Department of Education, University of Oxford

Factors impacting learners' proficiency in the Heritage Language Shanghainese

Chuchu Chen

Jesus college

Word count: 15423

Dissertation submitted to the University of Oxford in partial fulfilment of the requirements for the degree of MSc in Applied Linguistics and Second Language Acquisition

August 2022

Acknowledgement

I would like to thank my supervisor Elizabeth Wonnacott, who has been exceptionally professional, supportive, and insightful. Thank her for her (extremely!) quick replies to emails, her patient readings and inspirational, detailed comments on my essays, and most importantly her deep devotion to academia which makes her lectures all informative, thought-provoking, and intriguing.

Thank you to all my ALSLA people, who have composed a wonderful cohort. I feel so well supported and comfortable being myself around you all. This means a lot to me especially because it is my first time studying abroad in an English-speaking country and I wasn't overly familiar with the cultural environment here in the first place. Thank you so much for your kindness and impressive academic performance.

To Lingyi, Liuyi, and Xuechun, thank you for being my best friends in the cohort and my profession-level amateur academic tutors. Thank you Liuyi for being the best and funniest chef ever, Lingyi the most patient and supportive teacher-to-be, and Xuechun the most mentally stable, resourceful, and reliable linguist. We had wonderful times together and I am most certain that our paths will cross again.

To my amazing Jesus people, Daniella, Julia, Ruihan, and Wei, I feel so lucky to have met you (almost) the first days here. Thank you for being so smart, inspirational, and lovely. Most importantly, thank you for always being around and making me feel at home. Thank you Lily for being an angel in a human form.

To Lady Margret Hall, thank you for being a warm, comfy, liberal, lovely harbour during the term-times, for being my (self-entitled) honorary second college, and for providing the undisputedly best food at Oxford. I will strongly confront anyone who suggests otherwise.

To my family, thank you for always being my most stable and dependable support.

This year at Oxford hasn't been easy for me, during which I constantly faced new challenges, stepped out of my comfort zone, and encountered turbulences and uncertainties in life. I believe that the way to counter difficulties and to secure safe feelings is to love the people around me. That is also why I feel utterly lucky and blessed to have met all the amazing people. Thank you all so much for this meaningful and memorable year which will always be deeply engraved in my life.

Abstract

This study aims to better understand the current status of Shanghainese, a Chinese Heritage Language (CHL), and to help its preservation. It explores learners' usage patterns and their proficiency level in Shanghainese. It also investigates what factors impact the learners' Shanghainese competence by looking into the relationship between learners' HL proficiency and their attitudes and beliefs towards the dialect, engagement in dialect-related cultural activities, demographic backgrounds, as well as their home and community language environment. Most previous studies did not specify a regional dialect when researching CHLs, only employed relatively simple statistical analysis, relied on participants' self-report and focused on a very narrow range of variables that could impact the CHL proficiency. This study designed a questionnaire and a Shanghainese proficiency test to collect speaker information among 187 Shanghainese teenagers between Year 10 (16 years old) and Year 12 (18 years old) from five different high schools in Shanghai. The results showed that the participants generally preferred speaking Mandarin Chinese to Shanghainese in varied scenarios. Their self-perceived and tested Shanghainese proficiency were significantly positively correlated. The participants with a more positive attitude towards Shanghainese, more frequent attendance to dialect-related cultural activities, a more local demographic background, as well as a more dialect-oriented home language environment would have a higher competence level in the dialect Shanghainese. Implications for future research are also provided in this study.

Table of Contents

List of acronyms and abbreviations

AAS	Attitude Analog Scale
CHL	Chinese Heritage Language
HL	Heritage Language
L1	First language
L2	Second language
RQ	Research question
SMLC	Shanghai Municipal Language Committee

List of figures

List of tables

1. Introduction

1.1. Research background

In recent years, the attention paid to Heritage Languages (HL) has been on the rise. Scholars have long suggested that the development of HL is conducive to the learners and their cultural groups in that it helps speakers build their identity (e.g., Cho et al., 1997; Wong & Xiao, 2010; He, 2010; Leeman, 2005; Carreira, 2004), strengthens the bond between speakers and their ethnic community (Cho, 2000), as well as encourages linguistic and cultural diversity in the whole world (Jiang, 2021). Large numbers of schools and communities around the globe have correspondingly set up learning programmes to offer more solid support for HL learners.

Broadly speaking, there has been a shift in the definition of an *HL* as well as what counts as *HL speakers*. The term "heritage speaker" was first coined in Canada in 1977 (Cummins, 2005) and started to gain attention in the United States in the 1990s (Montrul, 2010). It was, however, originally applied exclusively in the context of a society where English is the dominant language. As Valdes (2001) defined it, heritage learners referred to speakers whose home language is not English but are able to speak English, and who understand their home language at least to some degree. Later on, the scope of the term "heritage learners" has been expanded to the context where English is not necessarily the mainstream language. That is, HL learners are those who grow up in families that speak a linguistically minoritised language and who are exposed to the majority language at the same time (Van Deusen-Scholl, 2003; Valdes, 2005; Montrul, 2010). Fishman (2001, 2006) pointed out that the range of HLs in the United States includes *immigrant languages* (e.g., Korean, Chinese, Indonesian), *colonial languages* (e.g., German, French), as well as *indigenous languages* (e.g., Zuni, Cherokee). As indicated by Jiang (2021), the scope of this categorisation could be expanded from the U.S. to the globe. Accordingly, the main focus of this essay, i.e., Shanghainese, belongs to the last category, i.e., *indigenous language*, in the context of Shanghai.

Studies that specifically targeted at Chinese Heritage Language (CHL) learners usually set their focus overseas, especially in the English-speaking countries, where the Chinese languages are immigrant languages. Among these studies, the focus on CHL has also experienced a shift from viewing them as a whole, i.e., *Chinese languages*, to breaking it down to specific regional dialects spoken in China (e.g., Cantonese, Shanghainese, Suzhou dialect) (Duff et al., 2017; Wong & Xiao, 2010). He (2006) was the first to explicitly investigate the identity construction of CHL learners and she referred to CHL speakers broadly as people who are bilingual in Chinese and English, and whose home language is Chinese (incl. Mandarin and dialects). She proposed an identity-based model of CHL development and indicated the importance of focusing on CHL development, which will be further discussed in the literature review chapter. In light of this, Wong and Xiao (2010) took a step further and examined the identity construction of CHL learners from different dialect backgrounds, and indicated that Mandarin speakers and dialect speakers are essentially different when constructing their identity. This, too, will be discussed in the next chapter as part of the literature review.

In terms of the studies carried out inside of China, where people's HLs normally are specific dialects, studies often directly look at dialects as HLs under Mandarin dominance. For example, Xu and Zhou (2016) investigated the dialect use and language attitudes towards the dialect Cantonese in Guangzhou, China. They found that although the increased use of Mandarin does not reduce the proportion of Cantonese users, teenagers and adolescents in Guangzhou preferred Cantonese over Mandarin in terms of their language attitude. Zhao (2018) examined the language variations and social identity mediated in the use of the Beijing dialect. She found that the use of the Beijing dialect conveyed a sense of localness and that speaker constructed their identity via daily dialect use. However, this specific focus on dialects has not yet attracted much attention either overseas or inside of China, as is especially the case for Shanghainese. The dominance of Mandarin over other Chinese dialects often causes difficulty in the preservation of the HL. This will be explicitly discussed in the following section.

1.2. The language power relationship in China

Within the broad scale of China, there are a large number of heritage languages. Moreover, many of those language varieties are mutually unintelligible (He, 2006; Li, 2006; Jiang & Dewaele, 2019). Speakers from different HL backgrounds may therefore construct different identities, have differentiated language experiences, and represent different cultures and traditions. This means that only referring to Mandarin Chinese when talking about Chinese HL, which is an extremely common phenomenon, is considerably problematic. On the individual level, studies have also shown that speakers of different HL groups take great pride in their own non-Mandarin linguistic variety (Wong & Xiao, 2010). These all point to the fact that local dialects should be explicitly differentiated from Mandarin Chinese.

Broadly speaking, the non-Mandarin language varieties in China apart from minority languages (e.g., Uygur, Tibetan) could be divided into six major Han (the largest ethnic group in China) dialect groups: *Min*, *Wu*, *Yue*, *Gan*, *Hakka* (*Kejia*), and *Xiang* (see. e.g., Chappell, 1992; Ramsey,1989; Wong & Xiao, 2010). Between each of the two groups there stand huge differences in terms of semantics, syntax, and phonetics, especially between languages varieties of the Northern (e.g., Mandarin) and Southern (e.g., Min, Wu, Yue) (Wang, 1973). Within each group, the HLs could be further divided into different variations. Shanghainese, for example, belongs to the Wu group while Cantonese, a linguistic variety that also has a large population of speakers, is part of the Yue group.

The dominance of Mandarin and the fact that other CHL varieties are often left neglected reveals the power relationship behind these languages. Since the 1950s, Mandarin has been established as the standard language nationwide by relevant policies (Chen, 1999; Guo, 2004), which has led to significant changes in the linguistics ecology in China (Xu et al., 2021). Law of the People's Republic of China on the Standard Spoken and Written Chinese Language (Standing Committee of the National People's Congress, 2000) was enacted in 2001 which dictated that Mandarin (or *Putonghua*) be used as the basic language in education and teaching in schools and other educational institutions. It also prescribed that Mandarin be used as the basic broadcasting language in broadcasts and TV.

Just as English could act as a "predator language" in some countries and endanger other languages (Hornberger, 1997), Mandarin also overshadows other language varieties in China with its prestigious and significant status (Wu & Leung, 2020). This is in line with the phenomenon that nation-states always hold the one-nation one-language policy and have the tendency to ignore the linguistic diversity within the nation (Bauman & Briggs, 2003; Blackledge et al., 2008; Wu & Leung, 2020). However, many scholars have noticed the linguistic hegemony in Chinese languages and have called for a more suitable multilingual approach in communities and classrooms to support all speakers from various dialect backgrounds (e.g., Duff et al., 2017; Lam, 2005; Wong & Xiao, 2010). In the section to follow, I will specifically talk about efforts on HL preservation in the context of Shanghai.

1.3. The relationship between Shanghainese and Mandarin Chinese

As is mentioned above, Shanghainese and Mandarin Chinese belong to different dialect groups in the Chinese language family. While Mandarin Chinese stands as its own group, Shanghainese belongs to the Northern *Wu* group of Chinese dialects. Since they differ to a great extent in their syntax, phonetics, and semantics, they are usually regarded as being mutually incomprehensible. For example, Shanghainese has five tones, while Mandarin Chinese only contains four. In comparison to Mandarin, Shanghainese also has significantly more consonants and vowels. Since Shanghainese has no standard written form, it is rarely used in writing. Although it is possible to transliterate Shanghainese into the Mandarin characters in the written form, the transliteration would in most cases make no sense due to the lack of their lexical similarity.

1.4. The HL preservation movement in Shanghai

Although there has been increased attention paid to the protection of endangered languages (e.g., Australian indigenous languages, American indigenous languages) worldwide, the preservation of regional dialects (e.g., Shanghainese, Suzhou dialect) has been largely left ignored (Gao, 2015; Schilling-Estes & Wolfram, 1999; Shen, 2016;

Tulloch, 2006;). This is also the case in Shanghai, one of the largest cities in China. With approximately 14 million speakers, Shanghainese is the largest variant under the *Wu* Chinese dialect group.

In Shanghai, the issue of dialect preservation has triggered heated social discussions. A lot of attention has been paid to Shanghainese, or *Hu Yu*, which is the dialect of Shanghai. In kindergartens and schools in Shanghai, the use of almost all the regional dialects, including Shanghainese, has been strictly prohibited to promote Mandarin Chinese (Xia & Shen, 2019). Under the assimilation effect of Mandarin, Shanghainese has experienced rapid changes and lost a large proportion of its speaker population, especially among the younger generations. The ethnographic report of Zhou (2012) showed that over 40 % of students in primary and middle schools in Shanghai were not able to speak Shanghainese well. In this light, the "*Save Shanghai Dialect*" movement was launched in 2012 and is still in progress (Shen, 2016). This movement calls for efforts to save the dialect both in academic and social contexts. On December 15th of 2011, 82 Shanghai-based scholars, including *Qian Nairong* (linguist and head of the Chinese Department at Shanghai University) and *You Rujie* (president of the Association of Shanghai Linguistics), signed a joint initiative requesting the scientific preservation of the Shanghai dialect (Shi, 2012). In this initiative, it was suggested that students in kindergarten, primary and middle schools in Shanghai should be free to speak Shanghainese as well as other regional dialects outside of the classroom. The teachers and parents should also encourage and lead the students to engage themselves in speaking Shanghainese.

In light of this, Shen (2016) interviewed 10 stakeholders and analysed their perceptions and experiences of the Shanghainese preservation movement. He found that although several relevant activities have taken place to promote the status of Shanghainese, more effort is urgently needed for it to exert an actual influence. This study, along with other current movements that aim to preserve and promote Shanghainese, will be discussed in detail in the literature review chapter to follow.

Based on what is discussed above, it is essential to investigate the language usage patterns of Shanghainese and speakers' proficiency thereof to critically evaluate the dialect preservation movement in Shanghai and to further facilitate this movement. Yu

and Yang (2016) investigated the linguistic pattern and dialect competence of local youths in Shanghai over the past 15 years. This study, too, found that Shanghainese is in a dire situation and is losing its speaker population drastically.

Given that the HL situation in Shanghai is severe and that scarce research has been done on its speakers' usage patterns, language proficiency, and the impact factors thereof, this dissertation aims to fill in the existing gap in the literature by looking at the following research questions:

RQ1. How can learners' Shanghainese HL usage patterns be characterised?

RQ2. How strongly correlated are learners' self-reported and tested HL proficiency?

RQ3. What factors are related to learners' proficiency level in Shanghainese?

1.5. Structure of the dissertation

This dissertation comprises 6 chapters. Chapter 2 reviews pertinent literature that investigates influence factors of HL proficiency. Chapter 3 outlines the methodology of this study. Chapter 4 presents the results of the data collected. Chapter 5 discusses the findings from the data and provides implications. Chapter 6 concludes the study.

2. Literature Review

In this chapter, I first introduce two studies that serve as the rationale for researching CHL development and proficiency. I then specifically evaluate several studies on the factors that influence HL proficiency, including support from family and community, HL schools and programmes, as well as learners' attitudes and motivations. I start with studies in the broader picture of HL research worldwide and then shift to the context of CHL. After that, I focus on relevant studies carried out specifically in the Shanghai context and point out the existing gap in the current literature.

2.1. Rationale for research into CHL development and proficiency

He (2006) was the first to explicitly look at the identity construction process of CHL speakers. Her study proposed an identity-based model for CHL development. This could serve as the rationale for the studies to come that focus on how social practises influence CHL speakers' identity and language proficiency. He (2006) drew on data from a couple of case studies where all the subjects were bilingual speakers of HL Chinese (Mandarin or dialects) and English. The participants all attended HL programmes on Mandarin Chinese in the United States. In light of the fact that CHL has received little attention in its theory building, He (2006) proposed a model of CHL development based on theories in *Language Socialisation* (which focuses on the process of becoming a culturally competent member of a social group through language use), *Second Language Acquisition* (to be more specific, the *Acculturation Model*, which states learners' acquisition of the second language is determined by their extent of acculturation to the target language group (Schumann, 1986) and the *Accommodation Theory* which considers the level of learners' motivation reflects how they define themselves in the ethnic terms (Giles & Byrne, 1982), both theories associate successful language acquisition with the learner's social and psychological adaptation to the target language community), and *Conversation Analysis* (which states that speakers' identity exists essentially in actions and language use). By analysing the speech data of participants talking about their HL learning and identities using the aforementioned theories, He (2006) pointed out that CHL development is grounded in speakers' engagement in social practice and their adaptation to the social activities and identities of the social communities they inhabit. That is to say, she theorised that CHL speakers' identity is dynamically connected with *social practise* (e.g., classroom activities, communication with family and peers, perceiving the language image in the environment, etc.). The article acknowledges the difference between Chinese dialects and Mandarin and pointed out that many of them are mutually incomprehensible. However, it focuses on CHL learners' identity construction only mainly in the context of Mandarin Chinese without further discussion concerning specific regional dialects. Though the model construction is overall logically well supported by the evidence analysed in the study, it remains unclear whether this theory could equally apply to

Mandarin and dialect speakers. As will be discussed below, the identity construction of speakers of Mandarin and other dialects are, in fact, essentially different. Thus, this study would be more valid if speakers from specific CHL backgrounds other than Mandarin could be investigated separately.

Wong and Xiao (2010) took a step further and investigated the identity construction of CHL learners from different dialect backgrounds. They interviewed 64 Chinese language students learning Mandarin at universities in the U.S. among which 37 were born in the U.S. and 27 immigrated to the U.S. with their parents at a young age (mostly between 0-10 yrs). 19 out of the 64 participants were Mandarin speakers while the others had a diverse CHL background (e.g., Cantonese, Jiangxi dialect, Shanghainese, Chaozhou dialect, Taiwanese). In the interview data, three conceptual categories emerged: *imagined community* (most learners imagine themselves to be members of the Chinese community, even though they inhabit an English-dominant society, and learn Mandarin to foster this connection), *linguistic hegemony* (the dominant cultural group, in this case the Mandarin speaker group, create a consensus among the minority groups that their language is the standard norm), and *language investment* (the students learn Mandarin to boost their career prospects in light of its powerful position in the global market).

