



Universiteit  
Leiden  
The Netherlands

## **The sociolinguistics of rhotacization in the Beijing speech community** Hu, H.

### **Citation**

Hu, H. (2022, September 21). *The sociolinguistics of rhotacization in the Beijing speech community*. LOT dissertation series. LOT, Amsterdam. Retrieved from <https://hdl.handle.net/1887/3464385>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3464385>

**Note:** To cite this publication please use the final published version (if applicable).

## **Chapter 3      Methodology**

This chapter presents the general methods and research design used for the collection, processing, and analysis of the data in this dissertation. The details concerning data manipulation and the statistical methods employed in different experiments will be introduced in the corresponding chapters.

Section 3.1 introduces the social variables of the participants examined and the actual number of participants in this study. Section 3.2 discusses the theoretical methods and practical considerations in participant recruitment for the fieldwork. The details of the recording procedures, including devices and equipment utilized in fieldwork, sample rate, and so forth, will be presented in Section 3.3. Section 3.4 introduces the two methods of sociolinguistic data collection in this study and discusses circumstances that may influence the data quality and objectivity in data collection.

### 3.1 Participants

#### 3.1.1 Age

One of the primary tasks of variationist sociolinguistics is to understand language change that occurs over time. Apparently, older people do not talk the same as younger people. Therefore, “establishing that a pattern of variation represents a change in progress typically requires the consideration of speakers of different generations” (Milroy & Gordon, 2003).

Variationist sociolinguists, thus, utilize two important methods of analysis to study language change in progress: (1) the real time method, conducting a longitudinal study to collect the speech of the same people in different life stages, by revisiting the same speech community in different periods; (2) the apparent time method, sampling speakers of different generations at a single point in time, and simulating and modeling real time change by using synchronic data (Meyerhoff, 2011). Age differences are believed to be able to reflect historical stages of language change in progress (Tagliamonte, 2012). The apparent time method is a more time-efficient and practical approach, because the diachronic data are usually not available to researchers and because time and money is usually insufficient to construct a real time corpus.

There are a few of studies about the effects of age on rhotacization in Mandarin. In the study of the speech dispersion of rhotacization in Beijing Mandarin, age is found to be an important factor (T. Lin & Shen, 1995; D. Sun, 1991). L. Wang’s (2014) study on the use of rhotacization by people of different generations finds that older people have more knowledge of the lexical and emotional connotations of rhotacization than younger people and also use more rhotacization in their speech, while some young people tend to have little knowledge of these aspects and thus use less rhotacization in spoken Mandarin.

In this study, the effects of generational change will be measured among three age groups, namely the Young (18 up to and including 27 years old), the Middle (38 up to and including 47 years old), and the Old (58 years old and older)

of both native and non-native speakers of Beijing Mandarin in the speech community of Beijing.

### 3.1.2 Gender

Gender has been proven to be an important social variable in sociolinguistic research. It has been found to have an effect on the preference of variation in many variationist studies.

The findings of stratification studies (Labov, 1963; Trudgill, 1972) on this variable suggest that there is generally a link between women and standard language use. Specifically, women tend to use more “prestige” or high-status language features, while males tend to use more vernacular language features. A number of explanations have been proposed to explain gender effects on language, including biology, cultural patterns, covert prestige, the social position of women, and societal norms and practices.

Gender differences have been proven to affect language use in different linguistic phases, an example of which is their effects on rhotacization in many languages. However, the opinions on the relationship of gender effects and language varieties are quite varied and are to some extent underestimated in sociolinguistic studies on Mandarin rhotacization. Lin & Shen’s (1995) study finds that there is no gender difference regarding the convergence of rhotacization among Beijing local people. Jing’s (2005) sociolinguistic study on variation and change in rhotacization doesn’t consider gender as a variable.

In the present study, we take gender as a social variable and try to determine if gender affects the use of rhotacization in the speech community of Beijing.

