. Furthermore, score 5 seems to be the watershed score separating younger and older speakers' Mandarin. Most younger speakers (white boxes) were scored 5 and over, whereas most older
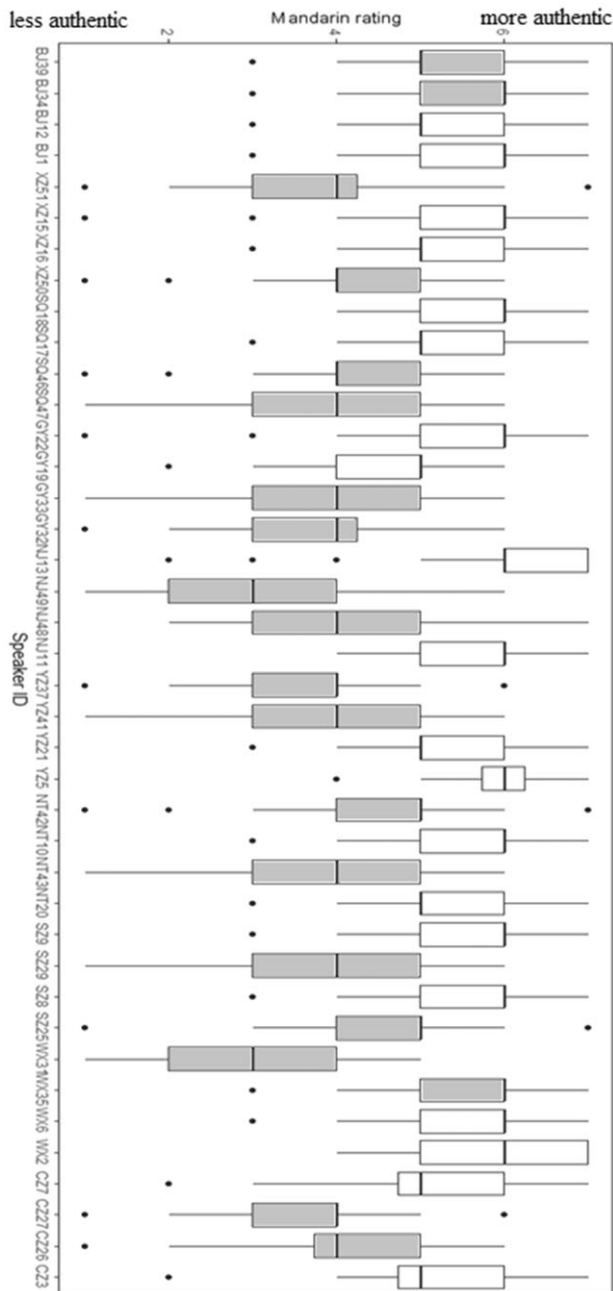
Figure 4. Mandarin rating for the forty speakers.

speakers (grey boxes) were under 5, suggesting higher Mandarin proficiency and perhaps lower dialect proficiency in younger speakers. Speaker age and Mandarin rating will be tested as predictors of identification accuracy below.

3.3 Factors affecting identification accuracy

3.3.1 Larger dialect areas
In order to see what factors significantly affected identification accuracy, we fit a binomial logistic mixed effects model (Baayen, Davidson & Bates, 2008) to the data using the lmerTest package (Kuznetsova, Brockhoff & Christensen, 2017) in R (R Core Team, 2019). The dependent variable was a binary, indicating whether or not the listener correctly identified the speaker's origin.

We tested the predictors of speaker age, speaker gender, speaker dialect (Beijing, Northern, Jianghuai, and Wu), speaker dialect proficiency, speaker Mandarin rating, listener dialect (Northern, Jianghuai, Wu, and non-Jiangsu), and listener dialect proficiency. Speaker dialect proficiency and listener dialect proficiency were both coded as a continuous variable in accordance with the five levels (1 = "cannot speak at all," 2 = "beginner," 3 = "intermediate," 4 = "advanced," and 5 = "native"). Speaker and listener were included as random intercepts. Each fixed effect and interaction was tested and compared using an ANOVA. Random slopes were added for the significant fixed effects.