The authors concluded that the identity construction of speakers from Mandarin and other different dialect backgrounds is essentially different. In terms of *imagined community*, the dialect speakers position themselves in a more detailed location, i.e., the local dialect community, rather than the broader Chinese community imagined by Mandarin speakers. Regarding *linguistic hegemony*, almost all the participants considered Mandarin as of a higher status even though the dialect speakers feel more comfortable with their own dialects. Although the dialect speakers are highly conscious of the official position of Mandarin, it is yet not popular nor favourite among them. As to *language investment*, dialect speakers are more motivated by being able to accumulate symbolic capital for a favourable position in the global job market compared to heritage Mandarin speakers, while the latter group is also driven by the cultural bond. In conclusion, this study indicates the need to offer separate and differentiated support for the maintenance of the dialects.

2.2. Factors impacting HL speakers' proficiency

Several factors have been suggested to have an impact on the speakers' proficiency in the HL.

First, *support from family and community* plays an essential role. Often, little formal instruction in the HL is available to the speakers (Jiang, 2021), which means that the family environment usually serves as the main source of HL learners' language input. Broadly speaking, family and parental support have long been acknowledged as a vital condition in children's language development and preservation (Han, 2020; Schwartz & Verschik, 2013; Spolsky, 2012; Tse, 2001). Current studies on CHL also stress the positive impact of family support on learners' HL development (e.g., Lao, 2004; Li, 2006; Jiang, 2021).

Among all the relevant studies conducted on the effect of family and community support on CHL proficiency, a study by Mu and Dooley (2015) shows a rigorous research design and has a comparatively large sample size. This study investigated the role of family support on CHL learners in Australia. 230 Chinese-Australian learners aged between 18 and 35 partook in the online survey which deployed Likert scales to collect information on their *self-reported language skills*, *language use patterns at home*, and the *family support* they received. Those participants spoke a range of CHLs at home, e.g., Mandarin, Cantonese, and other Chinese dialects. According to the regression model, the *family support variables* were more able to predict the participants' self-reported CHL proficiency than other *demographic variables* such as *the age of immigration*, *place of birth*, and *generation of immigration*. In the follow-up semi-structured interviews with five participants, three main aspects of family support emerged as important: *family encouragement*, *informal HL instruction* (e.g., through family discussion in HL, watching TV series in the HL, etc.), and *family sending children to schools for formal HL instruction*. A limitation of the methods of this study is that they used participants' self-report of their linguistic abilities, which might have resulted in a biased evaluation of their language proficiency. Scholarly discussions have revealed potential problems of the self-report method. To be more specific, participants' self-report might lack in credibility in comparison to direct tests. They might be subject to motives like consistency seeking, self-enhancement, as well as self-

representation (see. e.g., Paulhus, & Vazire, 2007; Robins & John, 1997). In light of this, a direct testing of speakers' language proficiency would be more accurate and desirable in reflecting their true competence level rather than sole reliance on the self-report. This is also considered when designing the methodology for the current dissertation, which combines learners' self-report with a direct test.

An ideal measure of HL proficiency would be direct tests, however, given that participants in this study spoke a range of HLs and that no standard tests were available that were matched across the languages, self-report seems to be a reasonable resort. Many scholars also validate the reliability of learners' self-report in language proficiency (e.g., Ma & Winke, 2019; Smith & Briggs Baffoe-Djan, 2022) and it has been found that self-reporting measures correlate highly with direct testing measures in terms of HL ability (see. e.g., Kang & Kim, 2012; Oh & Fuligni, 2010). In light of this, the conclusion of this study that family fulfils a crucial role in CHL learners' language development and acts as an indicator of their language proficiency seems fairly reasonable. It is worth noticing, however, that since the participants came from varied CHL backgrounds, it would be more insightful if a comparison between Mandarin speakers and other dialect speakers were provided.

Another study that has also focused on CHL proficiency and home language environment factors but used a direct testing measure rather than self-report is Zhang and Koda (2011). They distributed a *Home Literacy Environment Questionnaire* among 36 Grade-3 (average age = 9.12 yrs) Chinese-English bilingual children learning Mandarin Chinese as an HL at a weekend community school in the U.S. The questionnaire collected information on their *home language use* (the language the children and their parents use when talking to each other), *joint reading practise in CHL with family* (how often the family helps the children with their Chinese homework and how often they read Chinese books together) and *independent CHL reading practise* (how often do the children read Chinese books related and unrelated to their Chinese homework). To measure their proficiency in CHL, the researchers administered a word-knowledge test to the children that contained 1) a *Radical Identification Task* to identify the graphic components of a Chinese character and 2) a

yes/no checklist to identify whether or not the children know the meaning of the word (they are shown the character in its written form) item.

The key findings of this study were that the children who engaged more in the CHL schoolwork and who had more independent reading practise or joint reading practise with their family showed a higher level of CHL word knowledge. This study has an advantage over Mu & Dooley (2015) in more directly testing speakers' CHL proficiency rather than relying on a self-report. Nevertheless, there are limitations in the study design. The validity of the use of a checklist to identify children's knowledge of word meaning is very limited and it could be improved by combining it with other direct testing measures (e.g., the participants could perform a test where they are asked to tell the word's meaning in English) for triangulation. To be more specific, performing a checklist means it is possible that the participants assume themselves to know the correct meaning of the word but in fact do not. Another problem is that the overall low-frequency data in children's reading practise could possibly lead to a floor effect when running the data analysis. Besides, although this study specifically looks at the literacy aspect of language proficiency, it would yield more insight if the researchers could also consider testing the speaking or listening ability of the CHL speakers. It is worth noticing that this is especially important when it comes to Chinese dialects, which will be further discussed below in Section 3, since the dialects might possess essentially different acoustic features compared to Mandarin. To be more specific, the same character could match up to a completely different phonological word in two different dialects, since the dialects share the writing system but the spoken language could be mutually unintelligible.

Second, the importance of *HL schools and programmes* is stressed by various scholars of language development (e.g., Chinen & Tucker, 2005; Cummins, 2005; Fishman & Nahirny, 1966;). This is also the case for CHL learning, as schools and programmes provide a valuable opportunity for learners to acquire the HL (e.g., Chow, 2001; Du, 2017; Han, 2020; Li, 2006; Pu, 2006; Xiao, 1998; Zhou & Li, 2003). As even in the above-mentioned study by Zhang and Koda (2011), the engagement in schoolwork in HL could also contribute to learners' HL proficiency. In the following paragraphs, I will further look at evidence that could support this argument. Since

sparse research has focused on CHL regarding the factors mentioned in the following section, I will review studies in a broader HL context (e.g., Vietnamese, Japanese) that could provide insight into this topic.

Maloof et al. (2006) looked at the relationship between *HL school variables* and *HL competence* of 33 Vietnamese HL learners in the U.S. The HL school variables were *the length of attendance*, *regularity of attendance* (both reported by teachers), and *class participation* (self-reported by students), while *HL competence* consists of students' proficiency in the *communicative* domain (self-reported proficiency for understanding, speaking, reading, writing, and the ability to communicate in a variety of situations) and *cultural content* domain (answering questions about cultural artefacts such as proverbs and ethnic holidays in the HL, which then was turned into numerical scores to represent knowledge of the HL culture). Multiple regression models showed that the three *HL school variables* jointly accounted for 23% of the variance in participants' HL competence in Vietnamese. Although the sample size means this study was low powered, it offered suggestion that HL school attendance predicts the measure of HL communicative and cultural proficiency.

This study has the advantage that it looked at HL proficiency both in terms of speakers' communicative ability and their cultural knowledge. Cultural awareness and linguistic knowledge are inseparable and both important. Cultural competence could boost linguistic competence (by providing contextual support for speakers' linguistic ability and especially facilitating the process of comprehension (see. e.g., Kozhevnikova, 2014)), and vice versa (linguistic knowledge could enhance speakers' ability to engage in cultural activities). However, these two are separable constructs and it would be interesting to know the factors that predict them separately. Similar to the study of Mu and Dooley (2015), this study relied on self-report which is potentially problematic. This study could consider combining self-report with direct tests to achieve higher validity. The study also pointed out that where there is not sufficient school support for the HL, parents rather than trained teachers are usually the volunteered HL instructors to their children, especially in the context of dialect preservation. This indicates that the two factors investigated above, i.e., home and school factors, might be closely related. As will be discussed later, this is the typical

case in Shanghai, where the family offers the most linguistic support and functions as the language instructor of the children.

Third, HL learners' *attitudes and motivations* could also exert influence on their HL learning. Broadly speaking, attitudinal questionnaires and interviews are often employed in the field of Applied Linguistics (Dörnyei, 2007), and especially so in the context of multilingualism (Baker, 1992). Moreover, they play an essential role in language revitalisation (Bell, 2013). When the learner holds a positive attitude towards the HL and the ethnic group, the language acquisition process is positively supported (Duff & Doherty, 2019; Montrul, 2010; Smith & Li, 2020; Tse, 2001). Similarly, motivation is considered to be a predictor of success in language learning in the broader picture (Dörnyei, 2019; Dörnyei & Ryan, 2015). In the specific context of CHL, I will discuss several studies that examine the role of learners' attitudes and motivations on their language proficiency.

Mori and Calder (2015) investigated the role of motivation and identity on 116 HL Japanese learners' (aged 15-18 yrs) vocabulary development. They conducted a 120-item multiple-choice Japanese vocabulary test and a Likert-scale questionnaire survey with 25 motivational items (e.g., *You study Japanese because you will return to Japan someday and resume your education and live there*; *You study Japanese because it allows you to communicate with your relatives in Japan*.). The results showed that Japanese HL learners' preference for going back to their home country in the future had a significant positive correlation with their Japanese vocabulary knowledge, even though the regression model showed that it was not a significant predictor. To be more specific, when all the correlated factors are taken into account, only their *self-identification of being American* and their *career-oriented motivation* negatively predicted their Japanese vocabulary and respectively counted for 6% and 4 % variance of it. In other words, the more they identify themselves as being American, and the more they learn the HL Japanese mainly because they want to gain a favourable position in the job market, the less their Japanese vocabulary knowledge is. On a general level, the key finding in this study is that a favourable attitude towards the home country was associated with a higher level of language proficiency in the HL.

In a study by Smith and Li (2020), 58 students aged from 10 to 18 years old in a Chinese language weekend school in the U.S. learned Mandarin Chinese explicitly along with other subjects mediated in Mandarin. The participants completed 1) *Language learning attitudes and effort questionnaire*, 2) *Language and Social Background Questionnaire* and 3) *Heritage bilingual reading attitudes questionnaire*. The results showed that participants' self-reported Chinese reading ability correlated significantly positively both with their *extrinsic motivation* (the motivation to do something because of enjoyment) and *intrinsic motivation* (the motivation to do something to achieve rewards). This is in line with the findings in the previously reviewed studies that there is a significant correlation between HL learners' motivation and their language proficiency in the HL. It is also worth considering, however, that such a correlation does not reveal the direction of the association. That is to say, it is unclear whether the learners achieve higher proficiency levels because they are more motivated, or they are more motivated because they are better at the heritage languages. This phenomenon will be further discussed when considering the results from the current thesis.

2.3. The HL preservation movement in Shanghai

Shen (2016) conducted a qualitative study by interviewing 10 stakeholders and analysing their perceptions and experiences of the "*Save Shanghai Dialect*" movement mentioned in the introduction chapter. The 10 participants included two language policy officials in Shanghai, two linguists, two delegates of the local political advisory committee, two parents whose children were enrolled in the trial kindergarten programme aiming to preserve Shanghainese, and two members of a local volunteer group promoting Shanghainese preservation on varied social media platforms. The interview data shed light on the local insecurity about the status of Shanghainese. Both the grassroots and social elites considered the descending of Shanghainese as a menace to the local identity and results in a sense of insecurity. For example, the language policy officials indicated that 90 % of the petition letters and complaints they received between 2012 and 2014 were requesting for actions to save the Shanghai dialect and maintaining its vitality. Similarly, the interviewees who worked on promoting

Shanghainese online indicated that in a survey they conducted, over 83% of the 30,000 respondents agreed that Shanghainese was in a crisis. The two linguists in the interview suggested that Shanghainese is undergoing a deviation from its expected norm (e.g., its standard grammar) and that the younger generations are becoming considerably less fluent in Shanghainese. They thus undertook an enormous effort to work on the language policies in the hope to promote the local dialect. Other interviewees have also indicated their concerns based on their observations that the generational transmission of Shanghainese has almost stopped and that young people in Shanghai have lost interest in the local cultural activities, e.g., the Hu opera and *Huajixi*, the local comedy.

These local concerns have pushed the policy-makers to put forward corresponding actions. In 2016, the Shanghai Municipal Language Committee (SMLC) arranged with other sectors to provide Shanghainese broadcast in public transportations (e.g., on buses and some of the underground lines), several Shanghainese-medium TV and radio programmes as well as 24 trial kindergarten and school projects with the focus on dialect preservation and cultural experience. Nevertheless, multiple interviewees in this study, including linguistics working on the Shanghainese dialect preservation, famous actor in the field of Shanghai local comedies, and volunteers promoting Shanghainese online, indicated that these actions were far from enough to leave a real impact. For example, it is suggested that the trial kindergarten programmes which aim to raise children's awareness of Shanghainese is by no means sufficient for them to actually acquire the dialect and that they need more exposure to Shanghainese both in and out of class. Apart from this official initiative, however, most other movements are all informal and are only launched by e.g., local celebrities, people working in the local culture promotion industries, and local communities (see. e.g., Xia & Shen, 2019). It was also noted that under the overarching policy structure in China, the local policies aiming at saving the Shanghai dialect are still to a great extent constrained. The impacts of the movements initiated by grassroots and small-scale institutional measures are most likely to be meagre. In other words, these efforts are rendered merely trivial considering the unchallenged dominance of Mandarin. In sum, the preservation of Shanghainese has been an essential focal point of both local communities and some language policy-makers, albeit quite limited under the whole

language structure nationwide in China. It is obvious that greater effort in the dialect promotion is urgently needed.

The movement to preserve Shanghainese has been ongoing ever since, and in 2016 the Shanghai Municipal Language Commission (SMLC), in the aforementioned initiative, stressed the importance of 1) building and maintaining a web-based platform for Shanghainese and cultural resources that showcases the characteristics of Shanghainese, oral culture, and humanistic customs; 2) encouraging and supporting the construction of Shanghainese and culture museums in universities, and providing language services to the public in terms of knowledge of Shanghainese and culture; 3) expanding the scope of the trial programmes and continue to provide Shanghainese education and promote cultural experience in kindergartens; 4) guiding and supervising streets, towns and communities to play their part in the transmission of Shanghainese (Shanghai Municipal Education Commission, 2016). In light of this, it is essential to investigate the language usage patterns of Shanghainese and speakers' proficiency thereof for the movement to be properly evaluated and for further actions to take place.

2.4. Relevant empirical research in the Chinese context

As is discussed above, previous research has indicated that HL speakers' language proficiency is correlated with *support from family and community, HL schools and programmes*, and learners' *attitudes and motivations*. Since to the best of my knowledge, HL schools and programmes remain scarce and no research has focused on their efficacy so far, the following section will look at studies on the other two aspects, i.e., family and community language environment as well as learner attitudes.

Wang (2017) conducted a mixed-method study to investigate the role of speakers' attitudes on their linguistic behaviours. It also examines the formation process of a new language, i.e., the new vernacular combining Mandarin Chinese and the local dialect in Hohhot, China. The locally-born residents speak the Hohhot dialect while the immigrants speak Mandarin. Among the 67 participants (32 with a migrant background and 35 with a local background), a questionnaire was used to collect data on their attitudes and identity information. Meanwhile, information regarding the participants'

social contact with dialect speakers and other demographic information was collected through an interview. An elicitation task was deployed to collect speakers' linguistic behavioural data. In the tasks, the participants produced a set of disyllabic lexemes known as "*l-words*", which contain variables indicating a preference for Mandarin or dialect characteristics (e.g., a strong-weak stress pattern is typical of Mandarin while a weak-strong stress pattern is the standard of the dialect). The findings indicated that speakers' attitudes predicted their linguistic behaviour patterns. To be more specific, a more positive attitude towards the local dialect would generally indicate a higher probability of using the dialect. Also, this study suggested that the participants' attitudes and identities exert complex influence in the new vernacular formulation process.

It is worth noticing that an *Attitude Analog Scale* (AAS, Llamas & Watt, 2014) was adopted in this study rather than traditional Likert scales. The AAS elicited participants' attitudes towards the dialect and local community and could to a great extent mitigate the downside of self-report techniques such as Likert scales. To be more specific, Llamas and Watt (2014) emphasised that Likert Scales suffer from disadvantages including 1)the ambiguity of the mid-point (the mid-point in the scale could either indicate a neutral attitude towards the statement or that the participant was unsure about the statement), 2) the limitations imposed on responses (the pre-determined answering points limit the respondents' ability to deliver an opinion), 3) the central tendency bias (the respondents are generally prone to avoid the extreme categories like "1" and "5" on a scale from 1 to 5) and 4) it could only produce ordinal data (so that the data analysis to be carried out would be less precise and more constrained than that based on interval or ratio data). In light of this, the AAS with a magnitude continuum could yield a more fine-grained measurement of learners' attitudes. The participants' markings on the linear scale are converted into a numerical percentage.

Particularly related to the current dissertation is a study by Yu and Yang (2016) which investigated the diachronic changes in Shanghai local youths' dialect use patterns and dialect competence with relevant data from the past 15 years. They collected data on the language use, dialect competence, language attitudes, and

demographic information from 2515 local students in primary and secondary schools in Shanghai. They then compared these data with those from relevant surveys previously conducted from 2000 to 2007 (see., Jiang, 2006; Steering Group Office for Survey of Language Situation in China, 2006; Sun et al., 2007). The results indicated that local students' dialect competence and the status of Shanghainese have largely declined. For example, the proportion of speakers who acquired Shanghainese as their L1 has declined from 93% in 2000 to around 50% in 2016, and the majority of participants consider that they speak Mandarin Chinese more fluently than Shanghainese. Even in the family environment, over half of the participants choose to speak Mandarin rather than Shanghainese. They also found that speakers' belief in the usefulness of Shanghainese and Mandarin, their first language (Mandarin or Shanghainese), and the frequency of dialect use in family and public predicted their self-reported proficiency in Shanghainese. These findings are consistent with those of Mori & Calder (2015), Smith & Li (2020), and Wang (2017), which were discussed above.

