NORTHERN PINGHUA PROFICIENCY AMONG HERITAGE SPEAKERS: INFLUENCING FACTORS IN A MULTILINGUAL CHINESE CONTEXT

HUANG, W.[1,2] – SOON, C. T.[1*]

[1] *Centre for the Promotion of Knowledge and Language Learning, University Malaysia Sabah, Sabah, Malaysia.*

[2] *School of Humanities, Guilin University, Guangxi, China.*

**Corresponding author*
e-mail: soon[at]ums.edu.my

(Received 23rd March 2025; revised 06th July 2025; accepted 14th July 2025)

Abstract. Language proficiency plays a crucial role in heritage language maintenance by influencing language choice. Northern Pinghua, a heritage language spoken in rural multilingual areas of northern Guangxi, China, is undergoing a language shift. This study examines the self-rated proficiency of Northern Pinghua speakers and its relationship with demographic factors (age, gender, education, occupation, income) and home language use (frequency of Northern Pinghua, Guiliu Hua, and Mandarin). Data were collected from 34 participants using a closed-ended questionnaire. Results from group difference and correlation analyses indicate that: (a) age significantly affects proficiency, with speakers aged 36 and above reporting higher proficiency than younger counterparts; (b) males reported slightly higher proficiency than females; (c) home use of Northern Pinghua positively correlates with proficiency, while the use of Mandarin and Guiliu Hua shows negative correlations; (d) education and income levels do not significantly influence proficiency. These findings highlight the importance of home language practices and generational continuity in sustaining heritage language vitality. The study contributes to the understanding of heritage language dynamics in multilingual contexts and highlights the urgency of supporting language transmission within families.
Keywords: *language proficiency, heritage language, sociodemographic variation, home language use, multilingual rural China*

Introduction

Language diversity has been rapidly declining worldwide, with many languages at risk of extinction due to factors such as demographic shifts, political pressures, and socioeconomic changes (Gao, 2017). UNESCO estimates that there are approximately 7,000 languages spoken worldwide, among which about 40% of languages are threatened with extinction (UNESCO, 2024). These endangered languages generally are the heritage languages with a limited number of speakers, often from older generations, and are not being passed on to younger generations. Heritage speakers normally show different levels of language proficiency, ranging from passive overhearers to native-like speakers (Montrul, 2013). Those "overhearers" may have less willingness to use the heritage language and shift to use the most fluent language, which resulting in language shift and language endangerment. This global endangerment crisis is also evident in China. The process of the language loss and shift of minority languages and regional dialects has been in an abrupt way (Cao, 2014). Amid this global trend, Northern Pinghua, a lesser-studied, regional, rural Sinitic variety spoken in northern Guangxi, China, has experienced significant decline these years (UNESCO, 2024; Ministry of Education of the Peoples Republic of China, 2020).

Research problem

Northern Pinghua (桂北平话), a subgroup of the Chinese dialect Pinghua, is used mainly in the rural areas in northern Guangxi, China, especially in the multilingual rural Guilin city. Northern Pinghua refers to a group of dialects with low mutual intelligibility, and the term was put forward in The Outline of Chinese Dialects in Guangxi (mimeographed edition of Guangxi Normal University) for the first time in 1960. In 1985, The Chinese Dialect in Guangxi (Draft) (Yang et al., 1985) officially took it as an academic term. The Northern Pinghua-speaking region is a multilingual area where Mandarin serves as the official language, Guiliu Hua (a regional Chinese variety) functions as the lingua franca in Guilin city, and Northern Pinghua is predominantly used in rural areas. In early 2000s, Northern Pinghua was reported to be an active communicative language with the largest group of users in rural Guangxi, based on a large-scale study across 47 Pinghua-speaking locations (Li, 2007). However, subsequent research by Lin and Yu (2012) highlighted a decline in Northern Pinghuas geographical distribution and linguistic characteristics due to shifts in the linguistic ecological environment. Their findings align with the endangered/unsafe classification of Northern Pinghua in the UNESCO World Atlas (UNESCO, 2024) and the most recent assessment by the Chinese Language Resources Protection Project (Ministry of Education of the Peoples Republic of China, 2020). These studies indicate that Northern Pinghua is facing an urgent threat of language loss, yet a systematic assessment of Northern Pinghua language proficiency and its influencing factors in local Northern Pinghua speaking communities remains under-explored. In multilingual regions of China, heritage languages such as Northern Pinghua are facing increasing pressure from dominant and regional languages, particularly Mandarin and Guiliu Hua. Although research has identified language exposure and sociodemographic factors as key influences on heritage language maintenance, empirical evidence on Northern Pinghua proficiency and its influencing factors remains limited. Understanding how heritage language proficiency varies among heritage speakers and what factors affect it is essential for informing language policy, preservation efforts, and community-based planning.

