Social Attribute Rating of Mandarin Accents by Singaporean Listeners

Irene Yi
11 June 2023

1 Introduction

This study is a partial replication of previous work investigating language attitudes towards different accents of Mandarin. Previous research (Cavallaro et al. 2018; Chong & Tan 2013; a.o.) have used matched guise and verbal guise experiments to probe at Singaporean listeners' social evaluations of different Mandarin accents, such as Chinese Mandarin, Taiwanese Mandarin, and Singaporean Mandarin. These studies separated attributes into "Status" and "Solidarity" traits in their analyses. Thus, the present study is part of a broader project of investigating Singaporean listeners', Chinese listeners', and Taiwanese listeners' social attribute ratings of different accents of Mandarin. However, in this paper, only results from Singaporean listeners will be presented and discussed, making it more directly in conversation with predictions from previous literature.

In this study, a social attribute rating task was carried out in the form of a web-based experiment. Critical trials included clips of Singaporean Mandarin, Beijing Mandarin, Taiwanese Mandarin, and Standard (Chinese) Mandarin in clips of recorded speech (from a reading passage). Filler trials included clips of other Mandarin accents reading those same sentences. The purpose of this experiment was to replicate findings from previous literature on Singaporeans (Cavallaro et al. 2018; Chong & Tan 2013). Additionally, the present experiment used both Beijing Mandarin and Standard Mandarin, which are distinct from one another but often treated as the same (Chinese) variety in perception studies (or the use of one or the other is ambiguous as to which is being represented). The purpose of including these as distinct accents was to help tease apart the social attribute ratings for Beijing Mandarin and Standard Mandarin.

The GitHub repository for this study can be found here: https://github.com/irenegreenbean/social-attribute-rating-mandarin-accents. The OSF preregistration for this study can be found here: https://osf.io/vtn82.

1.1 Background

Languages, varieties, and accents have significant effect on the social evaluation of a given speaker by listeners (Cargile et al. 1994; Dragojevic et al. 2013; a.o.). Notably, accents play a "powerful social force" (Cargile and Giles 1997:195) which can lead to the development and reinforcement of social stereotypes. This is particularly relevant for communities where many languages and accents are represented, and it is made even more interesting when the languages in a given community are also spoken in other communities (e.g. other countries).

Mandarin provides a useful case study for investigating the social meanings that become tied to accents. Mandarin is spoken in many countries, such as China, Taiwan, Hong Kong, Singapore, Malaysia, etc., and spoken in even more diaspora communities across the world. Within each country, many different dialects and accents are present, each with its own social meanings attached to it. Accents such as Beijing Mandarin have been studied and shown to index a unique "Beijingness" (Zhang 2008), and there is even more nuance in who is privy to certain social meanings (Zhang 2017). To this end, it is valuable to have many listener populations (beyond Beijingers, and even beyond just Chinese Mandarin speakers) represented in the social evaluation of accents, so that the use and practice of these accents to index social

meanings can be teased apart in more nuanced ways. This paper begins to ask these questions by starting with a replication of social attribute rating tasks from previous literature, specifically with Singaporeans' perceptions of Mandarin accents.

Mandarin in Singapore has grown in the last few decades with the Singapoean government's implementation of the Speak Mandarin Campaign in 1979 (Bokhorst-Heng 1999). As a result of this campaign, from 1980 to 2010, the household use of Mandarin in Singapore grew over 30% (Department of Statistics 2011), and the sentiment grew that there was a new variety of Mandarin "nurtured in Singapore's soil" (Wang 2002:27). Alongside Singaporean Mandarin, the influence of Chinese Mandarin and Taiwanese Mandarin is also present in Singapore. Chinese Mandarin is seen as the gold standard of Mandarin education, and Taiwanese Mandarin is quite present in social media and pop culture, as many celebrities known in China and Singapore are Taiwanese artists and actors (Chong & Tan 2013). Thus, the question of Singaporean perceptions of different Mandarin accents is interesting and can speak to how the different nature of exposure to accents (e.g. in education, in a colloquial setting, in media, etc.) can influence social evaluations of those accents.

1.2 Predictions

Cavallaro et al. (2018) predicted and found that Singaporean listeners rated "Chinese" Mandarin (which appeared to be similar to Standard Mandarin) as higher in status than Singaporean Mandarin. Chong and Tan (2013) found that Singaporean participants rated Taiwanese Mandarin lower in status and solidarity than the guises they used for Beijing Mandarin, and that Singaporeans rated Singaporean Mandarin as higher in solidarity than Taiwanese Mandarin and Beijing Mandarin. Following this literature, the predictions of the present study are as follows:

Singaporean participants will rate one or both of the Chinese accents (Standard/Beijing) as higher in status than both Taiwanese Mandarin and Singaporean Mandarin.

Singaporean participants will rate Singaporean Mandarin as highest in solidarity out of all the accents. They will rate one or both of the Chinese accents (Standard/Beijing) as higher in solidarity than Taiwanese Mandarin.

Because many of these predictions compare against one of the Chinese accents, the baseline for comparison in the current paper's models was Beijing Mandarin.

2 The Experiment

This section consists of two subsections, Methods and Results.

2.1 Methods

2.1.1 Participants

In total, 74 Singaporean participants were hand-recruited through Telegram (a messenger app) in a Singapore-based research group chat called "NTU/NUS paid survey". Of this total, 27 participants were excluded due to failing attention checks. Data from the remaining 47 participants (28 female, 17 male, 2 prefer not to answer) were analyzed in this paper. These participants were all Singaporean citizens or Singapore Permanent Residents.

2.1.2. Materials

There were 40 audio stimuli in this verbal guise experiment. These clips came from a Mandarin reading passage, which 8 speakers recorded using their respective Mandarin accent, in a manner that was as

natural as possible. All 8 speakers recorded the same passage. Then, 5 sentences were identified from that passage to be extracted as stimuli clips. All passages had the same 5 sentences extracted. Each clip ranged from about 4 seconds long to 7 seconds long. The clips were then normalized for loudness (to 25.0 dB) using Audacity. Of the 8 total speakers, 4 speakers represented the crucial accents: Standard Mandarin, Beijing Mandarin, Singaporean Mandarin, and Taiwanese Mandarin. The other 4 speakers represented other Mainland Chinese accents, such as Tianjin, Northeast, Sichuan, and Guangdong; these constituted the filler trials of this experiment. Thus, there were 20 critical trials and 20 filler trials. The present paper will only discuss the critical trials.