Although this study gave a clear picture of Shanghainese teenagers' dialect use and it identified some factors which predict speakers' proficiency in CHL, the sole reliance on self-report might limit its validity in depicting participants' proficiency level, as is discussed above. Moreover, the language attitude and usage patterns collected in this study only focused on very narrow aspects. For example, when investigating the participants' language attitude toward Shanghainese and Mandarin, the researchers used only four broad questions of *whether it sounds good*, *whether it is useful*, *whether it feels intimate*, and *its social impact* to elicit the participants' responses. Utilising more detailed descriptions and breaking them down into varied scenarios might have led to more nuanced in-depth findings. Moreover, to be able to look deeper into participants' dialect competence, the study could further use a follow-up interview to collect some qualitative data so that the analysis could be better supported and the participants' language usage patterns could be better explained.

Since this study, however, no significant work has been done on the linguistic patterns and the HL proficiency of Shanghainese speakers. Given that previously discussed evidence has all pointed to the fact that Shanghainese is currently in an urgent

situation, this dissertation hopes to fill in the substantial gap in the literature and will be exploring speakers' language use patterns in Shanghainese as well as the impact factors of their proficiency levels which cover a more comprehensive range of aspects.

2.5. The current study

On a general level, this current study takes a step forward in comparison to the studies reviewed because 1) instead of merely using participants' self-reported for their language proficiency (see., e.g., Mu & Dooley (2015); Smith & Li (2020)), this study combined the participants' self-report with a specifically designed test containing different aspects of their HL proficiency (e.g., vocabulary, idioms, slangs) and also looked at the relationship between the tested and self-perceived language proficiency; 2) this study has a relatively large sample size of 197 participants (among which 143 finished both the questionnaire and the language proficiency test) compared to the sample size of e.g., 33 in Maloof et al. (2006), 36 in Zhang & Koda (2011), 58 in Smith & Li (2020); 3) this study recruited the participants from 5 different high schools in Shanghai rather than just one school (e.g., Mori & Calder (2015), Zhang & Koda (2011)), resulting in a more diversified and comprehensive participant background; and 4) instead of just investigating the correlation between HL proficiency and various impact factors (e.g., Smith & Li (2020)), this study further employs the regression model and examines the predicting factors of participants' proficiency in Shanghainese.

3. Methodology

This chapter provides a summary of the methodology of the current study, including the research design, participants, instruments, data collection and analysis. After that, the ethical concerns are discussed.

3.1. Research questions

The research questions of the current study and their respective hypotheses are as follows:

RQ1. How can learners' Shanghainese HL usage patterns be characterised?
This RQ has no hypotheses and aims only to identify patterns that emerged from the data collected.

RQ2. How strongly correlated are learners' self-reported and tested HL proficiency?
The hypothesis for this RQ is that learners' self-reported proficiency in Shanghainese and their tested proficiency are positively correlated.

RQ3. What factors are related to learners' proficiency level in Shanghainese?
a. more engagement in cultural activities will lead to a higher HL proficiency level
b. a more positive attitude will lead to a higher HL proficiency level
c. a more dialect-oriented home/ community language environment will lead to a higher HL proficiency level
d. a more local demographic background will lead to a higher HL proficiency level

3.2. Participants

3.2.1. Recruitment information

The target population of this study was Shanghainese high school and undergraduate students. The link to the questionnaire and the test were sent to their English teachers and then distributed to the students by the teachers. The completion of the questionnaire and test was voluntary, and the purpose of the study was explained

to the teachers and students beforehand along with the information sheet. They were also informed that this study had CUREC approval.

The five high schools are in the urban area of Shanghai and all are considered to be among the most elite secondary schools in Shanghai. The link to the questionnaire and test were distributed to the participants by their English teachers. The participants came from all three year groups in the high schools, i.e., from Year 10 to Year 12 (age 16 to 18). For participants in Year 12, this was after one of their major monthly tests and approximately one month ahead of their National College Entrance Examination, which means the participants had a relatively low workload that allows them to have time to undertake this study. For participants in Year 11 and 10, the time period that they finished the questionnaire and test also occurred when they did not have any of the major tests or heavy study burdens.

The undergraduate participant group came from a prestigious university in Shanghai. This time period was approximately 2/3 through the term and the participants had already finished the mid-term exams. This means that the questionnaire and test were not in conflict with any major exams.

Although we aimed for a balance of high school and university students, the final sample recruited 197 participants, among which 10 were undergraduate students and the other 187 were high school students. Since the link to the questionnaire and the test was distributed to the high school participants by their teachers, the students were likely to take them more seriously. However, presumably because the link was sent to the undergraduate students by their fellow students, the participation rate was considerably lower, with only 10 undergraduate students completing the questionnaire and the test. Since the sample size of the undergraduate group was too small to be analysed on its own, and the means of most aspects collected from this group varied significantly with the high school group, they were excluded from the data analysis.

The study also deployed G*power to calculate how many participants would be necessary to achieve a desirable level of power. Since none of the relevant studies reviewed in the literature review chapter explicitly indicated their effect sizes, this study adopted the medium level cut-off point, i.e., $r = 0.3$, suggested by Cohen (1992). According to Cohen (1992) and Field (2013)'s guideline, the α-level was set as 0.05

and the power as 0.8. The calculation (see. Appendix 8) showed that a total sample size of 84 would be desirable. This means that the current sample size of 197 (143 if running correlation with the participants' tested proficiency score) is more than ideal.

3.2.2. Participant drop out

Since the questionnaire and the test aimed to collect in-depth and rich information, it took approximately 15 to 20 minutes to finish all the questions. Even though the time estimate was clearly stated in the information sheet as well as in the first section of the questionnaire, this was still a considerable amount of time for some participants. As a result, among all the 197 recruited participants, 54 participants only finished the questionnaire and left the test unfinished. 143 participants completed both the questionnaire and the test. It is also possible that the former group did not take part in the Shanghainese proficiency test because they were not confident of their language competence in the HL. This could have resulted in the data collected being biased, as will be addressed in the discussion.

3.3. Measurement instruments

3.3.1. Questionnaire

The questionnaire was created on Gorilla and completed online. It consisted of six sections. All the information and instructions in the questionnaire were written in both English and Mandarin Chinese to guarantee a maximum level of understanding as possible. The full version of the questionnaire is included in Appendix 2.

The **first section** presented generic information about the study and asked for participants' consent.

The **second section** inquired about the demographic background of the participant. This includes their gender, age, current level of study, the hometown of their parents, their length of residence in Shanghai, and their generation of immigration in Shanghai, if applicable.

The **third section** collected information on their attitudes and beliefs about the HL Shanghainese. This section used 22 statements about Shanghainese and asked to what extent the participants agree with these statements using an Attitude Analogue Scale (AAS). The scale ranged from -50, meaning completely disagree, to 50, meaning completely agree. The participants could drag the slider to present their perception of the statement. The statements were adopted from a previously validated questionnaire investigating the same attitudinal aspects perceived by users of an HL (see., Wang, 2017). Some of the statements were deleted or rephrased to fit the current study. For example, original statements from (Wang, 2017) like "The reconstruction in the Old Town has destroyed the culture of Hohhot." were deleted because the Old Town vs. New Town scenario does not fit the situation in Shanghai. In addition, three original statements ("I have an SH accent when I speak Mandarin"; "I like my SH accent when I speak Mandarin"; "I can recognise the SH accent when I listen to other people speak Mandarin") designed by the current author were included in the questionnaire, which further probe into the participants' impression and attitudes towards Shanghainese. The statements were also modified to avoid ambiguity in meaning according to the result of the pilot test. Among the 22 statements, four were negatively worded where disagreement with the statement would indicate a favourable attitude towards the HL. The other 18 statements were positively worded.

The **fourth section** investigated the language usage patterns of the participants in their family and school environment, and their family members' language usage patterns. Analogue scales similar to the abovementioned AAS were adopted here as well to represent the proportion of the time the participants and their families spend speaking in Mandarin or the HL Shanghainese. Where languages other than Mandarin and Shanghainese are also spoken in the family, participants were asked to indicate all the languages used in the family and their respective proportion of time of the usage. This section was also designed by the author from scratch.

The **fifth section** asked for participants' self-perception of their language proficiency in Shanghainese, which also deployed the analogue scale. The participants rated their perception of speaking in and listening to Shanghainese on a scale ranging from "very easy" (50) to "very hard" (-50).

Finally, the **sixth section** used a 5-point Likert scale to investigate participants' engagement in media and cultural activities. This included their involvement in community- and school-hosted dialect activities (e.g., dialect knowledge competitions), their attendance to dialect-mediated cultural events (e.g., stand-up comedies), and their media use (e.g., listening to broadcast and TV shows mediated in Shanghainese).

3.3.2. Language proficiency test

After the abovementioned questionnaire, a language proficiency test containing 20 multiple choice questions was carried out to investigate participants' competence level in the HL Shanghainese. The test was also carried out online using Gorilla. The test was written in *Simplified Chinese* (the contemporary Chinese written language contains 1) *Traditional Chinese Characters*, which are used in Taiwan, Hong Kong, and Macau, and 2) *Simplified Chinese Characters*, which are employed in mainland China, and Singapore). The full version of the test is presented in Appendix 3.

The items in the test were designed from scratch by the researcher of this study and focused on participants' knowledge of Shanghainese vocabulary. The researcher referred to *Shanghai dialect vocabulary* (1991) and selected a range of words of different frequency levels to be tested. The chief editors of this book include Qian Nairong and You Rujie, who co-initiated the aforementioned "*Save Shanghai Dialect*" movement in the literature review section. Since the reference book was published three decades ago, the researcher made efforts to make sure that the items selected for the test were not outdated and obsolete. This included consulting high-proficiency Shanghainese speakers from the researcher's family and friend circle, as well as checking the number of search results of the words online to make sure the word item is up to date and of common use. The items eventually selected included daily object vocabulary, collocations, idioms, traditional lullabies, etc.

As is mentioned in the introduction section, although Shanghainese could be transliterated into the writing system of Mandarin Chinese, it would hardly make any sense since the two dialects contain significantly different lexical items. Hence, audio prompts were implemented in the questions (unless the question specifically tests the

participants' knowledge of the word pronunciation) to help participants understand the items and to alleviate the potential misleading from the transliteration into Mandarin Chinese. Each recording could be played up to two times. The audio prompts were recorded by a friend of the researcher, who possesses a high proficiency level in Shanghainese, specifically for the current study. In addition, all the pronunciations of the items were double-checked with local Shanghainese speakers across generations from the speaker's family to make sure they are correct and comprehensible.

To evaluate the statistical properties of this newly designed test, the normality of distribution was measured. The skewness and kurtosis indicated that the data from the test scores were normally distributed, while a histogram showed no observable floor or ceiling effect. This has suggested that the Shanghainese proficiency test was reasonably designed.

To test the reliability of the newly designed Shanghainese proficiency test, a correlation test was carried out between the first section (8 items) and the second section (12 items) of the test. Since participants' scores in the second section were not normally distributed, a Spearman's ρ test was operated. The results showed that the two sections were significantly positively correlated ($r = .592$, $p < .001$). According to Cohen (1988, 1992), this is a large effect size.

3.4. Pilot test

Both the questionnaire and the language proficiency test were piloted before being distributed to a separate set of participants to those taking part in the main study. This is in line with Gorard (2001) that a full pilot test of the complete research design is necessary.

The five pilot participants (four female, one male, average age = 22.4 yrs) were recruited through the researcher's own network. All of them were born and raised in Shanghai, fluent in both Chinese and English, and had different levels of knowledge of the HL Shanghainese. They came from different districts in Shanghai, so that potential regional differences could be detected and avoided in the test. Two of the pilot participants were undergraduate students, while the other three were in the first year of

their postgraduate study. Thus, the pilot participants' levels of study roughly matched those of the actual participants. The pilot participants provided feedback on the wording of the questionnaire items and the author made corresponding modifications. No significant issues emerged. The pilot test made sure that the questionnaire was of reasonable length, comprehensible, and easy to follow. Thus, a second pilot test was not necessary.

In terms of the language proficiency test, different perceptions of its difficulty were responded by the pilot participants which approximately matched their proficiency level in Shanghainese. This indicated that the test was able to differentiate between competence levels among the participants and was neither too easy nor too difficult. In this way, a potential ceiling or floor effect could be avoided. Besides, it was pointed out by the pilot participants that several lexical items in the original test were not used in their families. Thus, changes were made to make sure that all the items are generally used across different districts in Shanghai. The goal was, as far as possible, to avoid having regional differences as a potential founding factor. The pilot test also ensured that the audio prompts could be played without any technical issues and that the link to the survey readily works.

3.5. Data Analysis

The raw data from the questionnaire and the score of the language proficiency test was entered into SPSS for descriptive and inferential analysis. The negatively worded statements from the questionnaire were reverse coded.

3.6. Ethical considerations

Prior to the data collection phase, this study was approved by the University of Oxford's Central University Research Ethics Committee (see Appendix 1). Recruitment of this study commenced after ethical approval was secured. All adult participants, the minor participants, and their parents/ guardians in this study received comprehensive written information about this study. Consent was obtained online as part of the questionnaire from those who agreed to participate. For those under 18 years

old, an opt-out form was distributed to their parents or guardians by the teachers. The parents or guardians could sign the form so that their kids would drop out of the study. They were also encouraged to contact the researcher with any questions before, during, and after the data collection. It was clearly explained that the data would be completely anonymous and would be stored in a safe database of the University of Oxford. It was also made clear that they were able to withdraw from the research at any time. The link to the questionnaire was sent to the school teacher via the Chinese social media platform WeChat along with the information sheets.

Efforts were made to guarantee the anonymity and confidentiality of the participant information. All the digital data was stored on the researcher's computer in an encrypted and password-protected folder.

4. Results

First, the potential confounding variables are investigated. This chapter then presents descriptive statistics and, where appropriate, inferential statistics relevant to the RQs.

4.1. Confounding variables

4.1.1. Study level

The 197 participants in the study included high school students (n = 187) and undergraduate students (n = 10). The Independent-Samples Mann-Whitney U Test indicated that the two groups were significantly different in their scores for *attitudes and beliefs*, *tested proficiency*, *cultural activities*, their *self-reported speaking and listening proficiency*, as well as their *home language environment* (all $p < .05$). Since the sample size of the undergraduate students is not large enough to be analysed on its own, this part of the participants is not included in subsequent analyses addressing the research questions.

4.1.2. Gender

The 197 samples collected in the study included 88 female participants, 105 male participants, and 4 non-binary participants. Since the non-binary group (n = 4) was too small for statistical analysis, only the female and male participants were compared. According to an Independent-Samples T-Test, the female participants (M = 389.84, SE = 33.49) had a significantly higher score for *attitudes and beliefs* towards Shanghainese than their male counterparts (M = 299.64, SE = 29.88). The difference was significant ($p = .046$, $t(182) = 2.01$). Cohen's d was 0.298, suggesting a small to medium effect size. Among the other variables (*home language environment*, *community language environment*, *cultural activities*, *number of local parents*, *generation of immigration*), there was no significant difference between the two groups (all $p > .05$).

4.2. RQ1: Learners' language usage patterns

RQ1. How can learners' Shanghainese **HL usage patterns** be characterised?

This RQ has no hypotheses since it aims only to identify patterns that emerged from the data collected. The data from the "language usage patterns" section in the questionnaire was examined. This includes which language (Mandarin, Shanghainese, or other dialects used at home) the participants found easier to speak, and the proportion of the time they spoke in each language at home and with peers at school.

In terms of their perception of which language is easier to speak, Figure 1 shows that 74.9 % (n = 140) of the participants chose Mandarin, 21.4 % (n = 42) considered Shanghainese and Mandarin equally easy, while only 2.7 % (n = 5) of the participants thought of Shanghainese as easier to speak.

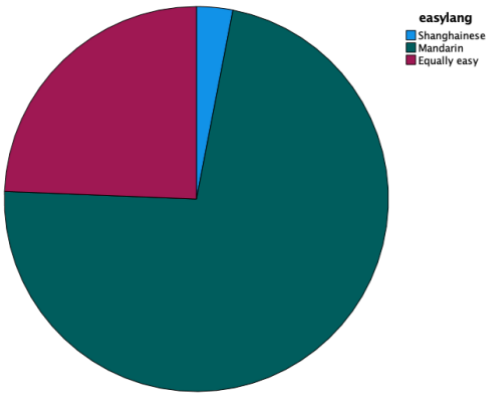


***Figure 1**. Which language did the participants find easier to speak*

In terms of what language the participants spoke the most at home, the majority indicated that they speak more Mandarin Chinese than the Shanghainese dialect. In cases where more than one dialect was spoken at home, all the participants but one also suggested that the proportion of time spent speaking Shanghainese outweighed other dialects.