The model was pruned to retain the effects that significantly improved model fit. The final model included an interaction between speaker dialect and listener dialect, as summarised in Table 1 and visualized in Figure 5. *Jianghuai* was used as the baseline level for both speaker dialect and listener dialect. Listeners were only significantly better at correctly identifying speakers from Beijing (p = 0.001), not from Northern or Wu dialect areas. Furthermore, listeners from Jianghuai dialect area were also found to be significantly better at correctly identifying speaker's origin than non-Jiangsu dialect listeners (p = 0.016). The interaction suggests that listeners from Northern dialect area recognized speakers from Beijing less accurately in comparison to listeners from Jianghuai dialect area, but this difference was only approaching significance (p = 0.053). More importantly, listeners of Wu dialect area recognized speakers of the same dialect area significantly better than Jianghuai listeners (p = 0.040).

To sum up, the effect of speaker dialect indicates that Beijing speakers were correctly identified most compared with Northern, Jianghuai, and Wu dialect area speakers. The effect of listener dialect suggests that Jiangsu dialect area listeners performed significantly better at the identification task than non-Jiangsu listeners. Moreover, listeners from Wu dialect area exhibited an additional advantage at the identification of Wu speakers.

3.3.2 Individual dialects
To investigate the factors affecting the identification of speaker origin within each dialect area, we subset the data into three individual datasets according to the coding of the speaker dialect ("Northern," "Jianghuai'," and "Wu") and fit three separate binomial logistic mixed effects models to the three datasets (each with 720 data points). For the three individual models, we used the binary "yes/no" response as the dependent variable and speaker and listener as random intercepts. The predictors of speaker age, speaker gender, speaker dialect (by city), speaker dialect proficiency, speaker Mandarin rating, listener dialect, and listener dialect proficiency were tested. The model for each individual subset was pruned to only include significant predictors.

For Northern dialect area data, the final model included the fixed effects of speaker gender, listener dialect proficiency, and speaker Mandarin rating (Table 2). The effect of speaker gender indicates that listeners identified the speaker origin of male speakers significantly more correctly than female speakers (Figure 6a). As listener dialect proficiency was coded as a continuous predictor, its effect suggests that listeners who had higher dialect proficiency in their own dialect identified the origin of speakers from Northern dialect area significantly less correctly than listeners who had lower proficiency (Figure 6b). The effect of speaker Mandarin rating revealed that higher Mandarin rating was related to lower identification accuracy. In other words, the higher the listeners rated the Mandarin of the speaker, the less correct the speaker's dialect was identified (Figure 6c).

Table 1. Summary for model of identification of speaker origin

	Estimate	Std. Error	z value	Pr(>|z|)	Sign.
(Intercept)	−1.820	0.276	−6.602	< 0.001	***
Speaker dialect_beijing	1.638	0.514	3.188	0.001	**
Speaker dialect_northern	0.019	0.382	0.049	0.961	
Speaker dialect_wu	−0.330	0.392	−0.843	0.399	
Listener dialect_northern	0.056	0.506	0.111	0.912	
Listener dialect_wu	−0.172	0.286	−0.601	0.548	
Listener dialect_non-jiangsu	−0.719	0.298	−2.414	0.016	*
Speaker dialect_beijing: Listener dialect_northern	−1.927	0.998	−1.932	0.053	.
Speaker dialect_northern: Listener dialect_northern	−0.001	0.684	−0.002	0.998	.
Speaker dialect_wu: Listener dialect_northern	−0.181	0.747	−0.242	0.809	
Speaker dialect_jianghuai: Listener dialect_non-jiangsu	0.169	0.467	0.362	0.717	
Speaker dialect_northern: Listener dialect_non-jiangsu	0.403	0.390	1.033	0.302	
Speaker dialect_wu: Listener dialect_non-jiangsu	0.069	0.431	0.159	0.874	
Speaker dialect_jianghuai: Listener dialect_wu	0.717	0.478	1.501	0.133	
Speaker dialect_northern: Listener dialect_wu	−0.382	0.407	−0.939	0.348	
Speaker dialect_wu: Listener dialect_wu	0.790	0.384	2.057	0.040	*
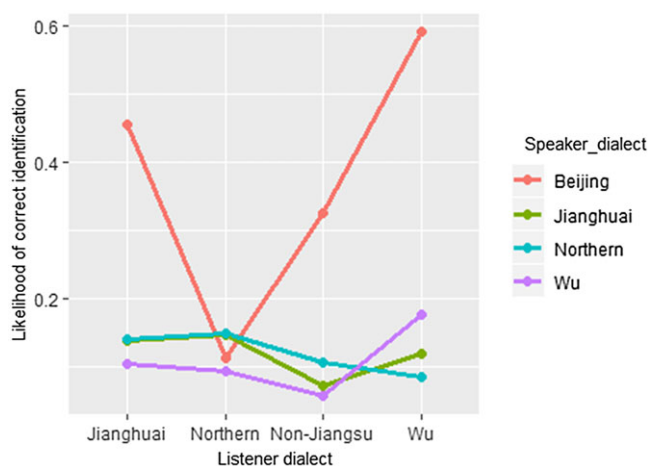
Figure 5. Correct speaker origin identification as predicted by an interaction between speaker dialect and listener dialect.