Table 1 shows the conditions that were used in the current experiment. The full list of conditions in the broader project (i.e. with Chinese and Taiwanese listener populations) is shown in the OSF preregistration given above.

Accent	Listener
Beijing Mandarin	Singaporean listener
Singaporean Mandarin	Singaporean listener
Standard Mandarin	Singaporean listener
Taiwanese Mandarin	Singaporean listener

Table 1: Conditions present in this experiment.

2.1.3 Procedure
The attribute rating task was coded as an experiment using jsPsych (de Leeuw et al. 2023). The whole experiment (consent, instructions, stimuli, follow-up questionnaire, etc.) was presented to participants in Mandarin Chinese, written with Chinese characters. The experiment is described here in English, but the Chinese prompts can be found on the GitHub repository for this experiment.

In the attribute rating task, participants were presented with a randomly selected clip (from the 40 clips described above), which they listened to. They were then asked to rate their agreement with 8 statements; each statement had the format of "This speaker is ____", where the blank line was replaced with one of 8 attributes. These 8 attributes were Kind, Honest, Humorous, Trustworthy, Confident, Intelligent, Educated, and Wealthy. Participants rated their agreement using a 7-point Likert scale, where the levels of agreement were (from 1-7) "Strongly Disagree", "Disagree", "Slightly Disagree", "Neutral", "Slightly Agree", "Agree", and "Strongly Agree". Participants had the option to replay the clip as many times as they wanted while they rated the statements. Participants were then asked to guess where they thought the speaker was from, by filling in a blank field. The dependent variable, then, was the social attribute rating score on the Likert scale.

The 40 clips were ordered randomly for every participant. The order of attributes was shuffled **between** participants, so each participant saw a different order of attribute statements from another participant, but the order of attribute statements stayed consistent **within** a participant, so participants always saw the same order for all the trials they went through.

After the attribute rating task, participants answered a follow-up questionnaire that asked them for their age, gender, citizenship status, language background, the place they currently live, and whether they have moved (and if so, where).

The experiment had 3 attention check trials which were randomly interspersed within the attribute rating task. For these trials, participants heard an audio clip which instructed them to select one specific level of agreement for *all* the statements in that trial. An attention check was considered failed if a participant selected the wrong level of agreement for one or more statements in that trial. Participants who failed one or more attention check trials were excluded from data analysis.

2.2 Results

After excluding 27 participants who failed attention checks, results from 47 participants were analyzed. The Likert scale ratings were coded from 1 to 7, where 1 was "Strongly Disagree" and 7 was "Strongly Agree". In other words, the coding was as follows: "Strongly Disagree" = 1, "Disagree" = 2, "Slightly Disagree" = 3, "Neutral" = 4, "Slightly Agree" = 5, "Agree" = 6, and "Strongly Agree" = 7. Figures 1-8 show the mean Likert ratings (with error bars) across stimuli accent, for each attribute. The Taiwanese Mandarin stimuli and Singaporean Mandarin stimuli were rated, on average, lower than the Standard Mandarin and Beijing Mandarin stimuli for most (but not all) attributes; models for each specific attribute are discussed below.
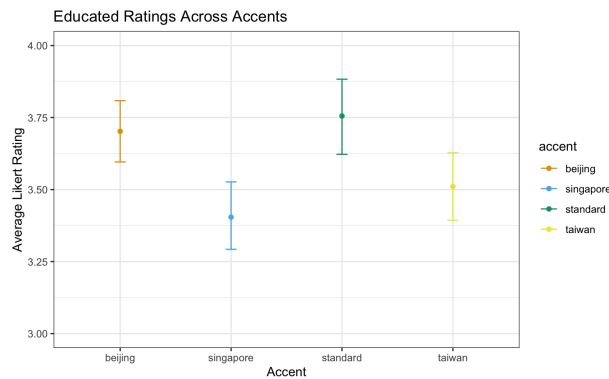

Figure 1: Mean ratings for attribute "Educated" across accents. Error bars represent 95% confidence intervals.
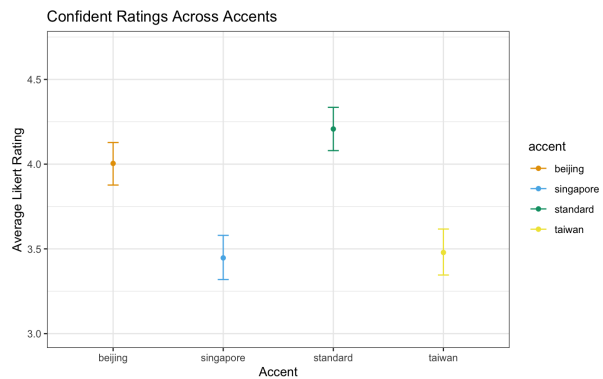

Figure 2: Mean ratings for attribute "Confident" across accents. Error bars represent 95% confidence intervals.
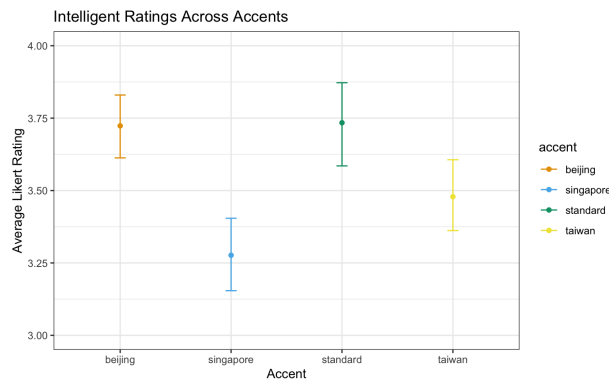

Figure 3: Mean ratings for attribute "Intelligent" across accents. Error bars represent 95% confidence intervals.
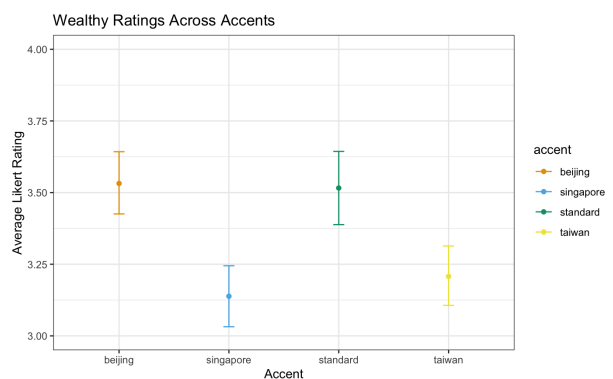