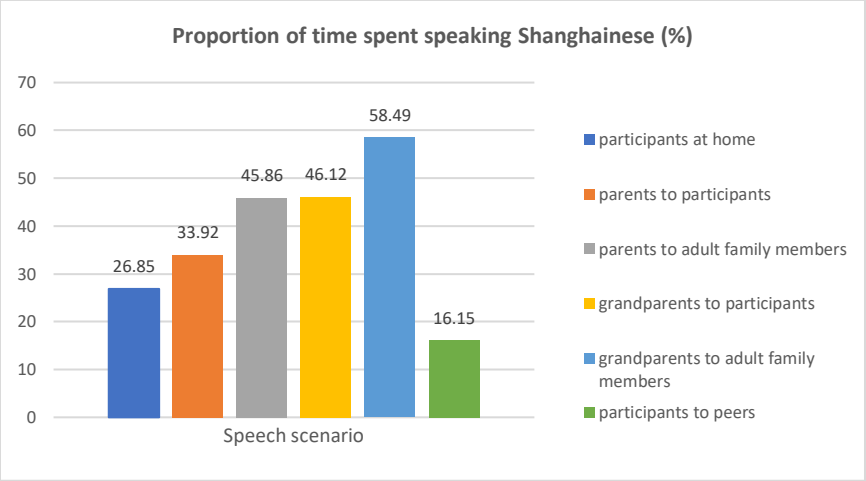


Figure 2*. The Proportion of time spent speaking Shanghainese by the participants and their family members*

This section of the questionnaire also asked the participants to indicate what proportion of time they and their family speak Mandarin and Shanghainese under different scenarios (***participants*** *at home*, ***participants*** *speaking to their* ***peers*** *at school*, *participants'* ***parents*** *speaking to* ***them***, *participants'* ***parents*** *speaking to* ***other adult family members***, *participants'* ***grandparents*** *speaking to* ***them***, and *participants'* ***grandparents*** *speaking to* ***other adult family members***). They were also asked to indicate what other dialects are spoken and how frequent it is, if applicable. The author then converted the data collected to the proportion of time that Shanghainese is spoken in each of these scenarios.

As can be seen in Figure 2, the older generation spoke more Shanghainese at home than the younger generation. The participants' parents and grandparents also spent more time speaking Shanghainese when talking to other adult family members than to the participants. In addition, the participants spoke more Shanghainese at home than at school.

4.3. RQ2: Learner's self-reported and tested proficiency

RQ2. How strongly correlated are learners' **self-reported** and **tested HL proficiency**?

The hypothesis for this RQ is that the participants' self-perceived competence in Shanghainese is positively correlated with their tested proficiency level.

Recall that the questionnaire asked participants to rate their perception of difficulty in listening and speaking on a scale between -50 (very hard) to 50 (very easy). Before directly addressing the research question, the author examined the normality of the measures as well as the relationship between these measures. The normality of participants' self-reported *listening* and *speaking* proficiency and their *tested proficiency* were checked by values of the skewness and kurtosis. The *z* scores of participants' *self-reported speaking* score and the *tested proficiency* fall between -2.58 and 2.58, so they would be regarded as being normally distributed (since the study has a large sample size) (see., Field, 2013). The *z* score of their self-rated *listening proficiency* indicates a non-normal distribution ($z < -2.58$). As can be seen in Figures 3 and 4, the participants' *tested proficiency* and *self-rated proficiency* in *speaking* approximately display a normal distribution.

On average, the participants' perceived *listening* to Shanghainese (M = 28.39, SD = 22.25) as easier than *speaking* in Shanghainese (M= -2.78, SD = 28.32). According to a Paired-Samples T-test, this difference was significant ($t(186) = 14.31$, $p < .001$, $r = .52$). This indicates a large effect size (Cohen, 1988, 1992). According to a Spearman's rho test, the participants' self-reported *speaking* and *listening* proficiency were significantly positively correlated ($r = .255$, $p < .001$).

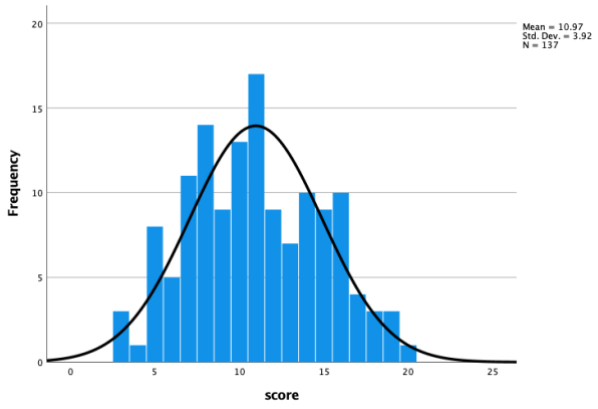


Figure 3. *Distribution of participants' tested Shanghainese proficiency*

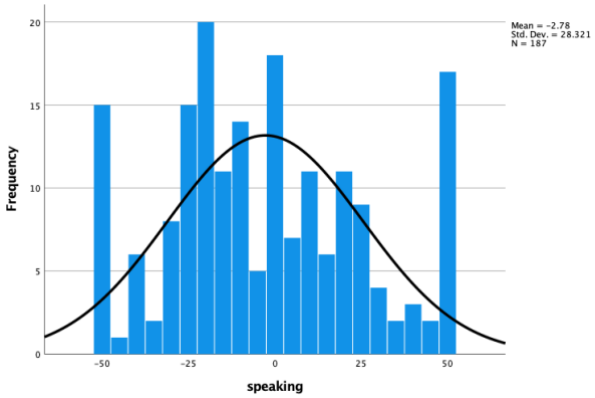


Figure 4. *Distribution of participants' self-reported proficiency in speaking*

Since the *self-reported* proficiency score for *speaking* and the *tested* proficiency score were normally distributed while the *self-reported listening* score was not, a Spearman's ρ was administered for the correlation between the self-reported *listening* score and the *tested proficiency* score, while a Pearson correlation was measured for the relation between the self-reported *speaking* score and the *tested proficiency* score. The statistical analyses revealed that 1) the *self-reported speaking score* and the *tested Shanghainese proficiency* were significantly positively correlated (r = .394, p < .001), and 2) the *self-reported listening score* and the *tested Shanghainese proficiency* were significantly positively correlated as well (r = .426, p < .001).

According to Cohen (1988, 1992), both of the correlation coefficients above would indicate a medium effect size. It also suggests that participants' perception of their speaking and listening proficiency had approximately the same strength of correlation with their tested proficiency scores.

4.4. RQ3: Factors impacting HL level

RQ3. What **factors** are related to learners' proficiency level in Shanghainese?

Hypotheses:

*a. more engagement in **cultural activities** would lead to a higher HL proficiency level*

*b. a more positive **attitude** would lead to a higher HL proficiency level*

c. a more dialect-oriented ***home/ community language environment*** *would lead to a higher HL proficiency level*

d. a more local ***demographic background*** *would lead to a higher HL proficiency level*

4.4.1. Hypothesis A: *more engagement in **cultural activities** would lead to a higher HL proficiency level*

The "cultural activities" section in the questionnaire investigated the frequency of participants' engagement in listening to broadcasts, watching local comedies, talk shows, Hu-opera, and TV shows, which were mediated in Shanghainese, as well as the frequency of their attending cultural events hosted by schools and local communities. The section also enquired about their general frequency of attendance at Shanghainese-related events.

Before addressing the hypothesis, the author examined the frequency of the activities included in the questionnaire. In order to address the hypothesis, an average score across the activities was computed. As can be seen in Figure 5, among the frequency of all the activities, the participants indicated the highest participation rate in *watching local TV shows* in Shanghainese (2.49) and *listening to broadcasts* in Shanghainese (2.45) on a scale from 1 (never) to 5 (always). In contrast, they spent the least amount of time *taking part in dialect-related activities* hosted by the local communities (1.48).

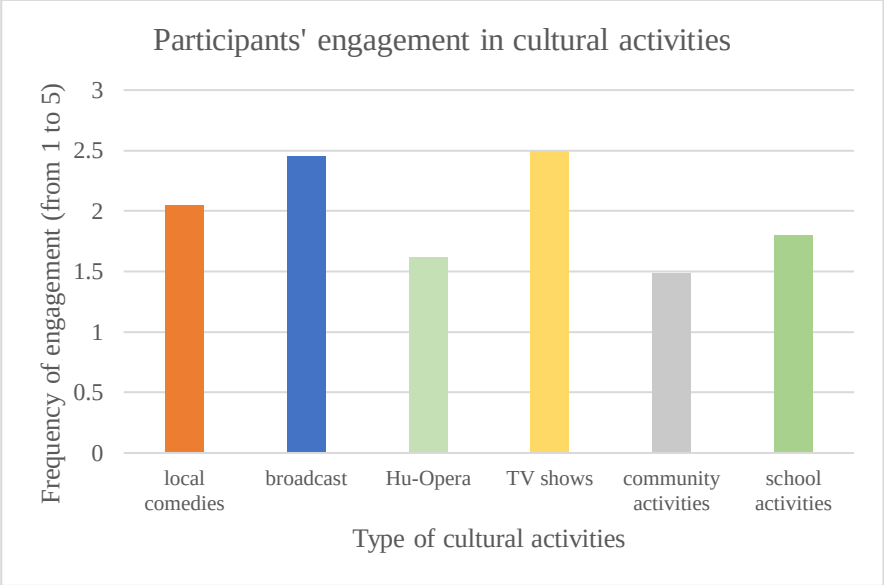


***Figure 5**. Frequency of participants' engagement in cultural activities*

To check the reliability of this section in the questionnaire, Cronbach's alpha (α) was calculated. Cronbach's alpha calculated for this section is 0.827. According to Field (2013), the value of $\alpha > .8$ would indicate good reliability.

According to the z-score ($z > 2.58$), the reported frequency of the *cultural activities* was not normally distributed. Hence, a Spearman's ρ test was administered. The statistics showed that the participants' engagement in dialect-related *cultural activities* was significantly positively correlated with their *tested Shanghainese proficiency* ($r = .307$, $p < .001$).

4.4.2. Hypothesis B: *a more positive **attitude** would lead to a higher HL proficiency level*

The next section in the questionnaire collected participants' attitudes and beliefs towards Shanghainese. Among the 22 items in this section, 18 were positively worded and 4 were negatively worded. All the responses collected from the Attitude Analogue Scale (AAS) were examined and adjusted (i.e., the negatively worded ones were reverse coded) so that a higher score would represent a more positive attitude towards

Shanghainese. Each response ranged from -50 (completely disagree) to 50 (completely agree), which makes the total score from -1100 to 1100.

To check the reliability of this section, Cronbach's alpha was measured ($\alpha = 0.812$). This would indicate a good internal consistency (Field, 2013).

The mean score of all the participants' responses in total was 334.21 (n = 187, SD = 306.43, Min = -700, Max = 951). According to the z-score ($z < 2.58$), their attitude scores were normally distributed, as can be seen in Figure 6.

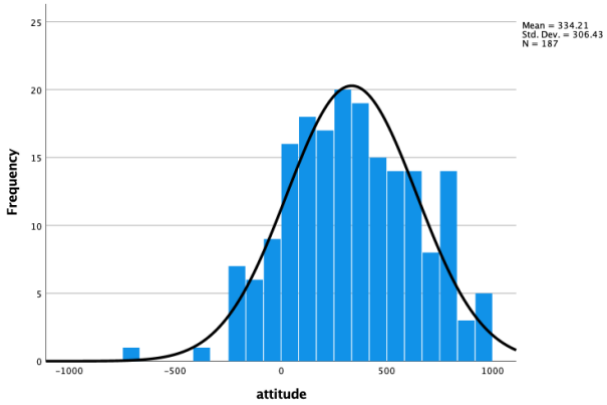


***Figure 6**. Participants' attitudes and beliefs toward Shanghainese*

Pearson's correlation test suggested that there is a significant positive correlation between participants' attitudes and beliefs in the HL Shanghainese and their tested proficiency score ($p < .001$, $r = .524$). According to Cohen (1988, 1992), this is a large effect size.

Statement	Score
It would be sad if Shanghainese disappears in the future.	36.52
I hope to see my children and grandchildren speak Shanghainese.	27.2
If my children spoke Shanghainese, I would oppose it.	-26.72

***Table 1**. Participants' most favoured attitudinal statements*

Statement	Score
If I had an opportunity to live in another city that is similar to SH in economic developments, I would choose to stay in SH.	-18.42
Immigrants should learn to speak Shanghainese.	2.83
If I am talking to another Shanghainese speaker, I will definitely speak the dialect.	3.14

***Table 2**. Participants' least favoured attitudinal statements*

Finally, to shed light on what aspects of attitudes might be provoking the highest levels of agreement and disagreement among the participants, the statements that gained the highest and lowest scores were investigated. Table 1 shows the statements that gained the highest scores (negatively worded items were compared using their absolute values). All three of these statements were concerned with the future of Shanghainese, and the participants showed a highly positive inclination towards the preservation of this HL. Table 2 presents the three statements that achieved the lowest scores. These statements will be further discussed in the next chapter.

4.4.3. Hypothesis C: *a more dialect-oriented **home/ community language environment** would lead to a higher HL proficiency level*

The data on participants' *home* and *community language environment* came from 1) the data on the proportion of time their parents and grandparents spoke each language at home in the section "language usage patterns" and 2) the frequency that the participants' school and community hosted dialect-related events in the section "cultural activities".

The proportion of time spent speaking Shanghainese at home by the participants, their parents, and grandparents were added up as an indicator of their *home language environment*. Since the home language environment scores were not normally

distributed, a Spearman's rho test was employed. Statistical analysis showed that it is significantly positively correlated with participants' tested Shanghainese proficiency (r = .667, p < .001). That is to say, the more the participants were exposed to the dialect at home, whether directly speaking Shanghainese or listening to family members speaking Shanghainese, the more competent they were in Shanghainese. The correlation coefficient indicated a strong relation (Cohen, 1988, 1992).

Regarding the *community language environment*, the participants were asked to indicate how frequently their local community (e.g., their local neighbourhood committee) and their schools hosted events mediated in Shanghainese or aimed to promote Shanghainese. Since the data was not normally distributed, a Spearman's ρ test was employed to test the relation between the community language environment and participants' tested proficiency level in Shanghainese. No significant correlation was revealed (r = .084, p > .05).

4.4.4. Hypothesis D: *a more local **demographic background** would lead to a higher HL proficiency level*

In the "demographic information" section of the questionnaire, the participants were asked to give information on both of their *parents' hometown*, their *length of living in Shanghai*, as well as which *generation of immigrants* in Shanghai they were, if applicable.

The participants reported on average 1.1 parents that are local Shanghainese, i.e., who is born and raised in Shanghai. 93.4 % of the participants indicated that they have always been living in Shanghai since birth. As can be seen in Figure 7, 39.0 % of the participants reported that their family has always been local Shanghai residents, while 40.1 % are the second generation in Shanghai and the other 20.8 % are the third immigration generation.

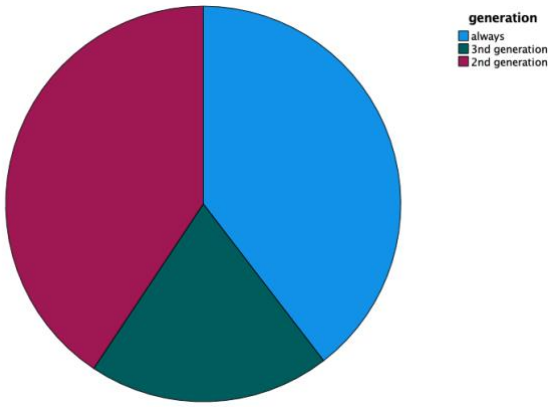


***Figure 7**. Participants' generation of immigration in Shanghai*

Spearman's ρ tests suggested that 1) *the number of local parents* that the participants have (r = .605, p < .001), as well as the participants' *generation of immigration* in Shanghai (2nd generation, 3rd generation, or non-applicable) (r = -.523, p < .001) were significantly correlated with their proficiency level in the HL Shanghainese. That is to say, the more local the participants' parents are, and the longer the participants' family has been living in Shanghai, the higher their proficiency level in Shanghainese was. Both correlation coefficients suggested a large effect size (Cohen, 1988,1992).

Similarly, the participants' *length of living* in Shanghai (more than 10 years, 5-10 years, 3-5 years, or less than 3 years) was significantly correlated with their tested proficiency scores in Shanghainese (r = -.218, p < .05). In other words, the longer they have lived in Shanghai, the more competent they were in the HL.

After the preliminary analysis of the data, a couple of Independent Sample t-Tests were used to compare the proficiency of groups of participants with zero, one, or two local parents. The tests indicated a significant difference in the Shanghainese proficiency scores between 1) participants with *zero* local parents and participants with *one* local parent (t(82) = -5.54, p < .001, r = .52), 2) participants with *zero* local parents and participants with *two* local parents (t(87) = -2.19, p < .05, r = .23), and 3) participants with *one* local parent and participants with *two* local parents (t(99) = -8.49, p < .001, r = .65). That is to say, participants with *two* local parents showed greater

proficiency in Shanghainese than participants with *one* local parent, and participants with *one* local parent showed greater proficiency than those with *zero* local parents. The correlation coefficients in 1) and 3) indicate a large effect size (Cohen, 1988, 1992).

4.4.5. Regression model: predicting factors of participants' language proficiency

After the correlation analysis of the association between participants' tested Shanghainese proficiency and the separate dimensions in 4.4.1 to 4.4.4, a regression analysis was operated to analyse which factors continue to predict the Shanghainese proficiency even when others were taken into account.

The variables of participants' *home language environment*, *community language environment*, *cultural activities*, *number of local parents*, *generation of immigration*, as well as their *attitudes and beliefs* were entered into the regression model in a stepwise fashion to predict their *tested Shanghainese proficiency*. The choice of the stepwise method is in line with Mori and Calder (2015) who explored a similar construct of how motivational factors predicted learners' HL Japanese proficiency level. In addition, Field (2013) suggests that this method is suitable for exploratory model building.