### 3.1.3 Dialect background

In this study, region of origin also plays a role, in view of language variation in China, especially in as far as it concerns rhotacization. Generally speaking, varieties of Chinese spoken in northern China tend to be rhotic, while southern varieties tend to be non-rhotic (B. Huang & Liao, 2017). The participants to our study belong to three different dialect background groups, namely, “Beijing native speakers,” “Rhotic speakers” and “Non-rhotic speakers”. The “Beijing native speakers” were born and raised in Beijing and still live there; crucially, their parents were also born and raised in Beijing. “Rhotic speakers” and “Non-rhotic speakers” were not born and raised in Beijing. They moved there from elsewhere in China, at the time of the fieldwork, they were working or studying in Beijing. Rhotic speakers come from areas in China where rhotacization is used in the regional dialect. These include speakers whose native dialect is, for example, Tianjin Mandarin, Harbin Mandarin (in Heilongjiang), or Shijiazhuang Mandarin (in Hebei). Non-rhotic speakers are originally from non-rhotic areas,

that is, areas in which rhotacization is not part of the regional varieties of Chinese. They hail from areas such as Jiujiang (in Jiangxi) (where they speak a variety of Mandarin), Guangzhou in Guangdong (Cantonese), and Fuzhou in Fujian (Northern Min).

It should be noted that in previous studies of Beijing Mandarin, Beijing native speakers were divided into two groups, “Old Beijingers” and “New Beijingers,” based on differences in their family language backgrounds (Y. Hu, 2011). Hu investigated social influences on the use of rhotacization in the speech of Beijingers, discovering that family language background is an important factor. In his study, Old Beijingers are native speakers who were born and raised in the urban area of the city. They are considered to be “real” Beijing natives, and the Mandarin they speak is taken as the “proper” native Beijing Mandarin. New Beijingers are Beijingers who were born and raised in Beijing, but whose parents were not born and raised there. In our study, Beijing native speakers are defined as Hu’s Old Beijingers, except that we don’t specify that they have to come from the urban center of Beijing. Due to the urbanization in past twenty years in Beijing, the urban area has expanded, and new city settlements continue to be developed. Local Beijingers have often moved from their original urban areas to newly developed areas in the city, where they were joined by native speakers and people who arrived in the city more recently.

#### 3.1.4 Ethnicity

The demographic census in 2018 shows that the permanent resident population of Beijing is about 19.6 million. Among them, 18.8 million are Han people, accounting for about 96% of the total population. In this study, we will not consider ethnicity as a social variable, due to the dominance of the Han ethnicity and Mandarin in Beijing.

#### 3.1.5 Summary<sup>9</sup>

The intended number of participants was four in each Dialect–Gender–Age subgroup, which means there should be 72 participants in total. However, it was

---

<sup>9</sup> Other variables, like ethnicity, are less relevant, as mentioned in Section 3.1.4. Level of education should be an important variable as well. However, the problem is that there is an overlap among the variables age, social status, and level of education. Cao (1991) observes that in China, there is a relation between level of education and age. Younger people are usually more educated than older people. Furthermore, those who have high social status are always highly educated. For this reason, level of education is not considered as a separate variable here.

not equally easy to get access to the intended number of participants for each group.

The speech data used in this study were collected during two fieldwork trips to Beijing, in October and November 2015 and September and October 2017. During these periods, 121 participants were recorded in total. However, not all the recordings could be used in data processing and analysis, due to practical and technical reasons. For example, in some cases, social factors of the participants, such as age and dialect background, turned out not correspond to the categories we needed in this project. In other cases, the recording quality of some two-participant conversations (“pair talking,” explained in Section 3.4.1) was not good enough to conduct an acoustic analysis, due to factors such as a noisy recording environment and interruptions. As a result, in the end the recordings of 76 participants were subjected to analysis in this study. Of these 76 participants, 31 are Beijing native speakers and 45 are migrants. They are all Han Chinese. On the basis of age and dialect background, the actual number of participants in each group is given in Table 3.1.