Figure 4: Mean ratings for attribute "Wealthy" across accents. Error bars represent 95% confidence intervals.
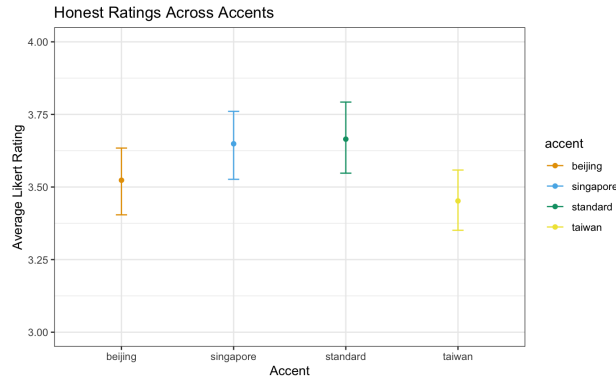
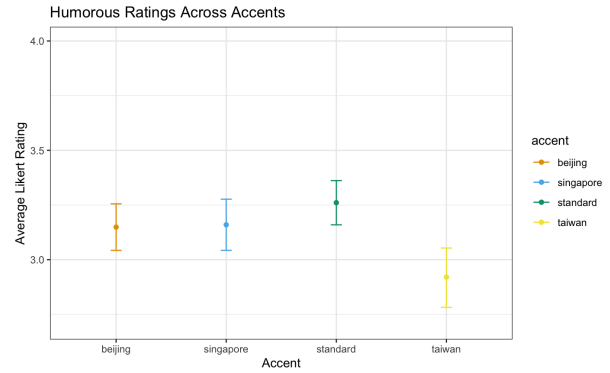Figure 5: Mean ratings for attribute "Honest" across accents. Error bars represent 95% confidence intervals.

Figure 6: Mean ratings for attribute "Humorous" across accents. Error bars represent 95% confidence intervals.
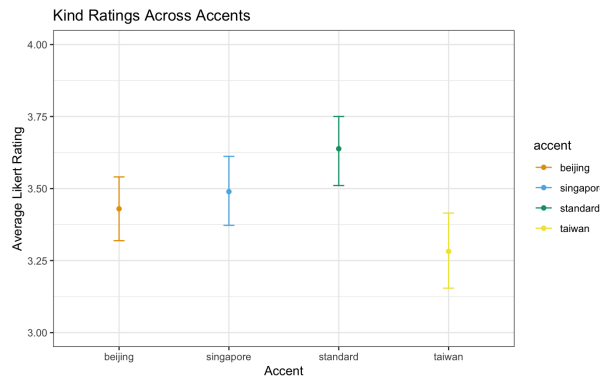

Figure 7: Mean ratings for attribute "Kind" across accents. Error bars represent 95% confidence intervals.
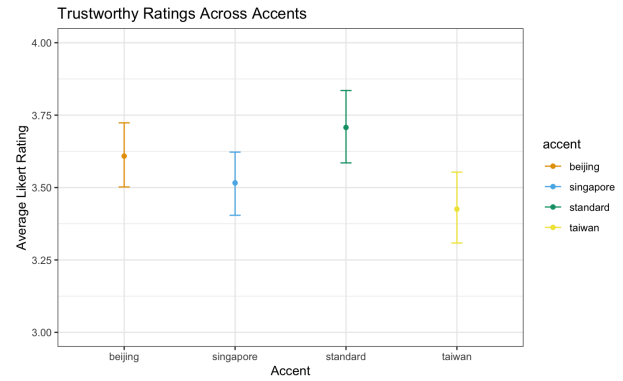
Figure 8: Mean ratings for attribute "Trustworthy" across accents. Error bars represent 95% confidence intervals.

In addition to examining each individual attribute, the 8 attributes were separated into Status traits and Solidarity traits; since the present study's hypotheses draw on differences in ratings for Status and Solidarity traits across accents, this separation is drawn from previous literature which used Principal Component Analyses to delineate these two clusters, Status and Solidarity (Chong & Tan 2013; Cavallaro et al. 2018). Table 2 shows the makeup of Status and Solidarity traits.

Status Traits	Solidarity Traits
Educated	Honest
Confident	Humorous
Intelligent	Kind
Wealthy	Trustworthy

Table 2: Status and Solidarity attributes.

Figures 9 and 10 show, respectively, the mean ratings for aggregated Status and Solidarity traits across stimuli accent.

Figure 9: Mean ratings for Status attributes across accents. Error bars represent 95% confidence intervals.

Figure 10: Mean ratings for Solidarity attributes across accents. Error bars represent 95% confidence intervals.
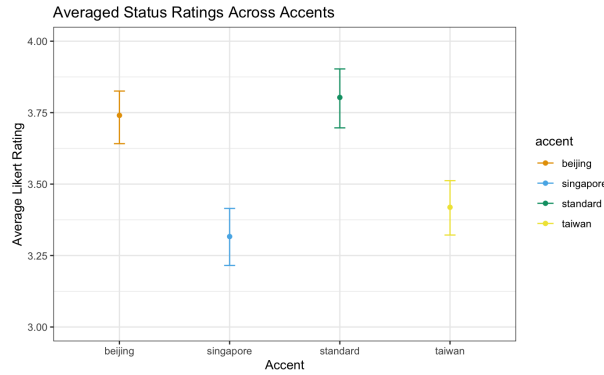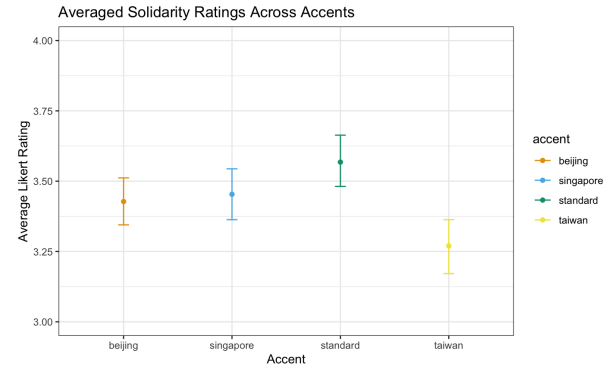
Using the brms package in R (Bürkner 2017), Bayesian ordinal regression models were fit to this data in order to test whether the accent of the stimuli had an effect on the Likert scale ratings. These ordinal regression models were fit for each individual attribute, and then for the aggregated Status and Solidarity traits. In selecting which ordinal regression model to use, a cumulative link model and an adjacent category model were run on the dataset for the attribute, Educated. Leave-one-out (LOO) cross validation (Vasishth et al. 2018) was used to compare these models.