To test the normality of the residuals in the regression model, this section also looked at the histogram and normal probability plot, as shown in Figure 8 and 9. The two figures indicated that the residuals were normally distributed.

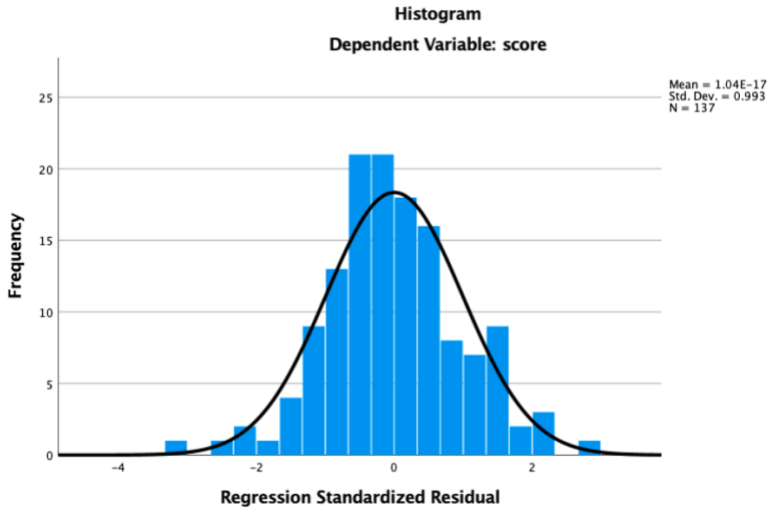


***Figure 8**. Histogram of regression standardised residuals*

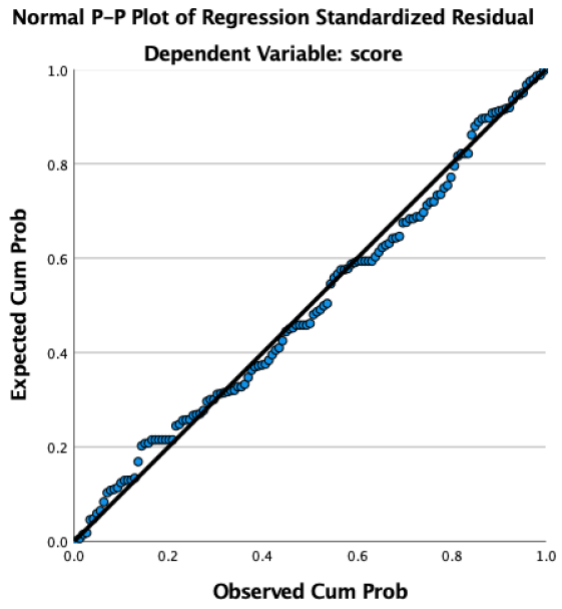


***Figure 9**. P-P plot of regression standardised residuals*

The regression model showed that among all the factors entered, the most prominent predictors for their tested proficiency level in Shanghainese were their *home language environment* and the *number of their parents being local Shanghainese*, as indicated in Table 3. These two factors could together account for 47.1 % of the variance in participants' tested Shanghainese competence. Since the Variance Inflation

Factor (VIF) was well below 10 and Tolerance not lower than 0.20, the regression model did not show a problem of collinearity (see., Field, 2013).

Summary of Regression Analysis Predicting Shanghainese proficiency

	B	**SE B**	**β**	***t***	***p***	**R^2**
constant	7.563	0.429		17.627	<.001	
home language environment	0.016	0.002	0.648	9.90	<.001	0.420
local parents	1.137	0.409	0.250	2.778	.006	0.452

Table 3*. Summary of Regression Analysis Predicting Shanghainese proficiency*

4.4.6. Participants with different parental backgrounds

a. Participants with no local parents

Since the regression model showed that the *home language environment* and the number of *local parents* played the most important roles in predicting participants' language proficiency in Shanghainese, it would be interesting to look into what factors are specifically related to the language competence of participants who do not have this support, i.e., who do not have local parents (n = 48). Since the test score of this subset of participants was not normally distributed, a Spearman's ρ test was administered. There was, however, no significant correlation between any of the variables and the test scores (all p >.05). The only marginally significantly correlated variable was the *home language environment* (p = .07). Hence, there was no need to employ a regression model.

b. Participants with two local parents

Similarly, it would also be interesting to investigate what specific factors contribute to the proficiency of speakers who already have a supportive home language environment, i.e., whose parents are both local Shanghainese.

The test scores of this subset of participants (n = 58) were normally distributed, so a Pearson's correlation test was operated. The results showed that their tested Shanghainese competence was significantly positively correlated with their *attitudes and beliefs* and their *home language environment*.

When entered into a regression model in a stepwise manner similar to the one above, participants' *attitudes and beliefs* were shown to be the most important predictor of the test scores, albeit an R^2 of 0.195 would indicate a relatively low predicting ability. This is shown in Table 4. The VIF value also showed no problem of collinearity.

Summary of Regression Analysis Predicting Shanghainese proficiency (local parent = 2)

	B	**SE B**	**β**	**t**	***p***	**R^2**
constant	10.932	.847		12.905	<.001	
attitudes and beliefs	.005	.002	.043	3.14	.003	.162

Table 4*. Summary of Regression Analysis Predicting Shanghainese proficiency (local parent = 2)*

5. Discussion

The previous chapter presents the learners' usage patterns of Shanghainese, and the data collected showed that speakers' self-reported and tested Shanghainese proficiency were significantly positively correlated. It also suggested that learners' engagement in cultural activities, their attitudes, home language environment, and demographic backgrounds are significantly corelated with their competence in Shanghainese. This chapter explores the results in more depth in the order of each RQ. It also considers the results in the context of the literature reviewed in Chapter 2. Limitations of the current study are discussed in this chapter as well.

5.1. RQ1: Learners' language usage patterns

RQ1. How can learners' Shanghainese **HL usage patterns** be characterised?

5.1.1. Comparison with previous literature

The overall pattern that the majority of the participants found Mandarin easier than Shanghainese and that they spent more time speaking Mandarin than speaking Shanghainese is expected. The findings of this RQ are in accordance with the study by Yu and Yang (2016) which suggested that the status of Shanghainese is at stake in comparison to Mandarin among teenagers in Shanghai. However, findings in this study have indicated that the current situation for Shanghainese is more serious than Yu and Yang (2016) have previously suggested. For example, over half of the participants in their study chose to speak Mandarin rather than Shanghainese at home. The data from this current study showed that 72.5 % of the participants found Mandarin easier to speak than Shanghainese and only 28.9 % of their time was spent speaking Shanghainese at home, which depicted a direr situation for the dialect than the previous study.

5.1.2. Comparison of language usage patterns across generations

The findings in the current study showed a decrease in the proportion of time spent speaking Shanghainese across generations in the family. This is consistent with the trend of Shanghainese losing its population. The fact that the participants' parents and grandparents preferred to speak more in Shanghainese to other adult family members than to the participants might be the reason for the participants not having enough opportunity to practise their dialect. On the other hand, it could also be due to the fact that the younger generations are not proficient enough in the dialect, leading to the family members adjusting their language usage patterns accordingly.

The fact that participants barely spoke in Shanghainese at school with their peers could be because 1) their proficiency level in Shanghainese is not high enough for them to do so; 2) a significant number of students, teachers, and staff at their schools are not local Shanghainese, which means they wouldn't be able to communicate in the dialect. Besides, the participants might find it rude to try talking to non-local peers in the dialect since the addressee would not be able to understand; and 3) as is mentioned in the literature review section, the use of regional dialects has been generally prohibited in schools by the local policymakers in Shanghai; 4) all the material and contents taught in schools are delivered in Mandarin Chinese, which means the schools offer a much more supportive environment for communication in Mandarin rather than Shanghainese. Future research could consider shedding light on this area with more detailed qualitative data, for example, with an interview, so that the Shanghainese learners' behaviour patterns could be better interpreted and understood.

5.2. RQ2: Learner's self-reported and tested proficiency

RQ2. How strongly correlated are learners' **self-reported** and **tested HL proficiency**?

Before addressing this RQ, this study looked at the relationship between participants' different evaluations of their speaking and listening proficiency in Shanghainese. Regarding how the participants perceived their competence in speaking and listening in Shanghainese, their average rating for speaking (-2.78) was considerably lower than listening (28.39) and the difference was statistically significant.

This could be because they spend more time on the receptive skill, i.e., listening, rather than the productive skill, i.e., speaking.

This is consistent with the findings in the language usage patterns in the participants' families. Even though the participants themselves did not invest a large amount of time in speaking Shanghainese, their parents and grandparents still spent a great proportion of their time speaking the dialect. That is to say, even though the participants might not be directly addressed in such scenarios, they were still exposed to sufficient language input, which honed their listening skills in Shanghainese.

The participants' self-reported proficiency in Shanghainese, either speaking or listening, was significantly correlated with their tested competence. This fits with the existing literature that participants' self-reported proficiency in HL is highly correlated with their proficiency tested in direct measures (see. e.g., Kang & Kim, 2012; Oh & Fuligni, 2010). The participants' self-reported listening and speaking proficiency both had a moderate strength of correlation with their tested proficiency, which indicates they had a similarly reliable judgement of the two abilities, even though their perceived speaking as more difficult than listening.

5.3. RQ3: Factors impacting HL level

RQ3. What **factors** are related to learners' proficiency level in Shanghainese?

5.3.1. Hypothesis A

*more engagement in **cultural activities** would lead to a higher HL proficiency level*

Presumably, the fact that participants regarded speaking in Shanghainese as harder than listening is coherent with the data on their participation in dialect-related cultural activities. To be more specific, among all the activities inquired, the participants suggested the highest frequency of listening to the local broadcast, watching local TV shows, as well as watching local comedies or talk shows. In contrast, they invested the least amount of time taking part in community-hosted cultural events. As a result, they hadd significantly more opportunities to enhance their receptive skill, i.e., listening, than their productive skill, i.e., speaking. It is unclear, though, whether the participants

engaged in more listening-related activities because they lack speaking proficiency, or vice versa.

5.3.2. Hypothesis B

*a more positive **attitude** would lead to a higher HL proficiency level*

The finding that participants' attitudes toward Shanghainese was significantly positively correlated with their language proficiency is in line with previous studies, e.g., Mori & Calder (2015) and Smith & Li (2020), which are discussed in the literature review section that suggested a more favourable attitude towards the HL is significantly positively associated with the language competence.

It should be acknowledged here, however, that a significant correlation does not indicate the direction of the association. To be more specific, it is not clear whether the participants held a more positive attitude towards Shanghainese so that they were more competent in this HL, or their positive attitude was a result of their higher proficiency. Hence, interpretations should be made carefully. Nevertheless, this uncertainty in the direction of correlation is true of all studies that take these measures (e.g., Mori & Calder (2015), Smith and Li (2020)).

As is shown in Table 1 and 2 in Chapter 4, the statements that achieved the highest scores generally showed participants' attitudes toward the future development of Shanghainese (e.g., *It would be sad if Shanghainese disappears in the future*.), while the statements with the lowest scores respectively showed the participants' future orientation (*If I had an opportunity to live in another city that is similar to SH in economic developments, I would choose to stay in SH*.), their attitudes towards immigrant identity (*Immigrants should learn to speak Shanghainese.*), and their beliefs on speaker identity (*If I am talking to another Shanghainese speaker, I will definitely speak the dialect*.).

The comparison between the two clusters suggests that 1) on a general level, the participants hoped to see the HL Shanghainese well preserved and to be passed on generations by generations, and 2) on an individual level, the participants believed that personal development, individual values and identities should be respected and

considered prior to the use and preservation of Shanghainese. Indeed, the task of HL preservation should not solely rely on individual efforts, instead, more policies should be enacted to promote the HL status.

5.3.3. Hypothesis C & D

a more dialect-oriented ***home/ community language environment*** *would lead to a higher HL proficiency* & *a more local* ***demographic background*** *would lead to a higher HL proficiency level*

Since the aspects of home/community language environment and demographic background are highly related, they will be discussed together. The data collected in this study, indicating that the learners' home language environment as well as their demographic background are significantly correlated with their proficiency level in Shanghainese, is consistent with previous studies. For example, Mu and Dooley (2015) found that the demographic variables (*place of birth*, *age of immigration*, and *generation*) each, individually, was significantly associated with participants' self-reported CHL proficiency.

It has to be noted, however, that the non-significant correlation between community language environment and the HL proficiency could be because 1) only two items were employed to collect the frequency of their neighbourhood community and school hosting dialect-related events, which might not be able to reflect the whole picture of the community language environment; and 2) participants' responses on the two items were both very low (respectively on average 1.89 and 1.97 out of 5), which could have resulted in a potential floor effect in the total score in this section. In this case, an association between their community language environment and competence level in Shanghainese would not be able to be detected.

5.3.4. Regression model

The result of the regression analysis showed that the most important predictors of participants' proficiency in Shanghainese, among all the variables, were the *home*

language environment and how many of their *parents* they are *local*. This is consistent with previous study results from Mu & Dooley (2015) which suggested that the *family support* variables were better able to predict participants' self-reported proficiency in CHL than other demographic variables including *the age of immigration, place of birth*, and *generation of immigration.* This has provided insight into what roles a range of different variables plays in predicting the competence in HL. Though it has been noticed that a regression model does not tell the direction of the relation, this pattern fills the literature gap because the majority of the studies on HL only looked at the relationship between HL proficiency and a specific aspect of variables (e.g., attitude, home language environment, demographic background, etc.).

The regression model showed the important role that home language support plays in HL speakers' language development. Hence, parents and grandparents could spend more time speaking Shanghainese at home, thus providing a more supportive home language environment. Even though the children might respond in Mandarin, their listening skills could still be practised, and this is significantly correlated with their overall Shanghainese proficiency, as is discussed in 5.2.1.

5.3.5. Participants with different local parent numbers

This section is also investigated to shed light on whether there continued to be a role for these aforementioned factors when looking at participants with/without local parents separately. For participants whose parents are both local, their *attitudes and beliefs* towards the HL most successfully predicted their competence. For those participants whose parents are both non-local in Shanghai, however, there was no significant predictors in the regression model.

This indicates that for HL speakers who already have a rich language environment and relatively strong support from the family language environment, what matters more is whether they think of the HL as worthy of practising. In this case, the government could invest more in the cultural industry in producing more dialect-relevant programmes that follow the current trend. For example, there could be more TV shows, stand-up comedies, broadcasts, or even pop songs delivered in Shanghainese. In recent

years, there have been successful examples of short videos mediated in Shanghainese that went popular on Chinese social media platforms, e.g., Bilibili (a platform like YouTube) and Douyin (the Chinese version of TikTok). If the dialect plays a more prominent role in pop culture, the young generations would have a more positive attitude towards it.

For those who do not have a supportive language environment as such, even if they have a favourable attitude towards the HL, they would not be able to possess a high competence in the HL given that there is not enough accessible language resource. Regarding this circumstance, the local communities could offer more dialect courses so that these learners would have more opportunities to learn the dialect in a formal setting, rather than merely picking up random expressions sporadically. It remains to be examined whether the kindergarten programmes mentioned in the literature review section could actually take effect.

5.4. Limitations and future research

It should be noted that several limitations exist in this current study. Hence, further research is necessary to explore the findings in this study more thoroughly.

First, the sample group in the study is to some extent flawed. The attrition in the Shanghainese proficiency test might cause the participant group which fulfilled the whole data collection process to be skewed. As is mentioned in Chapter 3, only 143 out of all 197 participants finished the test as well as the questionnaire. In other words, 27.4 % of the participants decided to quit the data collection after the questionnaire. This could be because of two reasons: first, the whole data collection process might be too long for some of the participants. Even though the estimated completion time of 15 to 20 minutes was clearly indicated in the general information section in the link sent to the participants, they might be impatient with going through all the sections. Second, some participants in the study were from younger immigrant generations in Shanghai, which means they have scarce exposure to the dialect Shanghainese in their daily life which in turn results in a relatively low proficiency level. In this case, they might feel

reluctant to finish the Shanghainese proficiency test because of too little knowledge thereof.

Though the high school participants were strongly suggested by their teachers to fulfil the questionnaire and the test, their participation in this study was voluntary. This might have resulted in a biased group with higher motivation and interest in the dialect Shanghainese.

As is mentioned in Chapter 3, the schools were recruited through the author's personal contacts and all the five high schools and the university was among the most elite schools in Shanghai. Hence, the sample group in the study might not be able to represent the general demographics of the Shanghainese teenagers. For example, Shanghai has a strong focus on its policies to attract talents from other provinces and overseas, especially in the field of science and technology. Correspondingly, these top-tier talents pay attention to their children's education and send them to the best high schools in Shanghai. As a result, these high schools would presumably contain a larger proportion of students from a non-local background. It remains to be investigated whether other high schools in Shanghai show a differentiated demographic background and a different pattern of the students' Shanghainese proficiency.

Thus, regarding participant recruitment in future studies, the researchers might consider replicating this study among students from more diverse schools in Shanghai. They could also recruit intact classes or employ random sampling so that the participants could be of different motivation and proficiency levels, making the findings more generalisable.

Second, the research instruments also have certain limitations. The Shanghainese proficiency test and several sections in the questionnaire were designed from scratch by the author of this study. According to the reliability test as well as the statistical analysis, the structure of the questionnaire is reasonably designed, and the participants' performance in the proficiency test was normally distributed without a noticeable floor- or ceiling-effect. However, there is no guarantee that the sections regarding participants' demographics, cultural practise, home and school language environment in the questionnaire could collect all the necessary information. Even for the section investigating participants' attitudes and beliefs towards Shanghainese, which was

developed from a previously validated study, the questions were not standardised. Similarly, the proficiency test might have only tapped into a limited scale of participants' knowledge of Shanghainese. Future research might consider recruiting more participants to further validate the questionnaire and the proficiency test used in designed study. In addition, a mixed-method design could also help further investigate the patterns that emerged from this current study. As is mentioned above in the discussion chapter, adding an interview to the current methodology would be beneficial for looking into why participants have certain linguistic behaviours when it comes to choosing between Shanghainese and Mandarin Chinese in their daily use under all sorts of different scenarios. Furthermore, future research could use separate tests for participants' proficiency in speaking and listening in Shanghainese, so that a more in-depth understanding of their language competence could be achieved.