**Table 3.1** Actual number of participants.

<i>Gender</i>	<i>Age</i>	<i>Beijing Mandarin</i>	<i>Rhotic dialect</i>	<i>Non-rhotic dialect</i>
Male	Young	8	4	4
	Middle	4	4	4
	Old	4	4	2
Female	Young	7	4	5
	Middle	4	4	3
	Old	4	2	5
Total		31	22	23

Based on the aspects of the sociological information of these participants listed above, three independent variables were set up for the testing of social conditioning effects on the use of rhotacization. In what follows, predictor names are presented in uppercase, and levels of categorical predictors are in italics. In our analyses, the social predictors include:

- (1) AGE (Young, Middle, and Old)
- (2) GENDER (Male and Female)
- (3) DIALECT BACKGROUND (Beijing, Rhotic, and Non-rhotic)

### 3.2 Participant recruitment

Three methods of participant recruitment were involved in this study: (1) random searching by the author, (2) using her social networks in the community,

and (3) the top-down method. Most participants were approached through the latter two methods.

When the author made her way into the community, she tried to simply apply a random sampling methodology to this study. “Participant recruitment leaflets” were posted on the bulletin boards on the campus of several universities, as well as in several residential areas in Beijing. The leaflets outlined the requirements of intended participants, payment, the general purpose of this research, and so forth. The leaflets were worded using general terms understandable to lay readers. It was stated that the object of this study is to observe the everyday languages spoken by residents of Beijing, while the real intent of studying rhotacization was not mentioned in the leaflets, lest people become overly self-conscious about pronouncing rhotacization in recordings.

However, this random sampling method turned out to be an inefficient way to generate contacts, due to the constraints of low trust between strangers in Chinese society. University students did contact the author, from which some people who met the experimental criteria were randomly chosen. However, although several people in residential areas got in touch with the author via the contact information on leaflets, they showed strong distrust towards the survey. The author constantly found herself having to explain what she was doing there in the community. After further explaining the purpose and motives of this study, only a couple of people were willing to participate.

Later, a “snowball sample” or “friend of a friend,” the well-known sociolinguistic method, proved to be an excellent way to find participants. The “friend of a friend” is “a community member with whom the researcher shares a common friend or acquaintance” (Bayley & Lucas, 2007; Meyerhoff et al., 2011; Milroy, 1980). This method enables fieldworkers to obtain more naturalistic and spontaneous speech than if they are viewed chiefly as researchers (Bayley & Lucas, 2007). Moreover, participants have higher motivation because they are helping their friends. The author thus turned to her personal networks that have ties to the community, such as previous classmates, friends, relatives, and family. These people were asked to contact their friends and acquaintances who fit into the required categories.

### 3.3 Data collection

The data collection was conducted during the two fieldwork visits in Beijing in 2015 and 2017. Both of the visits were planned for the autumn, the most comfortable season of the year (in the summer, temperatures can reach 40 degrees or more, while in the winter minus 10 is not unusual). In the autumn, people do not need air conditioners to keep cool or heaters to stay warm. In this way, we avoided the noise from air conditioners or heaters. Speakers also tended to be more patient and cooperative in the experiments under these conditions.

Speakers were interviewed in their own neighborhood. Their speech was recorded in a quiet room, which they were very familiar with (for more details,

see below), and the doors and windows were closed during the recording to reduce background noise.

The topics for participants to discuss were presented on two Microsoft PowerPoint slides on a 13.3-inch laptop screen. The font size was 50 points, easier for older participants to read when they sit about one meter away from the screen. The display laptop was a MacBook Air, whose weight was 1.35 kg with battery. It had up to 12 hours of battery life, which was more than enough for one-day use of PowerPoint displays. In view of the fanless design of the MacBook Air, there would be no loud fan noise as may have been the case with fan-equipped laptops.

The speech of the participants was recorded on digital audio tape using a portable TASCAM DR-07 recorder. The recordings were sampled at 48 kHz (24 bits).

The author of the present study was the interviewer, and she had two fieldwork assistants who were both native speakers of Mandarin. They were all graduate students of phonetics and had experience in dialectological fieldwork.