The results of the LOO cross validation are shown in Table 3. The LOO Information Criterion (LOOIC) for the cumulative link model (shown by m.logit.educated) is lower than the LOOIC for the adjacent category model (m.acat.educated), meaning the cumulative link model fits the data better (Bürkner & Vuorre 2019). Additionally, the assumption underlying the cumulative link model is that the variable being predicted comes from something underlyingly continuous, and not categorical. This fits with the present experiment, as the predicted variable is the agreement with an attribute statement, which likely stems from a continuous range of agreement degree.

Model	LOOIC Estimate	SE
m.logit.educated	1603.0	62.9
m.acat.educated	1673.7	66.7
m.logit.educated - m.acat.educated	-35.3	6.8

Table 3: Results of comparing the cumulative link model to the adjacent category model.

Cumulative link ordinal regression models were fit for each attribute, for a total of 8 models examining individual attributes. An example of the model for Intelligent is shown in Figure 11. The outcome variable was the rating for the respective attribute, with stimuli accent as the fixed effect. By-clip and by-participant random intercepts were also included. No priors were specified.

```
m.logit.intelligent = brm(intelligent ~ accent + (1|id) + (1|clip),
        data=d,
        family=cumulative(),
        cores=4,
        save_all_pars = TRUE)
```

Figure 11: brms model for Intelligent.

For each model, a null model was also generated and compared against the full model. Each respective null model was the same as the full model, except that they only included the random intercepts, and no fixed effects. The Bayes Factor was calculated to compare the null model to the full model. Table 4 shows a summary of the Bayes Factors for all the model comparisons. The model in the numerator was always the full model. Table 4 shows that all models had Bayes Factors greater than 3, which is generally taken as evidence in favor of the model in the numerator (Schmalz et al. 2021), though the full models for Honest and Kind have smaller Bayes Factors than the rest of the models.

Model	Bayes Factor
Educated	165.6667
Wealthy	1999
Intelligent	3999
Confident	>9999
Trustworthy	665.6667
Kind	69.17544
Honest	6.352941
Humorous	999
Status Attributes	147.1481
Solidarity Attributes	>9999

Table 4: Bayes Factor for all models.

The rest of this section will first present each attribute model individually, and then the aggregated Status and Solidarity models at the end. The first four attribute models will be individual Status attributes, and the last four attribute models will be individual Solidarity attributes.

2.2.1 Educated

The results of the ordinal regression model for the attribute Educated are shown in Table 5. This regression model, and all subsequent regression models discussed in this paper, were coded with **Beijing Mandarin as its zero level**. From Figure 1 above, the mean ratings for Beijing Mandarin and Standard Mandarin appeared to be higher than the mean ratings for Taiwanese Mandarin and Singaporean Mandarin.

Predictor (for attribute: Educated)	Estimate	Est.Error	l-95% CI	u-95%
First threshold, coded as Intercept[1]	-7.04	0.68	-8.49	-5.78
Second threshold, coded as Intercept[2]	-5.63	0.44	-6.53	-4.78
Third threshold, coded as Intercept[3]	-4.27	0.36	-5.01	-3.59
Fourth threshold, coded as Intercept[4]	-0.39	0.31	-1.00	0.21
Fifth threshold, coded as Intercept[5]	2.24	0.32	1.61	2.89
Sixth threshold, coded as Intercept[6]	6.19	0.54	5.17	7.34
Comparison between Singaporean and Beijing accent, coded as accentsingapore	-0.88	0.25	-1.37	-0.38
Comparison between Standard and Beijing accent, accentstandard	0.25	0.24	-0.22	0.72
Comparison between Taiwanese and Beijing accent, accenttaiwan	-0.67	0.24	-1.14	-0.20

Table 5: Regression model of attribute Educated.

The 95% credible interval for comparing ratings for Beijing Mandarin to Singaporean Mandarin was between -1.37 and -0.38, which does not include 0. The Estimate for this comparison is -0.88. This suggests that we can conclude with 95% confidence that Singaporean Mandarin is more likely to be rated lower than Beijing Mandarin for the Educated attribute.

The 95% credible interval for comparing ratings for Beijing Mandarin to Standard Mandarin was between -0.22 and 0.72, which does include 0. This suggests that we do not have sufficient evidence to conclude that Beijing Mandarin is rated differently from Standard Mandarin for the Educated attribute.

The 95% credible interval for comparing ratings for Beijing Mandarin to Taiwanese Mandarin was between -1.14 and -0.20, which does not include 0. The Estimate for this comparison is -0.67. This suggests that we can conclude with 95% confidence that Taiwanese Mandarin is more likely to be rated lower than Beijing Mandarin for the Educated attribute. Taken together, we can conclude that Beijing Mandarin is rated higher than Taiwanese Mandarin and Singaporean Mandarin on the Educated attribute, but that Beijing Mandarin is not rated as significantly different from Standard Mandarin.

2.2.2 Confident

The results of the ordinal regression model for the attribute Confident are shown in Table 6. From Figure 2 above, the mean ratings for Beijing Mandarin and Standard Mandarin appeared to be higher than the mean ratings for Taiwanese Mandarin and Singaporean Mandarin. Standard Mandarin appears to be rated slightly higher than Beijing Mandarin.

Predictor (for attribute: Confident)	Estimate	Est.Error	l-95% CI	u-95%
First threshold, coded as Intercept[1]	-7.29	0.65	-8.74	-6.12
Second threshold, coded as Intercept[2]	-6.06	0.43	-6.94	-5.24
Third threshold, coded as Intercept[3]	-4.21	0.33	-4.87	-3.57
Fourth threshold, coded as Intercept[4]	-1.19	0.28	-1.76	-0.65
Fifth threshold, coded as Intercept[5]	1.03	0.28	0.45	1.57
Sixth threshold, coded as Intercept[6]	4.53	0.38	3.79	5.27
Comparison between Singaporean and Beijing accent, coded as accentsingapore	-1.40	0.25	-1.91	-0.92
Comparison between Standard and Beijing accent, accentstandard	0.46	0.25	-0.04	0.95
Comparison between Taiwanese and Beijing accent, accenttaiwan	-1.41	0.25	-1.94	-0.92

Table 6: Regression model of attribute Confident.