Research objectives and research questions

This study aims to achieve the following objectives: (1) To investigate the self-rated proficiency level of Northern Pinghua among heritage speakers in a rural region of Guangxi, China; (2) To examine the extent to which demographic characteristics (e.g., age, gender, education, occupation, income) and home language use patterns (of Northern Pinghua, Guiliu Hua, and Mandarin) influence Northern Pinghua proficiency. This study tries to answer the following two research questions: (1) What is the self-rated proficiency level of heritage speakers in Northern Pinghua? (2) To what extent do demographic and home language use factors influence Northern Pinghua proficiency among heritage speakers? The findings will contribute empirical data to the broader discourse on heritage language maintenance, informed by Spolskys language management theory, which highlights the interaction between language proficiency and language use in shaping language behavior.

Literature review

This section reviews literature on language proficiency and language policy and planning, with a particular focus on the demographic and linguistic factors influencing language proficiency.

Language policy and planning and language proficiency

Language proficiency has been a core focus within the field of Language Policy and Planning (LPP). LPP is a multidisciplinary field that explores how languages are shaped, valued, and regulated through formal and informal mechanisms in various sociopolitical contexts (Ricento, 2000). Language maintenance, a key issue in LPP, refers to the preservation and continued use of minority or heritage languages across generations. A foundational theoretical perspective in this area is Spolskys Language Management Theory (Spolsky, 2021; 2009; 2004), which identifies three interrelated components of language policy: language practices, language ideologies, and language management. This framework highlights that language proficiency is an important part of language practices. Building on Haugen (1972) concept of the ecology of language, Spolsky (2021) emphasizes the impact of both linguistic and non-linguistic factors, including gender, age, education, and socio-economic status, on language practices. By incorporating Fishman (1991) notion of domain, Spolsky (2009) also identifies the family as a critical place for intergenerational transmission. This is evidenced by Vorobyeva and Bels (2021). They find that family language use is strongly linked to heritage language proficiency among young Russian speakers in Spain. In summary, language proficiency is an important part of heritage language practices, and is closely related to family practices and the immediate language environment. Therefore, when studying the heritage language maintenance, it is necessary to explore its speakers' language proficiency and the correlated linguistic and non-linguistic factors.

Linguistic and non-linguistic factors and language proficiency

Age is a key demographic factor in shaping language choice, particularly in the heritage language maintenance. Younger speakers now generally are less proficient in their heritage language or the traditional dialects than the elders. In China, for example, younger generations demonstrate lower proficiency in regional dialects as parental priorities increasingly favor Mandarinthe official languagefor educational and economic advancement (Wang and King, 2024). This "language shift" toward dominant varieties is driven by pragmatic concerns: Mandarin proficiency is perceived as essential for academic success and national integration, leading families to prioritize it over dialects in daily communication. Similar patterns are evident in Lithuania, where rural dialects are predominantly used by older generations, while younger Lithuanians increasingly rely on Standard Lithuanian or English in professional and social contexts (Ramonienė, 2018). This is also observed in the Shetland Islands, where younger speakers favor Scottish Standard English (Smith and Durham, 2011). This disruption of intergenerational language transmission poses serious threat to heritage language maintenance. Gender is also a basic demographic factor which requires to be studies in language practices. Previous studies show that there are gender differences in language skills and heritage language maintenance across diverse contexts, though exceptions exist. For example, a large-scale study (N=199,097) on 15-year-olds in 43 countries revealed that girls had better reading competence than boys in every nation surveyed. Similarly, in Panama, a 3% gender gap in language conservation emerged among 2429-

year-olds, with women showing slightly greater preservation of their ethnic language than men (Guevara et al., 2024). Besides, cross-continental evidence from New Zealand and Australia also shows that women in both first- and second-generation immigrant groups tend to keep using their ethnic languages longer than men (Holmes, 1993). Li et al. (2023) find that among Chinese-Canadian first graders (N=76), girls outperformed boys in both Chinese (L1) and English (L2) vocabulary. However, some research challenges this trend. A study of young Russian heritage speakers in Spain found no significant gender differences in heritage language proficiency (Vorobyeva and Bel, 2021).