For Jianghuai dialect area data, the final model included the fixed effects of speaker dialect, listener dialect, and speaker dialect proficiency. In a summary of the model in Table 3, *Nanjing* was the baseline for speaker dialect, and *Jianghuai* was the baseline for listener dialect. The effect of speaker dialect suggests that listeners recognized the Yangzhou dialect speakers significantly more correctly than speakers from Nanjing (Figure 7a). Compared to listeners from Jianghuai dialect area, non-Jiangsu listeners identified the speaker's origin significantly less correctly in Jianghuai dialect area data (Figure 7b). As speakers were all from Jianghuai dialect area in the Jianghuai data, this effect of listener dialect suggests that Jianghuai dialect area listeners were significantly better at correctly identifying Jianghuai dialect area speakers than non-Jiangsu dialect listeners. However, listener origin within Jiangsu province did not play a role. The effect of speaker dialect proficiency shows that the speakers with higher dialect proficiency were identified more correctly (Figure 7c).

For Wu dialect area data, the final model included the fixed effects of speaker dialect and listener dialect (Table 4). *Changzhou* was the baseline for speaker dialect, and *Jianghuai* was the baseline for listener dialect. Compared to Changzhou dialect speakers, the correct identification of Suzhou speakers was significantly better by the listeners (Figure 8a). In the effect of listener dialect, listeners from Wu dialect area were able to identify speakers of Wu dialect area significantly more correctly than listeners from Jianghuai dialect area. However, non-Jiangsu listeners identified Wu dialect area speakers significantly worse than listeners from Jianghuai dialect area (Figure 8b).

4. Discussion

To briefly sum up the results, the overall accuracy for correct identification in this study is 16.2%. There was variation in the correct identification accuracy with Beijing being the highest by far at 44.6%. This observation was supported by a significant effect of speaker dialect with Beijing speakers being more likely to be correctly identified in the analysis considering larger dialect areas.

Table 2. Summary for model of identification of speaker origin for Northern dialect area

	Estimate	Std. Error	z value	Pr(>|z|)	Sign.
(Intercept)	−0.294	0.612	−0.481	0.630	
Speaker gender_male	0.505	0.235	2.153	0.031	*
Listener dialect proficiency	−0.249	0.093	−2.676	0.007	**
Speaker Mandarin rating	−0.216	0.106	−2.036	0.042	*

Table 3. Summary for model of identification of speaker origin for Jianghuai dialect area

	Estimate	Std. Error	z value	Pr(>|z|)	Sign.
(Intercept)	−10.503	3.136	−3.349	0.001	***
Speaker dialect_nantong	−0.357	0.355	−1.00	0.316	
Speaker dialect_yangzhou	0.781	0.394	1.983	0.047	*
Listener dialect_northern	0.121	0.593	0.204	0.838	
Listener dialect_wu	−0.220	0.333	−0.659	0.510	
Listener dialect_non-jiangsu	−0.816	0.348	−2.345	0.019	*
Speaker dialect proficiency	1.760	0.630	2.795	0.005	**
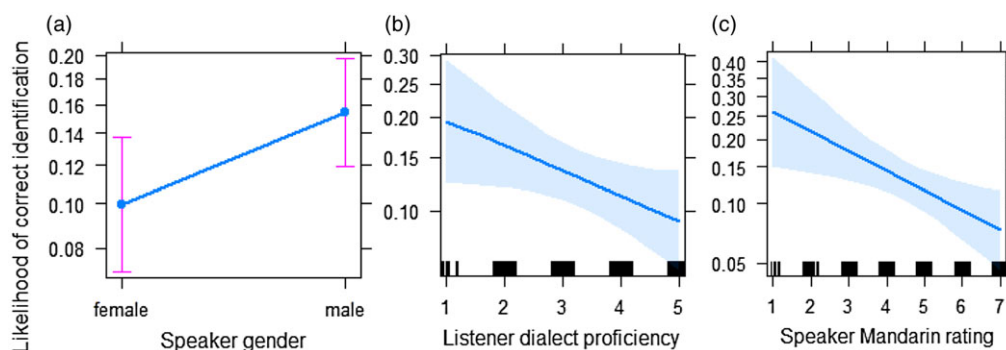
Figure 6. The effects of speaker gender (a), listener dialect proficiency (b), and speaker Mandarin rating (c) on the identification accuracy in Northern dialect area.