### **3.4 Research design**

#### **3.4.1 Pair talking**

The data collection methods of sociolinguistic fieldwork, such as group recordings, “target where one variant shows up by whom, what it is and how it changes” (Schilling, 2013). Furthermore, recording speakers in small group interactions is a way of obtaining more naturalistic, spontaneous speech and can address the observer's paradox. Some of the awkwardness associated with overt recording can be reduced, because the speakers are allowed to talk about daily topics and self-select who speaks when. Speakers may also feel more relaxed with familiar faces (Meyerhoff et al., 2011; Schilling, 2013). Pair talking is a form of such group recording for data collection in sociolinguistic fieldwork. Two speakers who are familiar with each other conduct a conversation on familiar topics in a familiar place (Schilling, 2013).

Data collection in this study was intended to obtain naturally occurring rhotacization from both native speakers and migrants in the speech community of Beijing. In the data collection, every time one person willing to participate was recruited, he/she was asked to self-select a partner to attend the pair talking, usually a friend, a member of their family, a roommate, a colleague, or a neighbor. They could also freely choose a quiet place that they were familiar with. Places such as one participant's study room at home, an office, or a conference room were usually chosen. The topics designed were all related to their everyday life in Beijing or memories of their childhood or hometown, such as “your experiences



## 34 THE SOCIOLINGUISTICS OF RHOTACIZATION IN THE BEIJING SPEECH COMMUNITY

with or opinions on the air and traffic in Beijing”<sup>10</sup> or “your experience with full marks/low marks on an exam when you were/your (grand)child was a student”. The complete list of given topics can be found in Appendix B.

### 3.4.2 Sociolinguistic survey

The sociolinguistic survey complements the pair talking and “allows for greater breadth of coverage” (Schilling, 2013). It can help to target features that may or may not be present in conversational settings and can also help the researcher glean detailed information on respondents’ characteristics, their identification of variants, varieties, and their attitudes toward the speakers who use them. The information elicited from sociolinguistic surveys can help to test if they are correlative with language variation. The information on perceptions and attitudes can be used to complement and explain production-based studies.

### 3.4.3 Respondents

The sociolinguistic survey was conducted immediately after each pair talk. The participants in the pair talks were, thus, the respondents of this sociolinguistic survey. Right after each pair talk, the author of this study interviewed one respondent while her fieldwork assistant interviewed the other respondent at the same time. However, instead of sitting face-to-face or next to each other as in pair talking (as introduced in Section 3.4.1), the two respondents were separated and interviewed in two different rooms. In this way, the respondent could not hear and thus was not influenced by the other parallel interview. All interviews were recorded with the respondents’ permission.

## 3.5 Related issues

### *Overcoming the observer’s paradox*

One important goal in collecting sociolinguistic data is “to observe how people speak when they are not being observed” (Labov, 1972b), because only when speakers are unobserved can one collect the natural and spontaneous speech that they would use in daily life. However, the presence of a fieldworker, a

---

<sup>10</sup> Beijing is notorious for its caustic air, caused by its rapid industrialization. People in Beijing have long suffered from the traffic congestion and hazardous air. Everyone in Beijing, rich and poor alike, both young and old, cannot ignore the effects of traffic and air quality. Therefore, this is a daily life topic that everyone can talk and complain about.

recording device, or the task in an interview can trigger the observer's paradox (Meyerhoff et al., 2011). The problem, faced by sociolinguists in particular, is that, in observing or interviewing people to find out about their habits of speech, investigators will, by their own presence and participation, tend to influence the forms that are used. For these reasons, sociolinguists utilize several methods to mitigate the effects of the observer's paradox in data collection. These include modifying the number of participants in an interview, topics discussed, and the task.

As mentioned in Section 3.4.1, group recordings, which record people in group interactions, is believed to be an effective method for mitigating the observer's paradox. Speakers may feel more relaxed with familiar interlocutors in a place familiar to them, leading them to produce naturalistic and spontaneous speech. The pair talking is one type of such group interaction, in which two people who are familiar with each other to talk about familiar topics. Moreover, such topics can often evoke participants' strong emotions connected with their past experience. This diverts participants' attention away from their speech and helps trigger their vernacular (Labov, 1972b).