Previous studies find that the use of dominant language in education has a negative impact on indigenous language proficiency and cultural continuity (Bromham et al., 2020). Using dominant languages (such as English) in education often leads to the replacement of heritage language acquisition by the dominant language, the decline in heritage language daily use, and the disruption of intergenerational transmission (Bromham et al., 2020). The relationship between income and language proficiency can be positive or negative, which deserves a close examination. First, that if a language can better income shows the market demands. In contexts which value cultural competence such as tourism, education, or public services, bilingual competence in both heritage language and the dominant languages (e.g., English or Spanish) may bring better income than monolinguals (De La Fuente Stevens and Pelkonen, 2023; Henley and Jones, 2005). Besides, in some regions where implement strong cultural protection policies and value heritage cultures, better heritage language ability is encouraged and may bring better economic benefits. However, in some labor markets such as globalized or standardized sectors which value dominant languages (e.g., English), the monolingual competence in heritage language may limit the speakers' opportunity to get higher income (Chiswick and Miller, 2007). In these regions, heritage language competence could not bring economic benefits, thus lowering the individuals' willingness to maintain their heritage languages and causing the intergenerational transfer to using dominant languages (Chiswick and Miller, 2007). In overall, if heritage language competence could bring better earnings is complex and the correlation between the two should be studied within the specific context.

Language proficiency is also closely related to the regulations governing language choices in professional fields. Public sectors like government and education often enforce the use of official languages (a mainstream or colonial lingua franca). For example, in Nigeria, government officials are required to use English due to a state-level strategy to standardize management (Ayeomoni, 2012). Similarly, in Malaysia, English is favored in education and professional sectors such as medicine and engineering, and this is driven by national policies to meet international academic and industry standards (Mei et al., 2016). In contrast, some professional fields may adapt language choice according to their audience's language characteristics. For example, Malaysian Tamil Christian youths use English in educational and professional contexts due to institutional pressures, but they use ethnic languages like Tamil or Malay persist in informal family or community interactions (Leo and Abdullah, 2013). Similarly, in Kenya, in home and community domains, the Kinubi language remains active though mainstream languages are preferred in institutional or professional fields. This reveals professions facing local clients or communities may prioritize the use of local languages to foster trust and enhance cultural competence (Adams et al., 2012). In conclusion, these patterns show that official/global sectors use mainstream languages for

standardization and accessibility, while professions working directly with local-language communities often adapt to the regional heritage languages.

The quantity and quality of heritage language exposure is a critical determinant of language proficiency. This exposure can be either passive contact to a certain language or the active language use. The increased exposure to dominant or majority languages like English Mandarin, or Spanish, often leads to the abandonment or the stagnation of native tongues, threatening cultural traditions and indigenous identities (Chui, 2023). For example, in Nigeria, English dominance has marginalized indigenous languages (Ajepe and Ademowo, 2016). What's more, heritage language competence is shaped not just by whether they can access languages, but by how often they use languages and in what situations (De Houwer, 2017). For example, Mulík et al. (2021) find that language use and proficiency are essential for maintaining the Hnanho language among bilingual speakers. Besides, among all the related domains, Researchers find that family-based exposure is closely related to heritage language proficiency. Vorobyeva and Bel (2021) also find that family language use significantly affects its proficiency. Without regular use, even high levels of exposure may fail to result in full language proficiency. Previous studies reveal that factors influencing language choice are diverse and interrelated, with language proficiency playing a pivotal role. This review emphasizes the need to investigate language proficiency particularly in relation to demographic factors and linguistic factors such as language exposure, which this study aims to explore.

Materials and Methods

Research design

This study employs a questionnaire to investigate the self-rated proficiency level of Northern Pinghua among heritage speakers in rural China, and to examine how a range of demographic and home language use variables, including age, gender, educational level, occupation, income, and use frequency of Northern Pinghua, Guiliu Hua, and Mandarin at home, are associated with Northern Pinghua proficiency.