Listener dialect was also found to be a significant predictor of accuracy with non-Jiangsu listeners fairing worse. Finally, Wu listeners were particularly better with Wu speakers. In the models focusing on individual dialect areas, speaker gender, listener dialect proficiency, and speaker Mandarin rating were significant predictors of accuracy in speaker origin identification for Northern dialect area data; speaker dialect, listener dialect, and speaker dialect proficiency were significant in Jianghuai dialect area data; and speaker and listener dialect in Wu dialect area data.

This study's overall accuracy for correct identification of 16.2% is much lower compared to 56% for the varieties in Enshi prefecture in Yan (2015), and 63–94% for Kunming in Blum (2004). It is also generally lower than the identification accuracy of English varieties in previous studies (cf., 30% for Welsh English in Williams et al., 1999). One reason might be that instead of dialects, we used local Mandarin, which would have fewer readily available cues for identification. Mandarin has been promoted and popularized in China for over half a century now, and it has become the national language. At least half the speakers and all of the listeners in this study were university-educated, providing them with abundant opportunity for exposure to Standard Mandarin in classroom

and among friends in daily life. The speakers' relatively high Mandarin rating supports this explanation. Thus, using local Mandarin for stimuli makes origin identification a more difficult task.

Considering the size of the variety under investigation and seeing higher identification accuracy for larger dialect areas than individual dialects (e.g., Xuzhou being correctly identified as Xuzhou 11.3% and as a Northern dialect 32.1%), we conclude that larger dialect areas are easier to identify than smaller ones. Similar results have been found elsewhere. In Clopper & Pisoni's (2004a) study, listeners were also found to be better at categorizing the regional dialects of American English into three broad dialect clusters instead of six smaller regions (six regional subvarieties within each of these larger clusters). This is not surprising as more precise identifications would require a more intimate knowledge of subtle dialect features.

We found significant effects of speaker dialect and listener dialect in the identification accuracy of speaker's origin of all speakers. For example, Beijing speakers were most correctly identified compared to speakers from the three dialect areas of Jiangsu province. One reason that listeners found it less difficult to separate Beijing

Table 4. Summary for model of identification of speaker origin for Wu dialect area

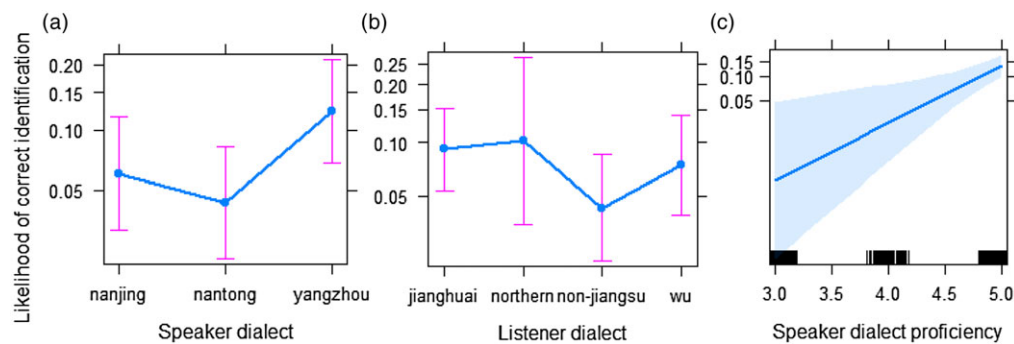|	Estimate	Std. Error	z value	Pr(>|z|)	Sign.
(Intercept)	−2.678	0.376	−7.134	0.000	***
Speaker dialect_suzhou	1.091	0.447	2.438	0.015	*
Speaker dialect_wuxi	0.611	0.459	1.332	0.183	
Listener dialect_northern	−0.122	0.570	−0.214	0.830	
Listener dialect_wu	0.609	0.268	2.276	0.023	*
Listener dialect_non-jiangsu	−0.644	0.321	−2.008	0.045	*