6. Conclusion

This dissertation aims to deepen the current understanding of speakers' proficiency in Shanghainese, a Chinese Heritage Language. Previous studies on Chinese Heritage Languages generally set their focus outside of China, investigated varied CHL as a whole, looked at the relationship between learners' proficiency level and a specific aspect (e.g., attitudes, length of immigration), and ran a relatively simple statistical analysis. In addition, they generally only employed a simplistic self-report to gauge speakers' proficiency level. Thus, a deeper insight into the current status of a specific Chinese Heritage Language and its promotion and preservation is undoubtfully necessary.

This dissertation fills the gap in the existing literature by looking at high school students in Shanghai who have some extent of proficiency in the language Shanghainese. It investigates their usage patterns of Shanghainese and the relationship between Shanghainese proficiency and varied aspects including speakers' participation in dialect activities, their attitudes toward the dialect, their home and community language environment, as well as their demographic background. By doing so, this study provides insights into the current status of Shanghainese among young people and presents implications for Heritage Language preservation as well as its development. The results of this study show that the participants generally prefer speaking Mandarin Chinese to Shanghainese. Their self-evaluated and tested Shanghainese proficiency are highly correlated. The participants with more engagement in cultural activities, a more positive attitude, a more dialect-oriented home language environment, and a more local demographic background would lead to a higher HL proficiency level.

Future research could build on this current research and engage a more diversified participant background, look at more specified aspects of their language proficiency, and combine qualitative data to develop a more nuanced understanding of speakers' proficiency in Shanghainese, so as to provide a deeper insight into its preservation.

7. References

Baker, C. (1992). *Attitudes and language* (Vol. 83). Multilingual Matters.

Bauman, R., Briggs, C. L., & Briggs, C. S. (2003). *Voices of modernity: Language ideologies and the politics of inequality* (No. 21). Cambridge University Press.

Bell, J. (2013). Language attitudes and language revival/survival. *Journal of Multilingual and Multicultural Development*, *34*(4), 399-410.

Blackledge, A., Creese, A., & Martin, P. (2008). Language ecology and language ideology. In *The Encyclopaedia of Language and Education-Ecology of Language*.

Chappell, H. (1992). Towards a typology of aspect in Sinitic languages. *Zhongguo jingnei yuyan ji yuyanxue: hanyu fangyan*, *1*(1), 67-106.

Chen, P. (1999). *Modern Chinese: History and sociolinguistics*. Cambridge, U.K: Cambridge University Press.

Chow, H. P. (2001). Learning the Chinese language in a multicultural milieu: Factors affecting Chinese-Canadian adolescents' ethnic language school experience. *Alberta journal of educational research*, *47*(4).

Chinen, K., & Tucker, G. R. (2005). Heritage language development: Understanding the roles of ethnic identity and Saturday school participation. *Heritage language journal*, *3*(1), 27-59.

Cohen, J. (1988). *Statistical power analysis for the behavioural sciences* (2nd ed.). New York: Academic Press.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *psychometrika*, *16*(3), 297-334.

Cummins, J. (2005). A proposal for action: Strategies for recognizing heritage language competence as a learning resource within the mainstream classroom. *Modern Language Journal*, 585-592.

Dancey, C. P., & Reidy, J. (2007). *Statistics without maths for psychology*. Pearson education.

Dörnyei, Z. (2007). *Research Methods in Applied Linguistics: Quantitative, Qualitative, and Mixed Methodologies*. Oxford University Press.

Dörnyei, Z. (2019). Towards a better understanding of the L2 Learning Experience, the Cinderella of the L2 Motivational Self System. *Studies in Second Language Learning and Teaching*, *9*(1), 19-30.

Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. Routledge.

Du, X. (2017). Rethink about heritage language learning: A case study of children's Mandarin Chinese learning at a community language school in Ontario, Canada. *Language and Literacy*, *19*(1), 4-20.

Duff, P. A., Liu, Y., & Li, D. (2017). Chinese heritage language learning: Negotiating identities, ideologies, and institutionalization. In *The Routledge Handbook of Heritage Language Education* (pp. 409-422). Routledge.

Duff, P., & Doherty, L. (2019). Learning "Chinese" as a heritage language: Challenges, issues and ways forward. In *The Routledge Handbook of Chinese Applied Linguistics* (pp. 149-164). Routledge.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*. sage.

Fishman, J. A. (2001). 300-plus years of heritage language education in the United States. *Heritage languages in America: Preserving a national resource*, 81-97.

Fishman, J. A. (2006). Acquisition, maintenance, and recovery of heritage languages. *Developing minority language resources: The case of Spanish in California*, *58*.

Fishman, J. A., & Nahirny, V. C. (1966). The ethnic group school and mother tongue maintenance. *Language loyalty in the United States*, 206-252.

Gao, X. (2015). The ideological framing of 'dialect': an analysis of mainland China's state media coverage of 'dialect crisis'(2002–2012). *Journal of Multilingual and Multicultural Development*, *36*(5), 468-482.

Gao, X. (2012). 'Cantonese is not a dialect': Chinese netizens' defence of Cantonese as a regional lingua franca. *Journal of multilingual and multicultural development*, *33*(5), 449-464.

Giles, H., & Byrne, J. L. (1982). An intergroup approach to second language acquisition. *Journal of Multilingual & Multicultural Development*, *3*(1), 17-40.

Gorard, S. (2001). *Quantitative methods in educational research: The role of numbers made easy*. A&C Black.

Guo, L. (2004). The relationship between Putonghua and Chinese dialects. In: Zhou M., Sun H. (eds) *Language Policy in the People's Republic of China: Theory and Practice Since 1949*, 45–53. Boston, MA: Kluwer Academic Publishers.

Han, L. (2020). A systematic literature review: Literacy practices and roles of Chinese parents and Chinese heritage language teachers in Chinese children's literacy learning. *Electronic Thesis and Dissertation Repository*. 6912.

He, A. W. (2006). Toward an Identity Theory of the Development of Chinese as a Heritage Language, *Heritage Language Journal, 4*(1), 1-28.

He, A. W. (2010). The Heart of Heritage: Sociocultural Dimensions of Heritage Language Learning. *Annual Review of Applied Linguistics, 30*, 66-82.

Hornberger, N. H. (1997). Literacy, language maintenance, and linguistic human rights: Three telling cases.

Jiang, B. (2006). Shanghaishi Zhongxaioxue Youeryuan Xuesheng Yuyan Shiyong Diaocha [Survey on language use of students in kindergarten, primary and secondary schools in Shanghai]. *Journal of Chinese Sociolinguistics, 1*.

Jiang, Y. (2021). Language Mixings in Heritage Language Education: A Systematic Review. *International Journal of Language and Literary Studies, 3*(2), 21–36.

Jiang, Y., & Dewaele, J. M. (2019). Language anxiety in Chinese dialects and Putonghua among college students in mainland China: the effects of sociobiographical and linguistic variables. *Journal of Multilingual and Multicultural Development, 40*(4), 289-303.

Kang, H. S., & Kim, I. S. (2012). Perceived and actual competence and ethnic identity in heritage language learning: A case of Korean-American college students. *International Journal of Bilingual Education and Bilingualism, 15*(3), 279-294.

Kozhevnikova, E. (2014). Exposing students to authentic materials as a way to increase students' language proficiency and cultural awareness. *Procedia-Social and Behavioral Sciences, 116*, 4462-4466.

Lam, A. S. (2005). *Language education in China: Policy and experience from 1949: Policy and experience from 1949* (Vol. 1). Hong Kong University Press.

Lao, C. (2004). Parents' attitudes toward Chinese–English bilingual education and Chinese-language use. *Bilingual research journal, 28*(1), 99-121.

Li, D. (2006). Chinese as a lingua franca in Greater China. *Annual Review of Applied Linguistics, 26*, 149-176.

Li, G. (2006). Biliteracy and trilingual practices in the home context: Case studies of Chinese-Canadian children. *Journal of Early Childhood Literacy*, *6*(3), 355-381.

Liang, S. (2015). Language Attitudes and Identities in Multilingual China: A Linguistic Ethnography. London: Springer.

Llamas, C., & Watt, D. (2014). Scottish, English, British?: innovations in attitude measurement. *Language and Linguistics Compass*, *8*(11), 610-617.

Ma, W., & Winke, P. (2019). Self-assessment: How reliable is it in assessing oral proficiency over time?. *Foreign Language Annals*, *52*(1), 66-86.

Maloof, V. M., Rubin, D. L., & Miller, A. N. (2006). Cultural competence and identity in cross-cultural adaptation: The role of a Vietnamese heritage language school. *International Journal of Bilingual Education and Bilingualism*, *9*(2), 255-273.

Montrul, S. (2010). Current Issues in Heritage Language Acquisition. *Annual Review of Applied Linguistics*, 30, 3-23.

Mori, Y., & Calder, T. M. (2015). The role of motivation and learner variables in L1 and L2 vocabulary development in Japanese heritage language speakers in the United States. *Foreign Language Annals*, *48*(4), 730-754.

Mu, G. M., & Dooley, K. (2015). Coming into an inheritance: Family support and Chinese heritage language learning. *International Journal of Bilingual Education and Bilingualism*, *18*(4), 501-515.

Oh, J. S., & Fuligni, A. J. (2010). The role of heritage language development in the ethnic identity and family relationships of adolescents from immigrant backgrounds. *Social development*, *19*(1), 202-220.

Paulhus, D. L., & Vazire, S. (2007). The self-report method. *Handbook of research methods in personality psychology*, *1*(2007), 224-239.

Pu, C. (2008). *Chinese American children's bilingual and biliteracy development in heritage language and public schools*. The University of Texas at San Antonio.

Robins, R. W., & John, O. P. (1997). The quest for self-insight: Theory and research on accuracy and bias in self-perception. In *Handbook of personality psychology* (pp. 649-679). Academic Press.

Ramsey, S. R. (1989). *The languages of China*. Princeton University Press.

Schumann, J. H. (1986). Research on the acculturation model for second language acquisition. *Journal of multilingual & multicultural development, 7*(5), 379-392.

Schilling-Estes, N., & Wolfram, W. (1999). Alternative Models of Dialect Death: Dissipation vs. Concentration. *Language, 75*(3), 486–521.

Schwartz, M., & Verschik, A. (Eds.). (2013). *Successful family language policy: Parents, children and educators in interaction* (Vol. 7). Springer Science & Business Media.

Shanghai Municipal Education Commission. (2016). *Notice on the Issuance of the 13th Five-Year Plan for the Reform and Development of Language in Shanghai*. http://edu.sh.gov.cn/xxgk2_zhzw_ghjh_01/20201015/v2-0015-gw_301132016007.html

Shen, Q. (2016). Saving Shanghai dialect: A case for bottom-up language planning in China. *The Asia-Pacific Education Researcher, 25*(5), 713-722.

Shi, J. (2012). Xuzhe lianming Changyi kexue baohu shanghaihua (Scholars proposing to protect Shanghai dialect in a scientific way). *Oriental Morning Post,* p. B3.

Smith, S. A., & Li, Z. (2020). Closing the enjoyment gap: heritage language maintenance motivation and reading attitudes among Chinese-American children. *International Journal of Bilingual Education and Bilingualism*, 1-18.

Smith, S. A., & Briggs Baffoe-Djan, J. (2022). It takes two to make a thing go write: self-reported Spanish reading and writing skill and relationships with digit span and stroop task performance among Spanish-English bilinguals. *International Journal of Bilingual Education and Bilingualism, 25*(1), 272-286.

Spolsky, B. (2012). Family language policy–the critical domain. *Journal of Multilingual and Multicultural Development, 33*(1), 3-11.

Standing Committee of the National People's Congress. (2000). *Law of the People's Republic of China on the Standard Spoken and Written Chinese Language*. Lawinfochina.com. http://www.lawinfochina.com/display.aspx?lib=law&id=6233&CGid=

Steering Group Office for Survey of Language Situation in China. (2006) *Zhongguo Yuyan Wenzi Shiyong Qingkuang Diaocha Ziliao* [Information on the survey on the use of Chinese language and script]. Beijing: Language & Culture Press.

Sun, X., Jiang, B., Wang, Y., & Qiao, Y. (2007). Shanghai Xuesheng Putonghua he Shanghaihua shiyong qingkuang diaocha [Survey on the use of Mandarin and Shanghainese among students in Shanghai]. *Yangtze River Academic*, *3*, 1-10.

Tse, L. (2001). Resisting and reversing language shift: Heritage-language resilience among US native biliterates. *Harvard educational review*, *71*(4), 676-709.

Tulloch, S. (2006). Preserving dialects of an endangered language. *Current Issues in Language Planning*, *7*(2-3), 269-286.

Van Deusen-Scholl, N. (2003). Toward a definition of heritage language: Sociopolitical and pedagogical considerations. *Journal of language, identity, and education*, *2*(3), 211-230.

Wan, K. F. (2017). *Language Attitudes of University Students and Young Professionals in Guangzhou, China* (Doctoral dissertation, The Ohio State University).

Wang, W. S. (1973). The Chinese language. *Scientific American*, *228*(2), 50-63.

Wang, X. (2017). Exploring the role of attitudes in new dialect formation in Hohhot, China.

Wong, K. F., & Xiao, Y. (2010). Diversity and difference: Identity issues of Chinese heritage language learners from dialect backgrounds. *Heritage Language Journal*, *7*(2), 314-348.

Wu, M. H., & Leung, G. (2020). 'It's not my Chinese': a teacher and her students disrupting and dismantling conventional notions of 'Chinese' through translanguaging in a heritage language classroom. *International Journal of Bilingual Education and Bilingualism*, 1-14.

Xia, T., & Shen, Q. (2019). Individual agency in language acquisition planning for multi-dialectism: A Shanghai story. *International Journal of Bilingual Education and Bilingualism*, *22*(3), 338-351.

Xiao, L. (1998). Chinese language maintenance in Winnipeg. *Canadian Ethnic Studies*, *30*, 86–96.

Xu, B. & Tang, Z. (1991). *Shanghai Fangyan Cihui*. [Shanghai Dialect Vocabulary]. Shanghai: Shanghai Education Publishing House.

Xu, F., Gu, J., & Shen, Q. (2021). Silencing Shanghai: language and identity in urban China.

Xu, H., & Zhou, Z. (2016). On language use and language attitude of the youths and adolescents in Guangzhou. *Applied Linguistics* (Chinese Journal), *99*(3), 20-29.

Yu, W., & Yang, J. (2016). Diachronic changes of dialect use and dialect competence among local youths in Shanghai over the past fifteen years and influencing factors. *Applied Linguistics* (Chinese Journal), *4*(1), 26–34.

Zhao, H. (2018). *Language variation and social identity in Beijing* (Doctoral dissertation, Queen Mary University of London).

Zhang, D., & Koda, K. (2011). Home literacy environment and word knowledge development: A study of young learners of Chinese as a heritage language. *Bilingual Research Journal, 34*(1), 4-18.

Zhou, H. (2012). 2012 *Shanghai Zhongxiao Xuesheng Chengzhang Qingkuang Zuixin Diaocha Baogao* [2012 demographic survey on primary and middle school students in Shanghai]. Shanghai: Shanghai Social Science Academy.

Zhou, M., & Li, X. Y. (2003). Ethnic language schools and the development of supplementary education in the immigrant Chinese community in the United States. *New Directions for Youth Development, 2003*(100), 57-73.

8. Appendices

Appendix 1: Ethical Approval

CUREC approval email

From: _____

Sent: Saturday 26th March, 2022 3:14 PM

To: _____

Cc: Student CUREC

Subject: Curec Approval CIA-22HT-023

Dear _____,

Your application for 'Factors predicting the Heritage Language proficiency in Shanghainese' has been considered on behalf of the DREC in accordance with the procedures laid down by the University for ethical approval of all research involving human participants.

I am pleased to inform you that, based on the information provided to DREC, the proposed research has been judged as meeting appropriate ethical standards, and accordingly, approval has been granted. I would like to inform you that you will be required to update us on any amendments to your study should you need to change your research methods and may need to complete a data protection impact assessment if you use online platforms to conduct and record interviews. There is an application for this, and it is kept separate from the ethics committee. Please see the link below for more information.
https://researchsupport.admin.ox.ac.uk/policy/data/checklist

Please continue to follow all current guidance issued by CUREC during the pandemic, notably COVID-19: CUREC guidance on research involving human participants, https://researchsupport.admin.ox.ac.uk/governance/ethics/coronavirus

If needed, please follow the guidance on online data collection and research methods issued by the University,
(1) https://researchsupport.admin.ox.ac.uk/covid-19/data#collapse2299911
(2) https://infosec.web.ox.ac.uk/article/guidelines-for-using-zoom

If relevant, please also check the CUREC website for their best practice research guides, https://researchsupport.admin.ox.ac.uk/governance/ethics/resources/bpg

Yours sincerely,

_____

Appendix 2: Questionnaire

a. Generic consent

Consent

Thank you for agreeing to take part in this experiment!