Participants

The respondents of this study were local villagers from various Northern Pinghua-speaking households located in the rural Dahe Village of Guangxi Zhuang Autonomous Region, China. In this village, people can use three languages with Northern Pinghua as the heritage language. The sample included participants from a range of age groups to capture a comprehensive picture of how language practices have evolved across generations, as well as how different family members contribute to the maintenance of Northern Pinghua. Participants were recruited using a combination of purposive and snowball sampling methods to ensure representation from a diverse cross-section of the Northern Pinghua-speaking community. *Table 1* provides a summary of demographic information of the respondents. A total of 34 individuals participated in the survey, comprising an equal number of females (n=17, 50%) and males (n=17, 50%). In terms of age distribution, 5 participants (14.7%) were under 18 years old, 17 (50.0%) were in group aged 18 to 35, 6 (17.6%) were aged 36 to 50, and both the 51-60 and over-60 groups included 3 participants each (8.8%). Regarding educational background, 6 participants (17.6%) had completed only primary education or had no formal schooling,

15 (44.1%) had completed junior middle school, and 9 (26.5%) had an education level equivalent to middle school. Additionally, 1 participant (2.9%) had attended junior college, while 3 (8.8%) held graduate or postgraduate degrees. In terms of occupation, 10 respondents (29.4%) were farmers or fishermen, 7 (20.6%) were students, and another 7 (20.6%) worked in service-related roles such as cook, hairdresser, driver, security, cleaner, or waiter. Furthermore, 6 participants (17.6%) were unemployed, 4 (11.8%) were employed as government officials or institutional staff, and 1 respondent (2.9%) identified as other. Monthly income levels also varied across the sample. Thirteen respondents (38.2%) reported having no income, 2 (5.9%) earned less than 1,400 RMB, 9 (26.5%) reported an income between 1,400 and 3,000 RMB, 8 (23.5%) earned between 3,001 and 5,000 RMB, and 2 (5.9%) reported earning more than 5,000 RMB per month.

Table 1. *Respondents' demographic information.*

Variable		Female	Male	Frequency (N)	Percentage (%)
Gender	Female	17	0	17	50.0%
	Male	0	17	17	50.0%
Age	<18 years	5	0	5	14.7%
	18-35 years	9	8	17	50%
	36-50 years	1	5	6	17.6%
	51-60 years	0	3	3	8.8%
	>60 years	2	1	3	8.8%
Education	Primary school/ uneducated	3	3	6	17.6%
	Junior middle school	7	8	15	44.1%
	Middle school/ equivalent	6	3	9	26.5%
	Junior college	0	1	1	2.9%
	Graduate/postgraduate	1	2	3	8.8%
Occupation	Farmer/fisherman	2	8	10	29.4%
	Student	7	0	7	20.6%
	Cook/ hairdresser/ driver/security/ cleaner/waiter/others	3	4	7	20.6%
	Unemployed	3	3	6	17.6%
	Government official/institution staff	2	2	4	11.8%
	Company/factory employee	0	0	0	0.0%
	Others	1	0	1	2.9%
Monthly Income	No income	10	3	13	38.2%
	<1400 RMB	0	2	2	5.9%
	1400-3000 RMB	4	5	9	26.5%
	3001-5000 RMB	3	5	8	23.5%
	>5000 RMB	0	2	2	5.9%

Respondents were multilingual villagers, a characteristic that reflects the multilingual context of the community (*Table 2*). Specifically, in Northern Pinghua, 25 respondents self-assessed their proficiency as high, 6 as moderate, and 3 as basic, with no participants reporting minimal proficiency. For Guiliu Hua, 28 respondents rated themselves as highly proficient, while 6 reported moderate proficiency; no individuals indicated basic or minimal proficiency in this language. Regarding Mandarin proficiency, 19 respondents classified their ability as high, 13 as moderate, and 2 as basic.