Figure 7. The effects of speaker dialect (a), listener dialect (b), and speaker dialect proficiency (c) on the identification accuracy in Jianghuai dialect area.
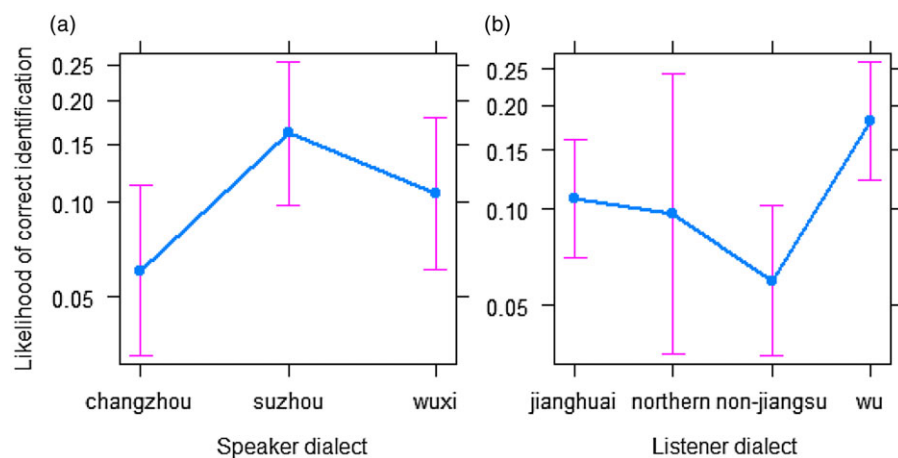

Figure 8. The effects of speaker dialect (a) and listener dialect (b) on the identification accuracy in Wu dialect area.

speakers from other localities might be because Beijing Mandarin is more salient and easily recognized by listeners than Mandarin spoken in the other three dialect areas, as Beijing is the cultural and political capital with a relatively large media presence in the daily lives of Chinese people (cf., Montgomery, 2012).

The finding of Jiangsu listeners being significantly better at correctly identifying the origin of speakers than non-Jiangsu listeners is compatible with the findings in many previous studies that local listeners are better at identifying local speakers than non-local listeners (Baker et al., 2009; Clopper & Pisoni, 2004a; Yan, 2015). The local advantage was also found in the individual models for Jianghuai and Wu dialect areas with non-Jiangsu listeners underperforming in comparison to Jiangsu listeners. When comparing dialect area listeners within Jiangsu province, there was a trend for Northern dialect area listeners to be worse than Jianghuai dialect area listeners at correctly identifying the origin of Beijing

speakers. This divergence may be due to the unbalanced distribution of listeners in this study, as over one-third of the listeners were from Jianghuai dialect area, but only three were from Northern dialect area. Future studies could consider having a more balanced demographic distribution of listeners, especially controlling listener dialect origin. Wu listeners being better at identifying Wu speakers might be due to Wu speakers sounding most different to Beijing Mandarin speakers, resulting in an easier recognition for Wu listeners. In Wu dialect area, listeners identified Suzhou speakers significantly more correctly than Changzhou speakers. Partly, it might be because there were four times as many Suzhou listeners as Changzhou listeners in the study. Another reason might be that Suzhou has a relatively larger amount of media presence nationally.

In addition to the effects of speaker and listener dialect, other significant predictors were identified when analyzing the data of

the three dialect areas separately. In Northern dialect area data, listeners identified the male speakers' origin significantly more correctly than that of female speakers. This finding appears to run contrary to the no gender effect result in our own data for the other two dialect areas and in previous literature, such as Clopper, Conrey and Pisoni (2005). One possible explanation for the effect might be that male speakers in Northern dialect area exhibit more dialect features than female speakers, which would be in line with some sociolinguistic literature showing that females are more likely to use prestige features and varieties (see Gal, 1978). A specific analysis of the speaker recordings and listener comments in questions 2 and 4 would be able to shed light on this issue in the future.