The 95% credible interval for comparing ratings for Beijing Mandarin to Singaporean Mandarin was between -1.91 and -0.92, which does not include 0. The Estimate for this comparison is -1.40. This suggests that we can conclude with 95% confidence that Singaporean Mandarin is more likely to be rated lower than Beijing Mandarin for the Confident attribute.

The 95% credible interval for comparing ratings for Beijing Mandarin to Standard Mandarin was between -0.04 and 0.95, which does include 0. This suggests that we do not have sufficient evidence to conclude that Beijing Mandarin is rated differently from Standard Mandarin for the Confident attribute.

The 95% credible interval for comparing ratings for Beijing Mandarin to Taiwanese Mandarin was between -1.94 and -0.92, which does not include 0. The Estimate for this comparison is -1.41. This suggests that we can conclude with 95% confidence that Taiwanese Mandarin is more likely to be rated lower than Beijing Mandarin for the Confident attribute. Taken together, we can conclude that Beijing Mandarin is rated higher than Taiwanese Mandarin and Singaporean Mandarin on the Confident attribute, but that Beijing Mandarin is not rated as significantly different from Standard Mandarin.

2.2.3 Intelligent

The results of the ordinal regression model for the attribute Intelligent are shown in Table 7. From Figure 3 above, the mean ratings for Beijing Mandarin and Standard Mandarin appeared to be higher than the mean ratings for Taiwanese Mandarin and Singaporean Mandarin.

Predictor (for attribute: Intelligent)	Estimate	Est.Error	l-95% CI	u-95%
First threshold, coded as Intercept[1]	-7.19	0.67	-8.64	-5.99
Second threshold, coded as Intercept[2]	-5.46	0.41	-6.28	-4.67
Third threshold, coded as Intercept[3]	-4.26	0.35	-4.95	-3.58
Fourth threshold, coded as Intercept[4]	-0.46	0.30	-1.02	0.13
Fifth threshold, coded as Intercept[5]	2.19	0.31	1.61	2.83
Sixth threshold, coded as Intercept[6]	6.05	0.54	5.06	7.12
Comparison between Singaporean and Beijing accent, coded as accentsingapore	-1.28	0.23	-1.73	-0.84
Comparison between Standard and Beijing accent, accentstandard	0.12	0.21	-0.28	0.54
Comparison between Taiwanese and Beijing accent, accenttaiwan	-0.81	0.22	-1.23	-0.39

Table 7: Regression model of attribute Intelligent.

The 95% credible interval for comparing ratings for Beijing Mandarin to Singaporean Mandarin was between -1.74 and -0.84, which does not include 0. The Estimate for this comparison is -1.28. This suggests that we can conclude with 95% confidence that Singaporean Mandarin is more likely to be rated lower than Beijing Mandarin for the Intelligent attribute.

The 95% credible interval for comparing ratings for Beijing Mandarin to Standard Mandarin was between -0.28 and 0.54, which does include 0. This suggests that we do not have sufficient evidence to conclude that Beijing Mandarin is rated differently from Standard Mandarin for the Intelligent attribute.

The 95% credible interval for comparing ratings for Beijing Mandarin to Taiwanese Mandarin was between -1.23 and -0.39, which does not include 0. The Estimate for this comparison is -0.81. This suggests that we can conclude with 95% confidence that Taiwanese Mandarin is more likely to be rated lower than Beijing Mandarin for the Intelligent attribute. Taken together, we can conclude that Beijing Mandarin is rated higher than Taiwanese Mandarin and Singaporean Mandarin on the Intelligent attribute, but that Beijing Mandarin is not rated as significantly different from Standard Mandarin.

2.2.4 Wealthy

The results of the ordinal regression model for the attribute Wealthy are shown in Table 8. From Figure 4 above, the mean ratings for Beijing Mandarin and Standard Mandarin appeared to be higher than the mean ratings for Taiwanese Mandarin and Singaporean Mandarin.

Predictor (for attribute: Wealthy)	Estimate	Est.Error	l-95% CI	u-95%
First threshold, coded as Intercept[1]	-6.26	0.50	-7.28	-5.36
Second threshold, coded as Intercept[2]	-4.29	0.38	-5.07	-3.55
Third threshold, coded as Intercept[3]	0.49	0.33	-0.16	1.15
Fourth threshold, coded as Intercept[4]	2.50	0.35	1.82	3.18
Fifth threshold, coded as Intercept[5]	5.79	0.52	4.78	6.88
Comparison between Singaporean and Beijing accent, coded as accentsingapore	-1.25	0.29	-1.86	-0.68
Comparison between Standard and Beijing accent, accentstandard	0.04	0.28	-0.54	0.60
Comparison between Taiwanese and Beijing accent, accenttaiwan	-1.08	0.30	-1.69	-0.51

Table 8: Regression model of attribute Wealthy.

The 95% credible interval for comparing ratings for Beijing Mandarin to Singaporean Mandarin was between -1.86 and -0.68, which does not include 0. The Estimate for this comparison is -1.25. This suggests that we can conclude with 95% confidence that Singaporean Mandarin is more likely to be rated lower than Beijing Mandarin for the Wealthy attribute.

The 95% credible interval for comparing ratings for Beijing Mandarin to Standard Mandarin was between -0.54 and 0.60, which does include 0. This suggests that we do not have sufficient evidence to conclude that Beijing Mandarin is rated differently from Standard Mandarin for the Wealthy attribute.

The 95% credible interval for comparing ratings for Beijing Mandarin to Taiwanese Mandarin was between -1.69 and -0.51, which does not include 0. The Estimate for this comparison is -1.08. This suggests that we can conclude with 95% confidence that Taiwanese Mandarin is more likely to be rated lower than Beijing Mandarin for the Wealthy attribute. Taken together, we can conclude that Beijing Mandarin is rated higher than Taiwanese Mandarin and Singaporean Mandarin on the Wealthy attribute, but that Beijing Mandarin is not rated as significantly different from Standard Mandarin.

2.2.5 Honest
The results of the ordinal regression model for the attribute Honest are shown in Table 9. From Figure 5 above, the mean ratings for Standard Mandarin and Singaporean Mandarin appeared to be slightly higher than the mean ratings for Taiwanese Mandarin and Beijing Mandarin.