#### *Speech convergence between interlocutors*

While overcoming the observer's paradox, we are also aware that there is the possibility of speech convergence and divergence between interlocutors in a conversation, which can affect participants' speech. Accommodation theory, proposed by Giles (1973) and Giles & Powesland (1975), suggests that speakers tend to adjust their speech toward or away from that of their interlocutors, although, conversely, they may also keep their speech uninfluenced. These phenomena are respectively called convergence, divergence, and maintenance. Whether the speaker accentuates the similarities or differences of his/her speech depends on the "psychological distance" they want to keep with their addressee and their attitudes towards him/her.

Speech convergence is found to be more frequent than divergence and maintenance, a phenomenon that is also known by the terms phonetic imitation, alignment, and entrainment.

Many studies have measured the influences of acoustic-phonetic factors on speech convergence. Data were collected in speech shadowing tasks and in conversational interaction. The measures of the AXB perceptual-similarity task and acoustic analysis are employed. In the perceptual task, listeners hear and then decide whether the pre-(A) or post-(B) exposure utterance sounds more similar to the middle utterance (X) in pronunciation (Babel et al., 2013, 2014; Goldinger & Azuma, 2004; Namy et al., 2002; Shockley et al., 2004). Acoustic attributes, such as vowel selection (Pardo, Gibbons, Suppes, & Krauss, 2012), duration (Gentilucci & Bernardis, 2007; Pardo et al., 2013), VOT (voice onset time) (Sanchez et al., 2010; Shockley et al., 2004; Yu et al., 2013), fundamental frequency (Babel, 2012; Garnier et al., 2013; Mantell & Pfordresher, 2013;

Postma-Nilsenová & Postma, 2013), F1 and F2 (Babel, 2010, 2012; Gentilucci & Bernardis, 2007; Pardo et al., 2013; Vallabha & Tuller, 2004; Walker & Campbell-Kibler, 2015), and the F1 × F2 vowel space, of converged utterances are also examined to see if there is a relation between perception and production in speech convergence. However, convergence was reported to be “subtle, variable, and inconsistent” (Pardo et al., 2017) due to the complexities of convergence itself and different methodologies in those studies.

Meanwhile, some studies also investigated the effects of factors related to the talkers themselves, such as gender, relationships between talkers, and attitudes towards model talkers (Pardo et al., 2012). Some found that the role of a talker affects the degree of phonetic convergence or convergence “moderately: relates to the relationship of the talkers (Pardo, 2006; Pardo et al., 2012).

Effects of other factors on phonetic convergence were also examined, including word frequency (Pardo, 2013; Pardo et al., 2017) and talkers’ experience with the words being examined in the experiments, as well as voice types (Babel et al., 2014), social preferences, and liking (Babel, 2012).

Above all, many studies have examined phonetic convergence from different perspectives. It is agreed that the phenomenon of convergence is perceived to exist between interlocutors and that it is observed in a laboratory setting. However, results of those studies on the various dimensions of convergence are still inconsistent. Issues such as the degree and scope of convergence, its acoustic correspondence and attributes, the phonemes subjected, and talkers’ effects are not commonly agreed upon.

We are aware of the phenomenon of rapid phonetic convergence in a conversational setting. However, in this present study, we would still utilize the speech data from pair talks, even if some participants do not share the same linguistic background. First, as mentioned above, those studies mentioned above concluded that the relation between perceptual and acoustic measures on convergence is holistic and variable. Furthermore, laboratory experiments on two previously unacquainted participants are different from a real social conversation. As mentioned above, in the present study, the pair-talk design is based on the principle of “people who know each other talking about familiar topics in a familiar place.” When they recalled their childhood memories and interesting life experiences, the participants were quite relaxed and could converse with each other in a positive and natural manner. Their attention was diverted to recalling their story and telling it. Accordingly, this design could trigger the natural and spontaneous speech of participants, which is believed to be the ideal speech data in a socio-phonetic study.

When recruiting participants and before doing the pair talking, participants were informed to talk in the way they do in their daily life.

As mentioned above, according to accommodation theory, both convergence and divergence can occur in an interactional conversation. We should not take for granted that convergence will necessarily occur. Furthermore, the mechanism of convergence and divergence in the context of Chinese culture is quite complicated. Due to rapid social development in Beijing, people differ greatly regarding to what they think highly of.