Additionally, we found that variation in proficiency in the dialect, or Mandarin, resulted in variable identification accuracy with higher dialect proficiency in listeners and lower dialect proficiency in speakers being associated with lower accuracy. For example, listeners with lower dialect proficiency identified the origin of speaker significantly more correctly than listeners with higher dialect proficiency in Northern dialect area data. This indicates that higher dialect proficiency of the listener hindered correct identification of speaker's origin. This is most likely reflective of Standard Mandarin exposure and amount of use with listeners being less sensitive to the features in Mandarin speech when they use their dialects more frequently. Similarly, listeners identified speakers with higher dialect proficiency significantly more correctly in Jianghuai dialect area. When speakers are more advanced in their dialect, it would probably make its dialect features more prominent in their Mandarin speech, leading to an easier identification by listeners. This explanation is further supported by the effect of speaker Mandarin rating suggesting that the higher their Mandarin rating, the less correctly the listener identified the origin of the speaker in the Northern dialect area data. The relation between speaker Mandarin rating and the identification accuracy implies that the better a person's Mandarin, the harder the identification is for the listener. The listeners probably rated a speaker's Mandarin higher if it was more "standard," with fewer local features, which in turn would make it more difficult to identify. Although we are not aware of proficiency being considered in other English or Chinese dialect identification studies, Gnevsheva (2018) had a similar observation in a foreign accent identification study such that the higher the speaker's proficiency in English as a second language the more difficult it was for listeners to identify their first language. It seems, then, that proficiency can be fruitfully applied to Chinese dialect identification research.

To bring all the results together, we find that similar factors predict dialect identification accuracy in both Chinese and English. One is participant region of origin. As with the finding in other studies that local listeners identify local varieties more correctly than nonlocal listeners (Baker et al., 2009; Clopper & Pisoni, 2004a; Yan, 2015), a local advantage was found not only in the larger dialect areas in this study (Jiangsu listeners being significantly better at identifying speaker origin than non-Jiangsu listeners), but also within the dialect areas of Jiangsu province (Wu listeners identifying speakers from their dialect better than non-Wu listeners). We also find an effect of cultural prominence (Montgomery, 2012) and geographical closeness. Speakers from culturally prominent Beijing were the most correctly identified compared to speakers from Northern, Jianghuai, and Wu dialect areas of Jiangsu province. In addition, Jianghuai dialect area listeners recognized "home" Jianghuai dialect area speakers and "near to" Wu dialect area speakers more easily. Finally, the found effect of participant dialect proficiency is something that sets this study

apart from previous English and Chinese dialect identification literature but makes it more comparable to foreign accent identification studies (Gnevsheva, 2018).

5. Conclusion

This study investigated origin identification in Jiangsu province, China. Speakers from nine localities (Xuzhou, Suqian, Ganyu, Nanjing, Yangzhou, Nantong, Suzhou, Wuxi, and Changzhou) in three dialect areas (Northern, Jianghuai, and Wu) of Jiangsu province and Beijing were recorded reading a passage in Mandarin. Listeners listened to the recordings of a total of 40 speakers and were asked to identify the origin of the speaker, score their Mandarin, and provide comments on their identification and Mandarin rating of the speaker. In this paper, we focused on the results of identification accuracy, Mandarin rating scores, and the factors that impacted the listeners' identification accuracy of speaker origin in general and for each dialect area of Jiangsu province.

To conclude, this study found significant effects of speaker dialect and listener dialect in the identification of speaker origin. The effect of listener dialect showed that local listeners were more accurate at identifying local speakers than nonlocal listeners. In addition to the effect of region of origin, we found significant effects of speaker gender, speaker dialect proficiency, listener dialect proficiency, and speaker Mandarin rating in the identification accuracy of speaker's region of origin in dialect area data. This paper hopes to make a contribution to the perceptual dialectology of China. Future work can be conducted on analyzing the phonetic characteristics of all recordings to reveal which phonetic features are salient to the identification of speaker origin by listeners, and a wider range of dialect areas consisting of northern and southern Chinese dialects will be attempted in a further dialect perceptual study.

Author Contributions. The first author is the lead author of the study. The second author contributed in study design and writing.

Conflicts of Interests. The authors declare that there are no potential conflicts of interest.

Notes

1 Bureau of Statistics of Jiangsu Province 2018, retrieved from the website of Jiangsu Provincial People's Government http://www.js.gov.cn/art/2019/3/7/art_34153_8258258.html.

2 There are ten dialects in China, including Northern Mandarin, Jin, Wu, Hui, Min, Cantonese, Hakka, Gan, Xiang, and Pinghua and Tuhua (Retrieved from Ministry of Education of the People's Republic of China http://www.moe.gov.cn/jyb_sjzl/wenzi/201903/t20190311_372965.html). Within the broader classification of Chinese dialects, Wu dialect covers speakers from Jiangsu, Zhejiang, Anhui, Jiangxi, and Fujian provinces, and Shanghai municipality. The Wu dialect area here only refers to the Wu dialect speakers within Jiangsu province.