Before we continue, we need your consent to the following:

1. I consent to performing the task online
2. I understand and consent to my responses are being recorded and stored securely in a database
3. I understand and consent to my responses may be used anonymously for secondary research in the future

同意书

非常感谢你参与本研究！

在开始前，我们需要获得你的许可：

1. 我同意在线上参与本研究。
2. 我理解并同意我的回答被记录，并被存储在安全的数据库中。
3. 我理解并同意我的数据被匿名使用于后续研究中。

☐ **I consent to items 1-3 above 我同意以上1-3条陈述。**

b. Demographics

Here are a set of questions collecting generic **demographic information**.

This section should take you approximately **2 minutes** to complete.

本部分将收集一些**基本个人信息。**

本部分预计耗时**两分钟。**

Gender

你的性别是

- ◯ Female 女性
- ◯ Male 男性
- ◯ Non-binary 非二元性别

You current level of study is:

你目前就读于：

- ◯ High school 高中
- ◯ Undergraduate 本科

Your father's hometown is

你父亲的家乡是

- ◯ Shanghai 上海
- ◯ Not Shanghai 非上海（请说明具体地区, e.g., 江苏，天津）

Your mother's hometown is

你母亲的家乡是

- ◯ Shanghai 上海
- ◯ Not Shanghai 非上海（请说明具体地区, e.g., 江苏，天津）

You have been living in Shanghai for

你在上海生活已经

- ◯ Always 一直在此
- ◯ More than 10 years 超过10年
- ◯ 5-10 years 5-10年
- ◯ 3-5 years 3-5年
- ◯ Less than 3 years 少于3年

You are the __st generation in SH (if applicable)

你是在上海生活的第___代移民

Please Select...

How old are you?

你的年龄是

Please Select...

c. Language usage patterns

Here are a set of questions about your and your family's **language usage patterns**.

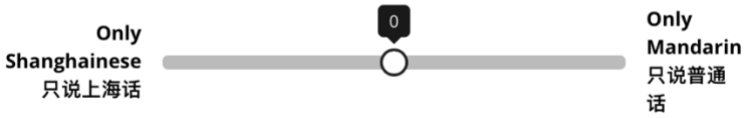
It should take you approximately **2 minutes** to finish this section.

本部分将会调查你以及你家人的**语言使用模式。**

本部分预计耗时 **2 分钟。**

What proportion of each language do **you** speak ***at home***?

在家时，你的语言使用模式是：



Please indicate any other language you speak, if applicable, as well as the proportion of the time that you speak each language:

e.g., *Shanghainese* (10), *Mandarin* (80), *Suzhou dialect* (10)

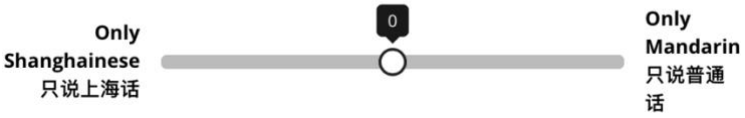
除了上海话和普通话，你在家还会使用**其他语言**吗?

若有，请指出该语言，及你使用它们的相对频率（括号内数值之和为100）。

例如：上海话（50），普通话（30），苏州话（20）

What proportion of each language do **your grandparents** speak ***at home*** when talking to **you**?

在家时，你祖父母与你交流时的的语言使用模式是：



Please indicate any other language they speak, if applicable, as well as the proportion of the time that they speak each language:

e.g., *Shanghainese* (10), *Mandarin* (80), *Suzhou dialect* (10)

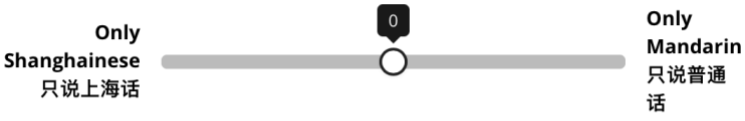
除了上海话和普通话，你祖父母在家还会使用**其他语言**吗?

若有，请指出该语言，及你祖父母使用它们的相对频率（括号内数值之和为100）。

例如：上海话（50），普通话（30），苏州话（20）

What proportion of each language do **your grandparents** speak ***at home*** when talking to **each other** or **other adult family members**?

在家时，你祖父母与彼此或其他成年家庭成员交流时的语言使用模式是：



Please indicate any other language they speak, if applicable, as well as the proportion of the time that they speak each language:

e.g., *Shanghainese* (10), *Mandarin* (80), *Suzhou dialect* (10)

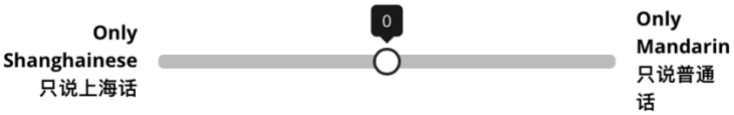
除了上海话和普通话，你祖父母在家还会使用**其他语言**吗?

若有，请指出该语言，及你祖父母使用它们的相对频率（括号内数值之和为100）。

例如：上海话（50），普通话（30），苏州话（20）

What proportion of each language do **you** speak ***at school*** **with peers** during daily communication?

在学校时，你和同学（在日常会话中）的语言使用模式是：



Please indicate any other language, if applicable, as well as the proportion of all the languages:

e.g., *Shanghainese* (10), *Mandarin* (80), *Suzhou dialect* (10)

除了上海话和普通话，你和同学在**学校**（日常会话中）还会使用**其他语言**吗?

若有，请指出该语言，及你使用它们的相对频率（括号内数值之和为100）。

例如：上海话（50），普通话（30），苏州话（20）

d. Attitudes

Here are 22 statements about the perception of Shanghainese that may or may not apply to you.

Please drag the slider to rate each statement to indicate the extent to which you **agree** or **disagree** with that statement.

For example, "-50" at the left end of the bar would indicate **completely disagree**, while "50" on the right end of the bar would suggest **completely agree**. If you don't have a clear opinion or you are not sure, you can leave it at the middle point.

There are no right or wrong answers.

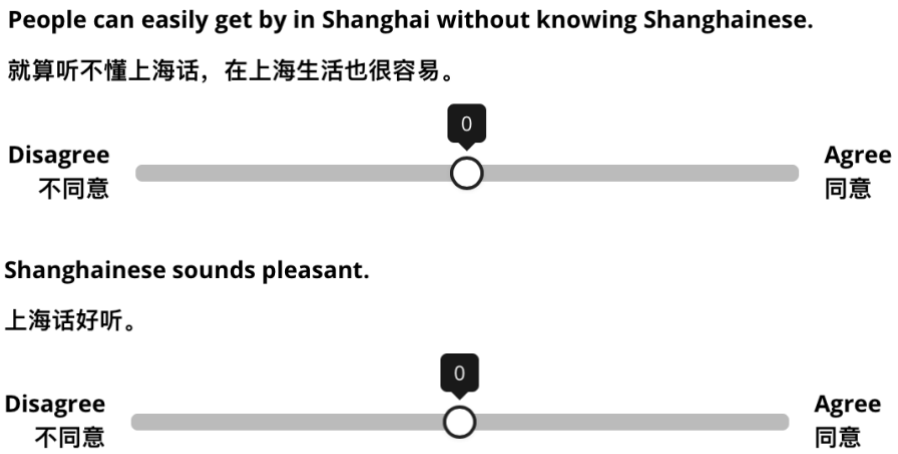
This section should take you approximately **3 minutes** to complete.

在本部分，你将会看到 22 句关于上海话的陈述。

它们会在不同程度上符合你的认知和判断。 请你通过**拖拽滑块**的方式，指出你对各个表述的**同意程度**。

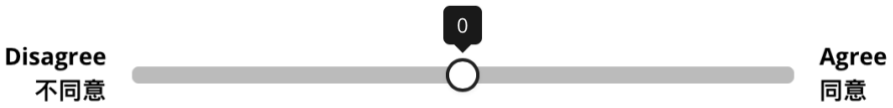
滑块最左端的"-50"表示**非常不同意**，最右端的"50"表示**非常同意**。若你没有明确的意见或不确定，可以把滑块留在正中间的"0"处。

题目没有**正确**或**错误**的答案，表达你的观点即可。
本部分内容预计耗时 **3 分钟**。



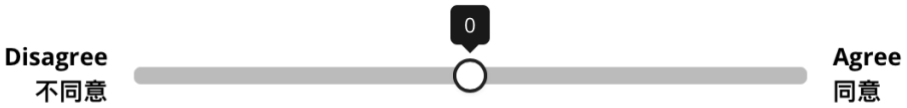
Immigrants should learn to speak Shanghainese.

移居到上海的人应该学说上海话。



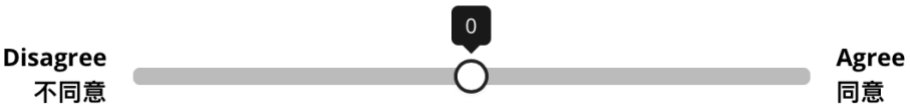
It would be sad if Shanghainese disappears in the future.

如果有一天上海话彻底消失了，我会感到悲哀。



If my children spoke Shanghainese, I would oppose it.

如果我的子女也说上海话，我会反对。



I love living in SH.

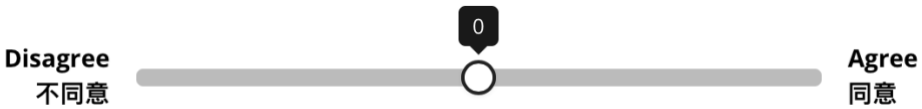
我喜欢在上海生活。

If I am talking to another Shanghainese speaker, I will definitely speak the dialect.

如果和一个会说上海话的人说话，我一定会用上海话。

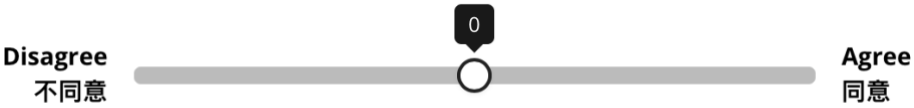
Shanghai dialect is witty and humourous.

上海话听上去很幽默有趣。



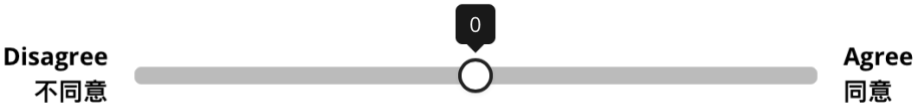
Local Shanghainese people can represent the Shanghai culture best.

上海本地人最能代表上海文化。



Immigrants in Shanghai also count as local.

上海移民也算是本地人。

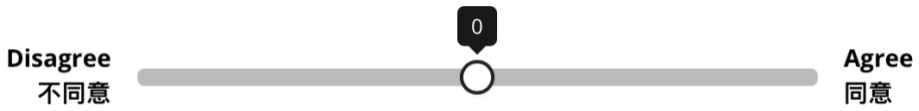


I'm proud of being a Shanghainese speaker.

我因会说上海话而感到自豪。

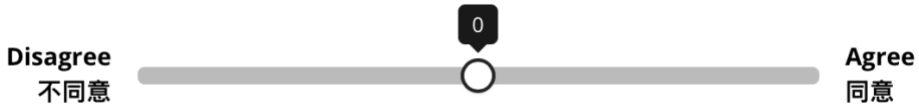
I can recognise the SH accent when I listen to other people speak Mandarin.

我可以听出别人普通话里的上海口音。



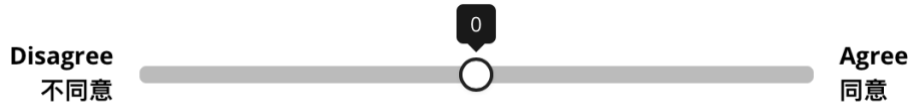
I like my SH accent when I speak Mandarin.

我喜欢我的上海口音。



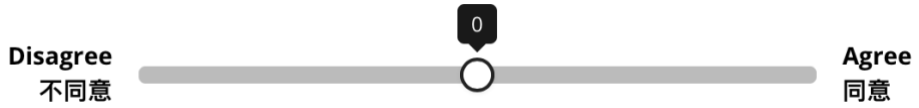
I have an SH accent when I speak Mandarin.

我说普通话有上海口音。



If I had an opportunity to live in another city that is similar to SH in economic developments, I would choose to stay in SH.

如果我有机会去另一个和上海发展水平相近的城市生活，我会选择留在上海。

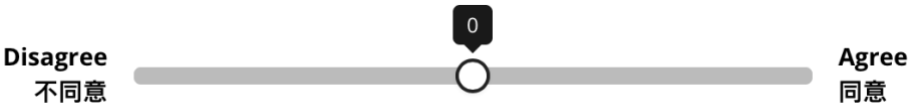


I hope my descendants can live and work in SH in the future.

我希望我的后代将来也在上海生活和工作。

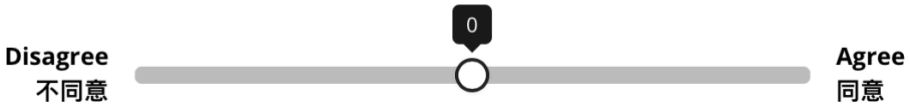
There should be some comedies in Shanghainese on TV and broadcasting.

电视或广播中应该有上海话喜剧、小品等节目。



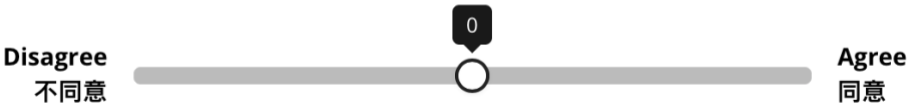
There should be some news reports in Shanghainese on TV and broadcasting.

电视或广播中应该有上海话新闻。



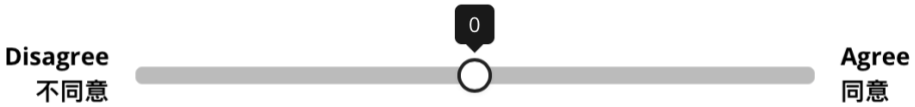
Shanghainese could represent the local culture.

上海话能代表上海文化。



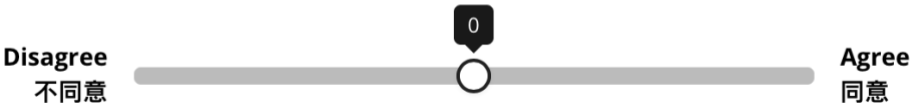
I hope to see my children and grandchildren speak Shanghainese.

我希望看到我的子辈和孙辈也讲上海话。



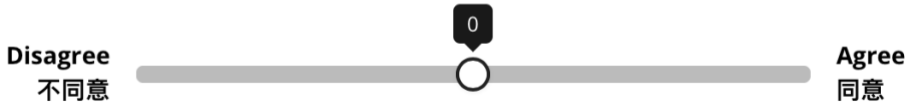
Shanghainese is vulgar.

上海话很粗俗。



I would love to keep living in SH and work here in the future.

我会选择继续在上海生活和工作。



e. Cultural activities

In this section, you will be asked several questions regarding your engagement with Shanghainese-related media and cultural activities.

They should take you approximately **2 minutes** to answer.

本部分将收集与上海话相关文化活动的信息。

该部分大约耗时**两分钟**。

I listen to **broadcasts in Shanghainese**:

我听**上海话广播**的频率是：

never 从不	seldom 很少	sometimes 偶尔	often 经常	always 总是

I watch **local comedies/talk-shows (e.g., stand-up comedies)**:

我看**本地喜剧/脱口秀**的频率是（包括线上和线下）：（e.g., 沪语脱口秀；papi的上海话视频）

never 从不	seldom 很少	sometimes 偶尔	often 经常	always 总是

I watch **Hu-opera**:

我看**沪剧**的频率是（包括线上和线下）：

never 从不	seldom 很少	sometimes 偶尔	often 经常	always 总是

I watch **local TV shows in Shanghainese**:

我看**上海话电视节目**的频率是：(e.g., 老娘舅, 上海台美食节目)

never 从不	seldom 很少	sometimes 偶尔	often 经常	always 总是

My community hosts events mediated in / promoting Shanghainese:

我居住的街道/社区组织沪语文化活动的频率是：

never 从不	seldom 很少	sometimes 偶尔	often 经常	always 总是	I don't know 我不清楚

I go to these events:

我参加这些活动的频率是：

never 从不	seldom 很少	sometimes 偶尔	often 经常	always 总是

For all the schools I attended, events mediated in / promoting Shanghainese were hosted:

我的学校（所有阶段，包括中小学和大学）组织沪语文化活动的频率是：

never 从不	seldom 很少	sometimes 有时	often 经常	always 总是	I don't know 我不清楚

I go to these events:

我参加这些活动的频率是：

never 从不	seldom 很少	sometimes 有时	often 经常	always 总是

Overall, I engage in activities related to Shanghainese:

总的来说，我参加沪语相关活动的频率是：

never 从不	seldom 很少	sometimes 偶尔	often 经常	always 总是

f. Self-perceived proficiency

Listening proficiency: When I **listen to** Shanghainese, I find it

我觉得**听**上海话

Very hard 很难 — 0 — **Very easy 很容易**

Speaking proficiency: When I **speak** Shanghainese, I find it

我觉得**说**上海话

Very hard 很难 — 0 — **Very easy 很容易**

Appendix 3: Shanghainese proficiency test

Instructions

In this section, you will be shown a set of questions.

Answer them using the box provided.

To guarantee the validity, please complete this test **independently**, thank you.

The audio prompt could be played up to **two** times.

This section should take approximately **5 minutes**.