Predictor (for attribute: Honest)	Estimate	Est.Error	l-95% CI	u-95%
First threshold, coded as Intercept[1]	-6.69	0.68	-8.14	-5.51
Second threshold, coded as Intercept[2]	-5.57	0.47	-6.51	-4.66
Third threshold, coded as Intercept[3]	-3.92	0.36	-4.64	-3.23
Fourth threshold, coded as Intercept[4]	0.08	0.31	-0.55	0.68
Fifth threshold, coded as Intercept[5]	2.88	0.33	2.25	3.55
Sixth threshold, coded as Intercept[6]	6.77	0.55	5.74	7.92
Comparison between Singaporean and Beijing accent, coded as accentsingapore	0.41	0.25	-0.08	0.90
Comparison between Standard and Beijing accent, accentstandard	0.36	0.25	-0.13	0.85
Comparison between Taiwanese and Beijing accent, accenttaiwan	-0.27	0.25	-0.76	0.23

Table 9: Regression model of attribute Honest.

The 95% credible interval for comparing ratings for Beijing Mandarin to Singaporean Mandarin was between -0.08 and 0.90, which does include 0. This suggests that we do not have sufficient evidence to conclude that Beijing Mandarin is rated differently from Singaporean Mandarin for the Honest attribute.

The 95% credible interval for comparing ratings for Beijing Mandarin to Standard Mandarin was between -0.13 and 0.85, which does include 0. This suggests that we do not have sufficient evidence to conclude that Beijing Mandarin is rated differently from Standard Mandarin for the Honest attribute.

The 95% credible interval for comparing ratings for Beijing Mandarin to Taiwanese Mandarin was between -1.76 and 0.23, which does include 0. This suggests that we do not have sufficient evidence to conclude that Beijing Mandarin is rated differently from Taiwanese Mandarin for the Honest attribute.

Taken together, we can conclude that Beijing Mandarin is not rated as significantly different from Singaporean Mandarin, Standard Mandarin, or Taiwanese Mandarin for the Honest attribute.

2.2.6 Humorous
The results of the ordinal regression model for the attribute Honest are shown in Table 10. From Figure 6 above, the mean ratings for Standard Mandarin appeared to be slightly higher than the rest, and the mean ratings for Taiwanese Mandarin appeared to be slightly lower than the rest.

Predictor (for attribute: Humorous)	Estimate	Est.Error	l-95% CI	u-95%
First threshold, coded as Intercept[1]	-6.73	0.60	-8.00	-5.63
Second threshold, coded as Intercept[2]	-4.34	0.35	-5.04	-3.68
Third threshold, coded as Intercept[3]	-2.71	0.30	-3.31	-2.12
Fourth threshold, coded as Intercept[4]	1.44	0.29	0.89	2.01
Fifth threshold, coded as Intercept[5]	3.68	0.34	3.02	4.35
Sixth threshold, coded as Intercept[6]	6.28	0.53	5.27	7.36
Comparison between Singaporean and Beijing accent, coded as accentsingapore	-0.02	0.23	-0.47	0.44
Comparison between Standard and Beijing accent, accentstandard	0.26	0.23	-0.20	0.72
Comparison between Taiwanese and Beijing accent, accenttaiwan	-0.87	0.24	-1.32	-0.40

Table 10: Regression model of attribute Humorous.

The 95% credible interval for comparing ratings for Beijing Mandarin to Singaporean Mandarin was between -0.47 and 0.44, which does include 0. This suggests that we do not have sufficient evidence to conclude that Beijing Mandarin is rated differently from Singaporean Mandarin for the Humorous attribute.

The 95% credible interval for comparing ratings for Beijing Mandarin to Standard Mandarin was between -0.20 and 0.72, which does include 0. This suggests that we do not have sufficient evidence to conclude that Beijing Mandarin is rated differently from Standard Mandarin for the Humorous attribute.

The 95% credible interval for comparing ratings for Beijing Mandarin to Taiwanese Mandarin was between -1.32 and -0.40, which does not include 0. The Estimate for this comparison is -0.87. This suggests that we can conclude with 95% confidence that Taiwanese Mandarin is more likely to be rated lower than Beijing Mandarin for the Humorous attribute.

Taken together, we can conclude that Beijing Mandarin is rated higher than Taiwanese Mandarin for the Humorous attribute, but that Beijing Mandarin is not rated as significantly different from Standard Mandarin and Singaporean Mandarin for the Humorous attribute.

2.2.7 Kind
The results of the ordinal regression model for the attribute Kind are shown in Table 11. From Figure 7 above, the mean ratings for Standard Mandarin appeared to be slightly higher than the rest, and the mean ratings for Taiwanese Mandarin appeared to be slightly lower than the rest.

Predictor (for attribute: Kind)	Estimate	Est.Error	l-95% CI	u-95%
First threshold, coded as Intercept[1]	-6.23	0.59	-7.51	-5.17
Second threshold, coded as Intercept[2]	-5.01	0.41	-5.83	-4.21
Third threshold, coded as Intercept[3]	-3.51	0.32	-4.16	-2.88
Fourth threshold, coded as Intercept[4]	0.58	0.28	0.03	1.14
Fifth threshold, coded as Intercept[5]	2.77	0.31	2.19	3.39
Sixth threshold, coded as Intercept[6]	6.28	0.50	5.35	7.30
Comparison between Singaporean and Beijing accent, coded as accentsingapore	0.26	0.23	-0.20	0.70
Comparison between Standard and Beijing accent, accentstandard	0.57	0.22	0.12	0.99
Comparison between Taiwanese and Beijing accent, accenttaiwan	-0.50	0.23	-0.95	-0.06

Table 11: Regression model of attribute Kind.

The 95% credible interval for comparing ratings for Beijing Mandarin to Singaporean Mandarin was between -0.20 and 0.70, which does include 0. This suggests that we do not have sufficient evidence to conclude that Beijing Mandarin is rated differently from Singaporean Mandarin for the Kind attribute.

The 95% credible interval for comparing ratings for Beijing Mandarin to Standard Mandarin was between 0.12 and 0.99, which does not include 0. The Estimate for this comparison is 0.57. This suggests that we can conclude with 95% confidence that Standard Mandarin is more likely to be rated higher than Beijing Mandarin for the Kind attribute.

The 95% credible interval for comparing ratings for Beijing Mandarin to Taiwanese Mandarin was between -0.95 and -0.06, which does not include 0. The Estimate for this comparison is -0.50. This suggests that we can conclude with 95% confidence that Taiwanese Mandarin is more likely to be rated lower than Beijing Mandarin for the Kind attribute.