Table 2. *Respondents' language proficiency distribution across three languages.*

Language Type	High Proficiency (n)	Moderate Proficiency (n)	Basic Proficiency (n)	Minimal Proficiency (n)	N
Northern Pinghua	25	6	3	0	34
Guiliu Hua	28	6	0	0	34
Mandarin	19	13	2	0	34

Instruments

A structured questionnaire with closed-ended questions was developed, as a self-reported questionnaire was useful to provide information from a large group of respondents, and can be objectively compared and interpreted through data analysis (Park and Kim, 1997). The design of the questionnaire was informed by Spolsky (2004) language management framework, which emphasizes the key role of language proficiency in influencing language use. Additionally, the selection of variables was informed by existing literature, which highlights demographic characteristics (e.g., age, gender, education, occupation, income level), self-rated proficiency and the use frequency at home across three commonly spoken languages, Northern Pinghua, Guiliu Hua, and Mandarin as key factors influencing heritage language proficiency. Respondents reported the language proficiency of their family members using a four-point scale: 1=Minimal proficiency (no comprehension), 2=Basic proficiency (comprehension without the ability to speak), 3=Moderate proficiency (limited communicative ability), and 4=High proficiency (full proficiency). Regarding the language use in family domain, respondents reported language use frequency using a four-point scale: 1=Rarely, 2=Seldom, 3=Sometimes, 4=Always.

Data collection and data analysis

The research was conducted during April 8th to April 23th, 2024. The paper questionnaires were administered in person, aiming to clarify any questions and ensure that respondents fully understood the questions. Before the participant recruitment, the purpose of the research was clearly explained. During the data collection and processing procedures, stringent confidentiality protocols were enforced to guarantee that the personal information of interviewees remained strictly protected against unauthorized disclosure, misuse, or exploitation. When presenting research findings, all data and discoveries were reported with complete honesty and precise accuracy, adhering strictly to ethical standards and scientific rigor. The survey data underwent the following statistical tests and analyses using IBM SPSS 29.0: descriptive analysis alongside inferential statistical techniques, including the Shapiro-Wilk test, Kendalls Tau correlation test, and Kruskal-Wallis test. First, since the sample size (N=34) was small, the normality of variables was tested by using the Shapiro-Wilk test (Yazici and Yolacan, 2007). Results indicated that all variables exhibited statistically significant deviations from normality ($p<.05$). Therefore, subsequent analyses were conducted under the premise of approximate non-normality. To answer the first research question, the descriptive analysis was used to show respondents self-rated Northern Pinghua proficiency. To address the second research question, a combination of Kendalls Tau correlation tests and group difference analyses (such as Kruskal-Wallis tests) were employed to understand the associations between demographic and linguistic factors and Northern Pinghua proficiency. Nominal variables such as gender and occupation were tested for group differences only, while ordinal variables (e.g., age group, education level, use frequency of Mandarin at home) were analyzed using both methods to capture general trends and specific group-level distinctions.

Results and Discussion

Northern Pinghua proficiency

To address the first research question, participants self-rated Northern Pinghua language proficiency was reported using a 4-point scale, from level 1 (minimal) to level 4 (high). *Table 3* provides a summary of Northern Pinghua language proficiency levels, revealing a generally high degree of self-reported proficiency among the 34 participants. The mean score was 3.65 (SD=0.65), with scores ranging from 2 to 4. Both the median and the 75th percentile equaled 4, indicating that at least half of the respondents rated themselves at the highest proficiency level (Level 4). This distribution suggests a strong overall command of the language within the community. The distribution of Northern Pinghua proficiency across age groups and genders was also computed. In the <18 age group (n=5; all female), three participants (60%) reported basic proficiency (Level 2), while two (40%) indicated moderate proficiency (Level 3). Among participants aged 1835 (n=17; 9 females, 8 males), female proficiency levels varied: four females (44.4%) reported moderate proficiency (Level 3), while the remaining five females (55.6%) and all eight males (100%) reported high proficiency (Level 4). In the older age groups3650 (n=6; 1 female, 5 males), 5160 (n=3; all male), and >60 (n=3; 2 females, 1 male) all participants reported high proficiency (Level 4). Notably, no respondents across any age or gender group reported minimal proficiency (Level 1). Across all age groups and genders, participants reported varying levels of Northern Pinghua proficiency. Younger participants (<18) reported basic to moderate proficiency levels. Among participants aged 1835, males predominantly reported high proficiency, while females showed both moderate and high levels. All participants aged 36 and above reported high proficiency. No participants reported minimal proficiency.