在本部分，你将会回答一些测试**上海话水平**的题目。请使用屏幕下方的选项回答。

为保证本测试的有效性，请在没有任何帮助的情况下**独立完成**，谢谢。

题目的语音提示最多可以播放**两遍**。

本部分预计**5分钟**。

水门汀 指 ▶ Play

A. 大桥
B. 水泥
C. 瓦片
D. 拱门

死蟹一只指： ▶ Play

A. 事情糟糕，不可挽回
B. 形容人精神状态不好，浑浑噩噩
C. 天气太热，昏了头
D. 境况很危险，性命攸关

搞七廿三指： ▶ Play

A. 眼花缭乱
B. 手脚不灵活
C. 胡搅蛮缠
D. 擅长手艺活

开大兴指： ▶ Play

A. 说假话，做浮事
B. 气量大，不计较
C. 比喻可尽情享受
D. 神通广大

煨灶猫指： ▶ Play

A. 比喻每个人做事有自己的余地
B. 比喻精神萎靡不振的人
C. 比喻舒服的生活状态
D. 形容人气量小

舍姆娘指：

▶ Play

A. 临时抱佛脚的人
B. 喝酒后胡乱说话的人
C. 坐月子的产妇
D. 负责人情往来的人

上只角指：

▶ Play

A. 狠下心来认真做事
B. 形容人优秀拔尖
C. 在城市里，居住条件好，生活水平高的地区
D. 一碰就坏的东西

一百零一指：

▶ Play

A. 唯一的一个
B. 最好的一个
C. 形容事情复杂繁琐
D. 形容人挑剔

水荡 指：

▶ Play

A. 小池子
B. 积水处
C. 小河滩
D. 河床

好户头指：

▶ Play

A. 指家境殷实
B. 老实软弱可欺者
C. 两面都讨好，不得罪
D. 人无十全十美

困扁头指：

▶ Play

A. 起早摸黑地工作
B. 比喻旅途辛苦
C. 晚上睡得太久导致白天没精神
D. 白日做梦，异想天开

乌苏天 指：

▶ Play

A. 干燥晴朗的天气
B. 忽冷忽热易感冒的天气
C. 十月里和暖似春的天气
D. 潮湿/闷热难受的天气

以下选项**不正确**的是：

A. 通关手：指手大而有力 ▶ Play

B. 黄鱼肚皮： 腿肚子 ▶ Play

C. 脚馒头：膝盖 ▶ Play

D. 小寒豆：豌豆 ▶ Play

以下选项**不正确**的是：

A. 文旦：柚子 ▶ Play

B. 黄芽菜：大白菜 ▶ Play

C. 地栗：荸荠 ▶ Play

D. 落脚菜：便宜的菜 ▶ Play

以下选项**不正确**的是：

A. 尽根牙：智齿 ▶ Play

B. 盘牙：磨牙 ▶ Play

C. 抠牙齿：向外凸出的牙齿 ▶ Play

D. 抢牙：牙齿斜出，遮住它牙 ▶ Play

表示**"总共"**的词汇有：

A. 夯拨郎当 ▶ Play

B. 一天世界 ▶ Play

C. 一塌刮子 ▶ Play

儿歌：

笃笃笃，卖__粥，三斤__四斤__。

A. 白，栗子，壳
B. 糖，蒲桃，壳
C. 白，栗子，肉
D. 糖，白桃，皮

训斥孩子时，上海家长说**我揍你**的沪语表述是：

A. 酸一记
B. 甜一记
C. 辣一记
D. 咸一记

上海话中，一下哪个字的发音和其他选项**不同**？（不考虑音调）

A. 糖
B. 汤
C. 挡
D. 烫

上海话中，以下哪个字的发音和其他选项**不同**?

A. 鹿
B. 六
C. 陆
D. 露

Thank you for taking part!

非常感谢你的参与！ 祝一切顺利 好运连连～

Appendix 4: Attitudes questionnaire reference from previous study (Wang, 2017)

呼市与呼市方言
Hohhot and Dialects in Hohhot

A. 基本信息
A. Basic information

性别 gender: 男 male /女 female

年龄 age: _____ 职业 occupation: _____

教育程度 level of education: _____

您父母的原籍 parents' hometown: 父亲 father _____ 母亲 mother_____

父母何时迁入呼市: _____
When did your parents move into this city:

您是否经常离开本地（半年以上）？ a. 是 b. 不是
Do you often go out of town (more than half a year)? a. Yes b. No

B. 语言能力
B. Language ability

请将您的普通话和此地话的听说能力在横线上标示出来。
Please indicate how well or badly you can speak Pǔtōnghuà and Jìn dialect by marking on the line.
例如 For example: **很好** Well ——┼—— **很差** Badly

1. 我的普通话说得 很好 Well ——— 很差 Badly
I can speak Pǔtōnghuà

2. 我听人说普通话 能听懂 Well ——— 听不懂 Badly
I can understand Pǔtōnghuà

3. 我的此地话说得 很好 Well ——— 很差 Badly
I can speak Jìn dialect

4. 我听人说此地话 能听懂 Well ——— 听不懂 Badly
I can understand Jìn dialect

C. 下面列出了一些人对呼市或呼市人的想法或态度，您怎么看这些想法？同意还是不同意他们的说法？他们的陈述是否符合您的情况？
C. Here are some statements about Hohhot or Hohhot people. What do you think of

them? Do you agree or disagree?

请将你对这些陈述同意（或不同意）的程度标示在该陈述下面的横线上，您标出的位置越靠近“同意”，表示您越同意上面的陈述，反之，标出的位置越接近“不同意”，代表您越反对上面的陈述。如果没有明确的意见或无法确定，可以划在横线的正中间。
Please indicate your level of agreement (or disagreement) with the following statements by marking on the straight line below. The closer you mark to “I agree”, the more you agree with the statement above it; and similarly, the closer you mark to “I disagree”, the more you disagree with the statement. If you don’t have a clear opinion or you are not sure, you can mark at the middle point of the line.

例： 奶豆腐很好吃。
For example: Cheese tastes nice.

同意 I agree ——————┼—————— **不同意** I disagree

1. 此地话好听。
Jìn dialect sounds pleasant.

同意 I agree —————————— **不同意** I disagree

2. 学说此地话很有意思。
Learning to speak Jìn dialect is very interesting.

同意 I agree —————————— **不同意** I disagree

3. 在呼市听不懂此地话也不影响日常生活。
People can easily get by in Hohhot without knowing any Jìn dialect.

同意 I agree —————————— **不同意** I disagree

4. 如果我在旧城找到一份工作，我会搬到旧城去住。
If I find a job in the Old Town, I’ll move there.

同意 I agree —————————— **不同意** I disagree

5. 我以自己是支边移民的后代而自豪。
I am proud of being a descendant of immigrants.

同意 I agree —————————— **不同意** I disagree

6. 如果聊天时发现对方也是支边移民的后代，我会感到很亲切。
I’ll feel close and intimate if I know someone is also a descendant of immigrants.

同意 I agree —————————— **不同意** I disagree

7. 旧城人很土。
Old Town people are vulgar.

同意 I agree ——————————— **不同意** I disagree

8. 此地话能代表呼市的文化。
The local Jìn dialect can represent Hohhot culture.

同意 I agree ——————————— **不同意** I disagree

9. 当有人问我是哪里人的时候，我会提到自己的父母是移民到呼市的。
When people ask me where I am from, I often mention that my parents are immigrants.

同意 I agree ——————————— **不同意** I disagree

10. 我喜欢在呼市生活。
I love living in Hohhot.

同意 I agree ——————————— **不同意** I disagree

11. 除了偶尔的购物、旅游，我很少去旧城。
Except for occasional shopping or touring, I seldom go to the Old Town.

同意 I agree ——————————— **不同意** I disagree

12. 如果我有机会去北京上海等大城市生活，我会选择去大城市。
If I have an opportunity to live in bigger cities like Beijing or Shanghai, I will choose to go.

同意 I agree ——————————— **不同意** I disagree

13. 此地话很土，很粗俗。
Jìn dialect is vulgar.

同意 I agree ——————————— **不同意** I disagree

14. 学说此地话很有用。
Learning to speak Jìn dialect is very helpful.

同意 I agree ——————————— **不同意** I disagree

15. 我更愿意与其他的支边移民后代交朋友。
I prefer to make friends with other descendants of immigrants.

同意 I agree ——————————— **不同意** I disagree

16. 如果我孩子的男/女朋友是旧城人，我会反对他们交往。
If my child is seeing or dating someone from the Old Town, I would oppose it.

同意 I agree ——————————— **不同意** I disagree

17. 此地话很幽默。
Jìn dialect is humorous.

同意 I agree ——————————— **不同意** I disagree

18. 我今后会一直在呼市生活和工作。
I will live and work in Hohhot in the future.

同意 I agree ——————————— **不同意** I disagree

19. 如果有一天此地话彻底消失了，我会感到悲哀。
It would be sad if Jìn dialect disappeared in the future.

同意 I agree ——————————— **不同意** I disagree

20. 电视或广播中应该有用此地话播报新闻的节目。
There should be some news reports in Jìn dialect on TV and broadcasting.

同意 I agree ——————————— **不同意** I disagree

21. 电视或广播中应该有用此地话表演喜剧、小品等节目。
There should be some comedies in Jìn dialect on TV and broadcasting.

同意 I agree ——————————— **不同意** I disagree

22. 我希望我的后代将来也在呼市生活和工作。
I hope my descendants can live and work in Hohhot in the future.

同意 I agree ——————————— **不同意** I disagree

23. 如果我有机会去另一个和呼市发展水平相近的城市生活，我会选择留在呼市。
If I have an opportunity to live in another city that is similar to Hohhot in economic developments, I will stay in Hohhot.

同意 I agree ——————————— **不同意** I disagree

Appendix 5: Example information sheet and opt-out form

Factors predicting the Heritage Language proficiency in Shanghainese

INFORMATION SHEET FOR PARENTS / GUARDIANS

Ethics Approval Reference: CIA-22HT-023

In partnership with researchers at the University of Oxford, your child's school has agreed to take part in a study investigating factors that predict Shanghainese teenagers' proficiency in the heritage language Shanghainese. We would like to invite your child to be part of this study. We very much hope you would like your child to take part, but before you decide, it is important that you understand why the study is being done and what it will involve.

What are we trying to find out?

The research aims to investigate Shanghainese teenagers' proficiency in Shanghainese, a Chinese Heritage Language. A mixed-method approach combining quantitative questionnaire and qualitative interview would be adopted. The target participant group is high school and university students in Shanghai who have some extent of proficiency in the language Shanghainese. We will look at the usage pattern of Shanghainese, the relationship between Shanghainese proficiency and speakers' participation in dialect activities at home and at school, motivations and attitudes, as well as some demographic factors. By doing so, this study will provide insights into the development and current status of Shanghainese among young people in Shanghai and could serve the purpose of Heritage Language preservation.

More information about the project can be obtained by contacting the research team (contact details overleaf).

Why has my child been invited to take part?

We are inviting your child to take part because they are a young person, aged between 15 and 18 years, attending Shanghai Datong High School.
We are inviting approximately 150 young people to take part.

Does my child have to take part?

No. You can ask questions about the study before deciding whether or not to allow your child to participate. If you do agree to participation, you may withdraw your child from the study

at any time, without giving a reason and without penalty, by advising the researchers of this decision.

What will happen if my child takes part?

If you agree your child to take part, your child will be asked to complete a questionnaire online which would take approximately 10 minutes. Your child may also be asked to participate in a follow-up interview online which would take approximately 20 minutes. The tasks are student-friendly and will not contain any sort of risks. The data collected will be anonymous.

What are the advantages / disadvantages of taking part?

There's no intended benefit for the participants directly. There is no risk for your child in partaking in this study.

What happens to the data provided? [1]

The information you or your child provide during the study is the research data. Any research data from which you or your child can be identified (date of birth, place of birth, audio recording etc.), is known as personal data. Personal data will be stored in a password-protected folder in Nexus 265 OneDrive for Business on a password-protected laptop as long as it is needed to conduct the research.

Opt-out forms will be retained by the school for the duration of the study, and for as long as the school determines appropriate after research activities have concluded at the school.

Other research data will be stored for a minimum of 3 years after publication or public release of the work of the research. The researcher and supervisor will have access to the research data. Fully anonymised and summarised data may be shared with other researchers beyond the university, via publications and via online forums such as the Open science framework, and it will not be possible to identify any particular individual within the dataset.

Will the research be published?

The research will be written up as a student's thesis. On successful submission of the thesis, it may be deposited both in print and online in the University archives to facilitate its use in future research. If so, the thesis will be openly accessible.

What will happen to any samples taken from my child?

We will use your child's samples for data analysis using e.g. SPSS, a statistical analysis software. When reporting results, the data will be anonymised and no personal details would be revealed.

[1] Please refer to CUREC's Best Practice Guidance on Data Collection and Management (BPG 09)

Who is conducting this research?

The research project is organised by _____ of Oxford University, who is a MSc student in Applied Linguistics and Second Language Acquisition, Department of Education, University of Oxford. This study has been reviewed by, and received ethics clearance through, the University of Oxford's Central University Research Ethics Committee, CIA-22HT-023.

What if there is a problem?

If you have a concern about any aspect of this study, please contact _____ (_____@education.ox.ac.uk) or _____ (_____@education.ox.ac.uk), and we will do our best to answer your query. We will acknowledge your concern within 10 working days and give you an indication of how it will be dealt with. If you remain unhappy or wish to make a formal complaint, please contact the Chair of the Research Ethics Committee at the University of Oxford who will seek to resolve the matter as soon as possible:

Chair, **Social Sciences & Humanities Inter-Divisional Research Ethics Committee**; Email: ethics@socsci.ox.ac.uk; Address: Research Services, University of Oxford, Wellington Square, Oxford OX1 2JD

Data Protection

The University of Oxford is the data controller with respect to your child's personal data and, as such, will determine how your child's personal data is used in the study.

The University will process your child's personal data for the purpose of the research outlined above. Research is a task that we perform in the public interest.

Further information about your rights with respect to your child's personal data is available from https://compliance.web.ox.ac.uk/individual-rights.

What should I do next?

Please fill in the enclosed form and return it to your child's class teacher if you would not like your child to take part in this study. Please remember that you may withdraw your child at any time, without penalty and without giving a reason, by notifying the researcher.

If you would like to discuss the research with someone beforehand (or if you have questions afterwards), please contact:

_____
Department of Education
15 Norham Gardens, Oxford
Tel: +44 _____
Email: _____@education.ox.ac.uk

Factors predicting the Heritage Language proficiency in Shanghainese

关于影响上海话水平的因素的调查

INFORMATION SHEET FOR PARENTS / GUARDIANS (Chinese version)

研究信息(家长阅读版)

Ethics Approval Reference: CIA-22HT-023

家长您好!

我是_____，目前正在攻读牛津大学应用语言学专业的硕士学位。 您的孩子正在参与的研究是我的毕业论文项目。本研究已通过了牛津大学研究道德委员会的审核(编号:CIA-22HT-023)。

本研究探究**上海青少年的上海话水平**及对**其产生影响的因素**。

本研究的**参与者**为上海本地对上海话有一定掌握程度的高中生及本科生。

本研究将关注上海话**使用情况**，并探索**上海话水平**与使用者**情感态度**、**家庭语言环境**、**学校语言环境**、**方言文化活动参与情况**、**其他相关个人信息**等因素的关联。

本研究意图关注上海话在青少年中的现状，从而保护并发展该方言。

您的孩子在本研究中提供的信息完全**匿名**。在本研究的任何环节，我们都不会询问或记录您孩子的姓名。此外，我们将会把信息存储在牛津大学的安全数据库中妥善保存。

参与本研究不会对您的孩子造成任何意义上的伤害。此外，他们有权利在任何阶段停止参与。 若您对本研究有任何疑问，欢迎联系我(_____@education.ox.ac.uk), 我会尽全力解答。

完成本问卷大约耗时 12-15 分钟。 非常感谢您的耐心阅读!祝一切顺利~

_____
Department of Education
15 Norham Gardens, Oxford
Tel: +44 _____
Email: _____@education.ox.ac.uk

Factors predicting the Heritage Language proficiency in Shanghainese

OPT-OUT FORM FOR PARENTS/GUARDIANS

Ethics Approval Reference: [CIA-22HT-023]

If you **DO NOT** want your child to participate in the above-named research study, please fill out the form below and return it to the school by [01/06/2022].

If we do not receive an opt-out form from you by this date, your child may be invited to take part in this study, as described in the accompanying information sheet.

I, the undersigned, hereby DO NOT give permission for my child to take part in the study titled Factors predicting the Heritage Language proficiency in Shanghainese.

Name of child: _____

Name of parent/guardian: _____

Signature: _____ Date: _____

Name of researcher: _____

Appendix 6: Circulation Email

Dear [teacher name],

I am _____, a MSc student in programme Applied Linguistics and Second Language Acquisition at the University of Oxford. I am currently working on my thesis about students' proficiency in Shanghainese, the Shanghai dialect, and would like to recruit the students from [school name] as my participants.

My research investigates speakers' proficiency in Shanghainese, a Chinese Heritage Language. A mixed-method approach combining quantitative questionnaire and qualitative interview would be adopted. It will look at the usage pattern of Shanghainese, the relationship between Shanghainese proficiency and speakers' participation in dialect activities at home and at school, motivations and attitudes, as well as some demographic factors. By doing so, this study will provide insights into the development and current status of Shanghainese among young people in Shanghai and could serve the purpose of Heritage Language preservation.

Please be assured that the University of Oxford has strict ethical procedures regarding research concerning young people and adults. I will inform parents/guardians, students and teachers before commencing the research. Parents'/guardians' opt-out form for their children participating in the research will be sought. The participation of students will be entirely voluntary. Throughout the research, the students are free to withdraw at any point in time without any consequences.

It would be lovely if you could circulate this email to other class teachers in [school name] and see whether their students are willing to partake in my research. I look forward to hearing back from you and we could further discuss about the details of the timing and procedures of this study.

Thank you for your precious time and hope you have a great day.

Yours sincerely,

_____
Department of Education
15 Norham Gardens, Oxford
Tel: +44 _____
Email: _____@education.ox.ac.uk

Appendix 7: Power analysis results