Taken together, we can conclude that Beijing Mandarin is rated higher than Taiwanese Mandarin and lower than Standard Mandarin for the Kind attribute, but that Beijing Mandarin is not rated as significantly different from Singaporean Mandarin.

2.2.8 Trustworthy
The results of the ordinal regression model for the attribute Intelligent are shown in Table 12. From Figure 8 above, the mean ratings for Beijing Mandarin and Standard Mandarin appeared to be slightly higher than the mean ratings for Taiwanese Mandarin and Singaporean Mandarin.

Predictor (for attribute: Trusthworthy)	Estimate	Est.Error	l-95% CI	u-95%
First threshold, coded as Intercept[1]	-7.1	0.78	-8.81	-5.77
Second threshold, coded as Intercept[2]	-5.46	0.45	-6.39	-4.62
Third threshold, coded as Intercept[3]	-3.73	0.33	-4.39	-3.11
Fourth threshold, coded as Intercept[4]	-0.17	0.29	-0.72	0.39
Fifth threshold, coded as Intercept[5]	2.41	0.3	1.83	3.01
Sixth threshold, coded as Intercept[6]	6.03	0.5	5.15	7.06
Comparison between Singaporean and Beijing accent, coded as accentsingapore	-0.27	0.22	-0.69	0.17
Comparison between Standard and Beijing accent, accentstandard	0.22	0.22	-0.2	0.65
Comparison between Taiwanese and Beijing accent, accenttaiwan	-0.64	0.22	-1.09	-0.21

Table 12: Regression model of attribute Trustworthy.

The 95% credible interval for comparing ratings for Beijing Mandarin to Singaporean Mandarin was between -0.69 and 0.17, which does include 0. This suggests that we do not have sufficient evidence to conclude that Beijing Mandarin is rated differently from Singaporean Mandarin for the Trustworthy attribute.

The 95% credible interval for comparing ratings for Beijing Mandarin to Standard Mandarin was between -0.20 and 0.65, which does include 0. This suggests that we do not have sufficient evidence to conclude that Beijing Mandarin is rated differently from Standard Mandarin for the Trustworthy attribute.

The 95% credible interval for comparing ratings for Beijing Mandarin to Taiwanese Mandarin was between -1.09 and -0.21, which does not include 0. The Estimate for this comparison is -0.64. This suggests that we can conclude with 95% confidence that Taiwanese Mandarin is more likely to be rated lower than Beijing Mandarin for the Trustworthy attribute.

Taken together, we can conclude that Beijing Mandarin is rated higher than Taiwanese Mandarin for the Trustworthy attribute, but that Beijing Mandarin is not rated as significantly different from Standard Mandarin and Singaporean Mandarin for the Trustworthy attribute.

2.2.9 Status Attributes

The Status traits Educated, Confident, Intelligent, and Wealthy were aggregated into Status traits. These were aggregated by taking the means of every rating for these four traits. The results of the ordinal regression model for Status attributes are shown in Table 13. From Figure 9 above, the mean ratings for Beijing Mandarin and Standard Mandarin appeared to be higher than the mean ratings for Taiwanese Mandarin and Singaporean Mandarin.

Predictor (for Status attributes)	Estimate	Est.Error	l-95% CI	u-95%
Intercept	3.74	0.09	3.56	3.90
Comparison between Singaporean and Beijing accent, coded as accentsingapore	-0.42	0.06	-0.55	-0.29
Comparison between Standard and Beijing accent, accentstandard	0.06	0.06	-0.06	0.19
Comparison between Taiwanese and Beijing accent, accenttaiwan	-0.32	0.06	-0.44	-0.20

Table 13: Regression model of Status attributes.

The 95% credible interval for comparing ratings for Beijing Mandarin to Singaporean Mandarin was between -0.55 and -0.29, which does not include 0. The Estimate for this comparison is -0.42. This suggests that we can conclude with 95% confidence that Singaporean Mandarin is more likely to be rated lower than Beijing Mandarin for Status attributes.

The 95% credible interval for comparing ratings for Beijing Mandarin to Standard Mandarin was between -0.06 and 0.19, which does include 0. This suggests that we do not have sufficient evidence to conclude that Beijing Mandarin is rated differently from Standard Mandarin for Status attributes.

The 95% credible interval for comparing ratings for Beijing Mandarin to Taiwanese Mandarin was between -0.44 and -0.20, which does not include 0. The Estimate for this comparison is -0.32. This suggests that we can conclude with 95% confidence that Taiwanese Mandarin is more likely to be rated lower than Beijing Mandarin for Status attributes. Taken together, we can conclude that Beijing Mandarin is rated higher than Taiwanese Mandarin and Singaporean Mandarin on Status attributes, but that Beijing Mandarin is not rated as significantly different from Standard Mandarin.

2.2.10 Solidarity Attributes
The Solidarity traits Honest, Humorous, Kind, and Trustworthy were aggregated into Solidarity traits. These were aggregated by taking the means of every rating for these four traits. The results of the ordinal regression model for Solidarity attributes are shown in Table 14. From Figure 10 above, the mean ratings for Standard Mandarin appeared to be slightly higher than the rest, and the mean ratings for Taiwanese Mandarin appeared to be slightly lower than the rest.

Predictor (for Solidarity attributes)	Estimate	Est.Error	l-95% CI	u-95%
Intercept	3.42	0.08	3.27	3.57
Comparison between Singaporean and Beijing accent, coded as accentsingapore	0.03	0.06	-0.09	0.15
Comparison between Standard and Beijing accent, accentstandard	0.14	0.06	0.02	0.26
Comparison between Taiwanese and Beijing accent, accenttaiwan	-0.16	0.06	-0.28	-0.04

Table 14: Regression model of Solidarity attributes.

The 95% credible interval for comparing ratings for Beijing Mandarin to Singaporean Mandarin was between -0.09 and -0.15, which does include 0. This suggests that we do not have sufficient evidence to conclude that Beijing Mandarin is rated differently from Singaporean Mandarin for Solidarity attributes.

The 95% credible interval for comparing ratings for Beijing Mandarin to Standard Mandarin was between 0.02 and 0.26, which does not include 0. The Estimate for this comparison is 0.14. This suggests that we can conclude with 95% confidence that Standard Mandarin is more likely to be rated higher than Beijing Mandarin for Solidarity attributes.