References

Baayen, Harald R., Doug J. Davidson & Douglas M. Bates. 2008. Mixed-effects modelling with crossed random effects for subjects and items. *Journal of Memory and Language* 59. 390–412.

Baker, Wendy, David Eddington & Lyndsey Nay. 2009. Dialect identification: The effects of region of origin and amount of experience. *American Speech* 84(1). 48–71.

Bayard, Donn, Ann Weatherall, Cynthia Gallois & Jeffery Pittam. 2001. Pax Americana? Accent attitudinal evaluations in New Zealand, Australia and America. *Journal of Sociolinguistics* 5(1). 22–49.

Blum, Susan D. 2004. Good to hear: Using the trope of standard to find one's way in a sea of linguistic diversity. In Minglang Zhou (ed.), *Language policy in the People's Republic of China: Theory and practice since 1949*, 123–141. Boston: Kluwer Academic Press.

Cai, Huaxiang. 2015. 浅谈江苏方言的儿化和儿尾 [Discussion on the suffixation of nonsyllabic 'r' in Jiangsu dialects]. *Modern Chinese 4*. 31–33.

Chen, Yachuan. 1991. "地方普通话"的性质特征及其他 [The nature and other features of 'Local Mandarin']. *Chinese Teaching in the World 1*. 12–17.

Chen, Zhangtai. 1990. 关于普通话与方言的几个问题 [Few issues about Mandarin and dialects]. *Language Planning 4*. 27–29.

Clopper, Cynthia G. & Ann R. Bradlow. 2009. Free classification of American English dialects by native and non-native listeners. *Journal of Phonetics* 37(4). 436–51.

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.

Clopper, Cynthia G. & David B. Pisoni. 2004a. Some acoustic cues for the perceptual categorization of American English regional dialects. *Journal of Phonetics* 32(1). 11140.

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

Clopper, Cynthia G. & David B. Pisoni. 2006. Effects of region of origin and geographic mobility on perceptual dialect categorization. *Language Variation and Change 18*(2). 193–221.

Clopper, Cynthia G. & David B. Pisoni. 2007. Free classification of regional dialects of American English. *Journal of Phonetics 35*(3). 421–38.

Coupland, Nikolas & Hywel Bishop. 2007. Ideologised values for British accents. *Journal of Sociolinguistics 11*. 74–93.

Long, Daniel. 1999a. Geographical perceptions of Japanese dialect regions. In Dennis R. Preston (ed.), *Handbook of perceptual dialectology* (Vol. 1), 177–198. Amsterdam & The Netherlands: Benjamins.

Long, Daniel. 1999b. Mapping nonlinguists' evaluations of Japanese language variations. In Dennis R. Preston (ed.), *Handbook of perceptual dialectology* (Vol. 1), 199–226. Amsterdam & The Netherlands: Benjamins.

Long, Daniel & Yim Young-Cheol. 2000. Perceptions of regional variation in Korean. 言语研究 *Gengo Kenkyu 117*. 37–69.

Gal, Susan. 1978. Peasant men can't get wives: Language change and sex roles in a bilingual community. *Language in Society 7*(1). 1–16.

Gao, Jiayin; Pierre Hallé & Christoph Draxler. 2019. Breathy voice and low-register: A case of trading relation in Shanghai Chinese tone perception? *Language and Speech*. 1–26.

Gnevsheva, Ksenia. 2018. Variation in foreign accent identification. *Journal of Multilingual and Multicultural Development 39*(8). 688–702.

Inoue, Fumio. 1999. Classification of dialects by image: English and Japanese. In Dennis R. Preston (ed.), *Handbook of perceptual dialectology* (Vol. 1), 147–59. Amsterdam & The Netherlands: Benjamins.

Inoue, Fumio. 2018. Continuum of Fujian language boundary perception: Dialect division and dialect image. *Dialectologia 20*. 147–80.

Jiangsu Local Chronicles Compilation Committee. 1998. 江苏省志·方言志 [*Local chronicles of Jiangsu Province · Dialect chronicle*]. Nanjing: Nanjing University Press.

Jin, Jian. 2010. 广州方言与东海方言平调的感知研究 [A perception research of level tone patterns of Cantonese and Donghai dialect]. *Dialect 2*. 145–55.

Jin, Jian. 2011. 垫江方言单字调的感知研究 [A perception research on citation tones of Dianjiang dialect]. *Studies in Language and Linguistics 31*(4). 107–14.