Table 3. *Summary of Northern Pinghua proficiency.*

Variable	Range	Min	Max	Mean	SD	25%	50% (Median)	75%
Northern Pinghua Language Proficiency	2	2	4	3.65	0.65	3	4	4

Language use frequency at home

To address the second research question, the distribution of language use frequency at home across three languages, the group differences analysis and correlation analysis were conducted. This part presented the results of language use frequency. *Table 4* shows that the use frequency of Northern Pinghua at home, scored on a scale from 1 to 4 (with higher scores indicating more frequent use), exhibited the greatest variability (SD=1.04). The mean score was 3.32, with a median of 3 and a range from 1 (rarely) to 4 (always). The 25th percentile score of 1 suggests that 25% of participants reported rarely using Northern Pinghua at home. The use frequency of Guiliu Hua at home had a mean of 2.79 (SD=0.88), with scores ranging from 1 to 4. The median was 3, and the 75th percentile was 4, indicating that while 25% of participants seldom used Guiliu Hua at home, the majority reported using it either sometimes or always. Mandarin use at home showed the lowest mean frequency (M=2.48, SD=1.15), with a range from 1 to 4. The median was 3, and the 25th percentile score was 1, indicating that a quarter of participants rarely used Mandarin at home, while the remainder reported using it sometimes or more frequently. The data shows that Northern Pinghua was the most frequently used language at home, while Mandarin showed the lowest average usage, with notable variability across all three languages. *Figure 1* displays the frequency of language use among participants across four categories: Rarely, Seldom, Sometimes, and Always. For Northern Pinghua, most participants (n=22) reported using it always, while fewer indicated using it sometimes (n=4), seldom (n=5), or rarely (n=3). In

contrast, Gui Liu was reported as always used by 9 participants, with 10 indicating sometimes, 14 seldom, and only 1 reporting rarely. For Mandarin, responses were more evenly distributed: 8 participants reported always using it, 9 sometimes, 7 seldom, and 9 rarely. These figures reflect varying usage patterns across the three languages.

Table 4. *Distribution of language use frequency at home across three languages.*

Variable	Range	Min	Max	Mean	SD	25%	50% (Median)	75%
Use Frequency of NP Language at Home	3	1	4	3.32	1.04	3	4	4
Use Frequency of GL Language at Home	3	1	4	2.79	0.88	2	3	4
Use Frequency of Mandarin Language at Home	3	1	4	2.48	1.15	1	3	3