The 95% credible interval for comparing ratings for Beijing Mandarin to Taiwanese Mandarin was between -0.28 and -0.04, which does not include 0. The Estimate for this comparison is -0.16. This suggests that we can conclude with 95% confidence that Taiwanese Mandarin is more likely to be rated lower than Beijing Mandarin for Solidarity attributes. Taken together, we can conclude that Beijing Mandarin is rated higher than Taiwanese Mandarin and lower than Standard Mandarin on Solidarity attributes, but that Beijing Mandarin is not rated as significantly different from Singaporean Mandarin.

3 Discussion

This paper investigated the social attribute ratings of Standard Mandarin, Beijing Mandarin, Singaporean Mandarin, and Taiwanese Mandarin by Singaporean listeners. This question was motivated by the more broad issue of social meaning, and what listeners have access to certain social meanings, especially when a given language is spoken broadly by a wide variety of speakers (across many regions). This study found that not every attribute had significant differences between Beijing Mandarin and the other three accents, though there were more often differences between Beijing Mandarin and Singaporean/Taiwanese Mandarin, than there were differences between Beijing Mandarin and Standard Mandarin. The fact that Singaporean listeners treated Beijing Mandarin and Standard Mandarin similarly is interesting, and future studies with Chinese and Taiwanese listener populations may find greater differences in the treatment of these two accents, which would speak to what social meanings are significant to which listener populations.

The prediction following from previous literature that one or both of the Chinese accents (Standard/Beijing) would be rated as higher in Status attributes than both Taiwanese Mandarin and Singaporean Mandarin was mostly replicated. We found that Beijing Mandarin was significantly more likely to be rated as higher in every individual Status trait, as well as the aggregated Status straits, than Taiwanese Mandarin and Singaporean Mandarin. However, since Beijing Mandarin was the zero level/baseline, we could only compare the other accents against this one.

The prediction that Singaporean Mandarin would be rated as highest in Solidarity out of all the accents was not replicated. In fact, none of the Solidarity attributes individually showed a significant difference in rating between Singaporean Mandarin and Beijing Mandarin. The aggregate Solidarity traits did not show this difference either. However, the prediction that one or both of the Chinese accents (Standard/Beijing) would be rated as higher in Solidarity than Taiwanese Mandarin was partially found. Taiwanese Mandarin was rated as significantly less likely to be rated higher than Beijing Mandarin in 3 out of the 4 individual Solidarity attributes, as well as the aggregate Solidarity attributes.

This shows that the factors that are predicted to affect how high in Status or Solidarity a given accent is rated, e.g. education standards, media exposure, colloquial language and "localness" (Chong & Tan 2013; Cavallero 2018), have more nuanced and dynamic effects on social evaluations. Additionally, the present results lay out more questions for future directions of this research, such as how other listener populations may or may not distinguish between Standard Mandarin and Beijing Mandarin more than Singaporean listeners. This is not an unreasonable prediction given Zhang's (2017) detailed observations and analyses that the social meanings of Standard Mandarin and Beijing Mandarin are changing alongside social change in China; thus, if certain listener populations can access this social distinction, it may speak to their participation in given parts of social change.

4 References

Bokhorst‑Heng, W. (1999). "Singapore's Speak Mandarin Campaign: Language Ideological Debates and the Imagining of the Nation." In Language Ideological Debates, ed. by J Bloomaert, 235–265. New York: Mouton de Gruyter.

Bürkner, P.-C. (2017). brms: An R package for Bayesian multilevel models using Stan. Journal of Statistical Software, 80 (1), 1-28. doi: 10.18637/jss.v080.i01

Bürkner, P.-C., & Vuorre, M. (2019). Ordinal regression models in psychology: A tutorial. Advances in Methods and Practices in Psychological Science, 2 (1), 77-101. doi: 10.1177/2515245918823199

Cargile, A. C., Giles, H, Ryan, E. B., & Bradac, J. J. (1994). "Language attitudes as a social process: A conceptual model and new directions." Language & Communication 14(3): 211–236.

Cavallaro, F., Seilhamer, M. F., Ho, Y. Y., & Ng, B. C. (2018). Attitudes to Mandarin Chinese varieties in Singapore. Journal of Asian Pacific Communication, 28(2), 195‑225. doi:10.1075/japc.00010.cav

Chong, R. H.-H., & Yan, Y.-Y. (2013). Attitudes toward accents of Mandarin in Singapore. Chinese Language and Discourse 4(1) DOI: 10.1075/cld.4.1.04cho

de Leeuw, J. R., Gilbert, R. A., & Luchterhandt, B. (2023). jspsych: Enabling an open-source collaborative ecosystem of behavioral experiments. Journal of Open Source Software, 8 (85), 5351

Department of Statistics (2011). Census of population statistical release 1: Demographic characteristics, education, language and religion. Singapore: Department of Statistics, Ministry of Trade and Industry.

Dragojevic, M., Giles, H., & Watson, B. M. (2013). Language ideologies and language attitudes: A foundational framework. In H. Giles & B. M. Watson (Eds.) The social meanings of language, dialect and accent: International perspectives on speech styles (pp. 1-25). New York: Peter Lang.

Lindeløv, J. K. (2018). How to compute Bayes factors using lm, lmer, BayesFactor, brms, and JAGS/stan/pymc3.

Schmalz, X., Biurrun Manresa, J., & Zhang, L. (2021). What is a Bayes factor? Psychological Methods. Advance online publication. https://doi.org/10.1037/met0000421

Vasishth, S., Nicenboim, B., Beckman, M. E., Li, F., & Kong, E. J. (2018). Bayesian data analysis in the phonetic sciences: A tutorial introduction. Journal of phonetics, 71, 147–161. https://doi.org/10.1016/j.wocn.2018.07.008

Wang, H. D. (2002). Xīnjiāpō huáyǔ tèyǒu cíyǔ《新加坡华语特有词语》[An investigation of the Singapore Mandarin Vocabulary], In C. H. Chow (Ed.), xīnjiāpō huáyǔ cíhuì yǔ yǔfǎ《新加坡华语词汇与语法》[Vocabulary and grammar in Singapore Mandarin] (pp. 25–73). Singapore: Lingzi Media.

Zhang, Q. (2008). Rhotacization and the 'Beijing Smooth Operator': The social meaning of a linguistic variable. Journal of Sociolinguistics 12: 201-22.

Zhang, Q. (2017). Language and social change in China: Undoing commonness through Cosmopolitan Mandarin. Taylor and Francis. https://doi.org/10.4324/9781315886251