Jin, Jian. 2015a. 杭州方言多个升调的感知研究 [A perception study on multiple rising tones in Hangzhou dialect]. *Journal of Sun Yat-Sen University (Social Science Edition) 55*(1). 70–82.

Jin, Jian. 2015b. 同一方言内部起点同高的两种降调——来自闽语惠来方言和赣语茶陵方言的个案 [Two kinds of falling tones of Huilai dialect in Guangdong province and Chaling dialect in Hunan province]. *Dialect 1*. 24–31.

Jin, Jian & Qisheng Shi. 2010. 汕头谷饶方言多个降调的声学分析和感知研究 [An acoustic analysis and perception research on the falling tones in the Shantou Gurao dialect]. *Studies of the Chinese Language 6*. 544–76.

Kingstone, Sydney. 2015. "Scottish", "English" or "foreign": Mapping Scottish dialect perceptions. *English World-Wide 36*(3). 315–47.

Kuznetsova, Alexandra, Per B. Brockhoff & Rune Haubo Bojesen Christensen. 2017. LmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software 82*(13). 1–26.

Lance, Donald M. 1999. Regional variation in subjective dialect divisions in the United States. In Dennis R. Preston (ed.), *Handbook of perceptual dialectology* (Vol. 1), 283–314. Amsterdam & The Netherlands: Benjamins.

Leach, Hannah, Kevin Watson & Ksenia Gnevsheva. 2016. Perceptual dialectology in northern England: Accent recognition, geographical proximity and cultural prominence. *Journal of Sociolinguistics 20*(2). 192–211.

Li, Yuming. 2010. 论中国语言资源有声数据库的建设 [On the establishment of the audio database of Chinese language resources]. *Studies of the Chinese Language 4*. 356–84.

Montgomery, Chris. 2007. *Northern English dialects: A perceptual approach*. Sheffield, UK: University of Sheffield dissertation.

Montgomery, Chris. 2012. The effect of proximity in perceptual dialectology. *Journal of Sociolinguistics 16*(5). 638–68.

Preston, Dennis R. 1989. *Perceptual dialectology: Nonlinguistics' view of areal linguistics*. Dordrecht & The Netherlands: Foris.

R Core Team. 2019. *R: A language and environment for statistical computing* (Version 3.3.3). R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

Tang, Chaoju. 2009. *Mutual intelligibility of Chinese dialects: An experimental approach*. The Netherlands: LOT Publications.

Tang, Zhiqiang & Li Shanpeng. 2018. 扬州方言入声区别性特征的感知研究 [Perceptual studies on the distinctive features of Rusheng Tone in Yangzhou dialect, Jiangsu province]. *Dialect 4*. 411–20.

Wang, Yizhi & Chen Zhongmin. 2016. 吴语全浊塞音声母的感知研究——以上海话为例 [Perceptual studies on voiced initial stops in Wu dialect: Illustrated by the example of Shanghai dialect]. *Studies in Language and Linguistics 36*(2). 44–50.

Weatherall, Ann, Cynthia Gallois & Jeffery Pittam. 1998. Australasians identifying Australasian accents. *Te Reo 41*. 153–62.

Williams, Angie, Peter Garrett & Nikolas Coupland. 1999. Dialect recognition. In Dennis R. Preston (ed.), *Handbook of perceptual dialectology* (Vol. 1), 345–58. Amsterdam & The Netherlands: Benjamins.

Yan, Qingyang. 2015. The perceptual categorization of Enshi Mandarin regional varieties. *Journal of Linguistic Geography 3*. 1–19.

Yao, Youchun. 1989. 应该开展对"地方普通话"的研究 [Research on 'Local Mandarin' should be carried out]. *Language Planning 3*. 18–21.

Zhang, Ruifeng & Kong Jiangping. 2014. 河南禹州方言声调的声学及感知研究 [Acoustic and perceptual studies on the four tones of Yuzhou dialect, Henan province]. *Dialect 3*. 206–14.

Appendix

"The North Wind and the Sun" passage in Chinese.

　　有一回，北风跟太阳正在那儿争论谁的本领大。说着说着，来了一个过路的，身上穿了一件厚袍子。他们俩就商量好了，说，谁能先叫这个过路的把他的袍子脱下来，就算是他的本领大。北风就卯足了劲儿，拼命地吹。可是，他吹得越厉害，那个人就把他的袍子裹得越紧。到末了儿，北风没辙了，只好就算了。一会儿，太阳出来一晒，那个人马上就把袍子脱了下来。所以，北风不得不承认，还是太阳比他的本领大。