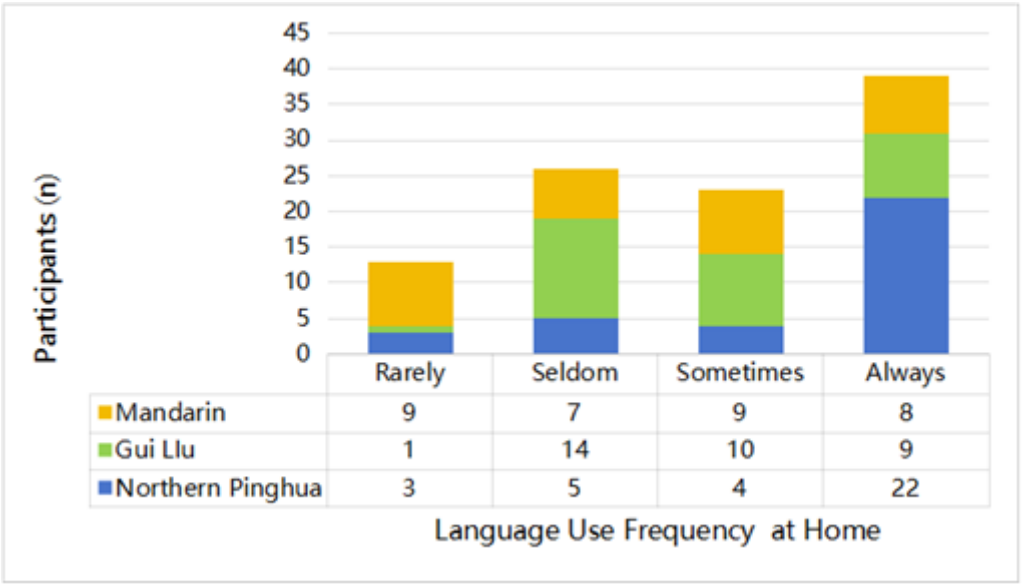

Figure 1. *The distribution of language use frequency at home.*

Group differences analysis on Northern Pinghua language proficiency

Group differences in Northern Pinghua proficiency were examined using Kruskal-Wallis H tests and descriptive statistics. A series of Kruskal-Wallis tests were performed to assess whether demographic variables (gender, age group, education level, occupation, income level) and sociolinguistic variables (language use frequency of Northern Pinghua, Guiliu Hua, and Mandarin at home) significantly influenced participants proficiency in Northern Pinghua. The results revealed that age group, gender, and use frequency of Northern Pinghua and Mandarin at home demonstrated statistically significant differences in Northern Pinghua proficiency at the $p<.001$ level. Additionally, occupation, education level, and use frequency of Guiliu Hua at home showed significant group differences at the $p<.05$ level. The following section presents the group differences in Northern Pinghua proficiency in relation to each individual variable. In terms of demographic variables, age group showed a significant effect on Northern Pinghua proficiency, $H(4)=20.172$, $p<.001$. Participants under 18 years had the lowest mean rank (3.8), suggesting the lowest level of Nothern Pinghua proficiency, while participants in older age categories (3650, 5160, and above 60) consistently showed highest mean ranks (22), indicating highest Northern Pinghua proficiency.

Results revealed a statistically significant difference between males and females, H(1)=11.634, p<.001, with males (Mean Rank=22) scoring higher than females (Mean Rank=13). This means that males reported higher Pinghua proficiency than females. Occupation was significantly associated with the Northern Pinghua proficiency as well, H(6)=20.616, p=.002. Students had the lowest mean rank (28.21), suggesting lower proficiency, whereas individuals in labor-oriented roles (farmers, unemployed, cleaners, service roles, etc.) generally had higher mean ranks (18.9-22), indicating higher proficiency. A statistically significant difference was also found across educational levels, H(4)=9.996, p=.041. Participants with higher education had the highest average rank (22.00), followed by those with basic education (primary or junior middle school), whereas those with intermediate middle school education had the lowest rank (10.67), indicating relatively lower proficiency. No statistically significant difference was found across income levels, H(4)=7.676, p=.104. In terms of linguistic background, a statistically significant difference was found across the use frequency of Northern Pinghua at home, H(3)=29.955, p<.001. Participants who always used Northern Pinghua at home had the highest mean rank (22), followed by those who sometimes used the language, with a higher mean rank (18.13), indicating higher Northern Pinghua proficiency. However, those who rarely or seldom used the language at home had the lower mean ranks (2-6.5), indicating lower Northern Pinghua proficiency.

Significant group differences were also found across frequencies of GL use at home, H(3)=15.354, p=.002, participants who seldom or sometimes used Guiliu Hua at home had the highest mean rank (20.89, and 20.45 respectively), indicating higher Northern Pinghua proficiency level. In contrast, those who always used Guiliu Hua at home had the lower mean rank (10.67), indicating lower Northern Pinghua proficiency level. As for official language, a significant difference was observed across the use frequency of Mandarin at home, H(3)=20.09, p<.001. Participants who never used Mandarin at home had the highest mean rank (21.5), follow by those who seldom (mean rank=19.36) and who sometimes (mean rank=19.83) used Mandarin, indicating these group of participants had higher level of Northern Pinghua proficiency level. However, those always using Mandarin showed the lowest mean rank (6.69), indicating the lowest Northern Pinghua proficiency. Overall, these findings collectively suggest that age group, gender, educational attainment, occupation, use frequency of Northern Pinghua, Guiliu Hua, and Mandarin at home are important factors influencing Northern Pinghua proficiency. However, income level showed no significant association.

Correlation analysis on language proficiency

A correlation analysis was conducted to investigate which demographic factors and linguistic factors had significant correlation with Northern Pinghua proficiency. The variables were listed in *Table 5*. A series of Kendall's Tau correlation (τ) analysis were conducted and the complete results were presented in *Table 6*. The results revealed that Northern Pinghua proficiency had significant positive correlations with age group, use frequency of Norther Pinghu, Guiliu Hua, and Mandarin at home (p<.05 for all), had a significant negative correlation with use frequency of Mandarin at home (p<.001) and Guiliu Hua (p=.035), but with no significant relationship with the education level and the income level of the sample size.

Table 5. *Variable pairings for correlation testing.*

Variable 1	Variable 2

Northern Pinghua proficiency	(1) Demographic factors: Age group, education level, income level (2) Linguistic factors: Use frequency of Northern Pinghua at home; use frequency of Guiliu Hua at home; use frequency of Mandarin at home

***Table 6**. Kendall's tau correlation analysis results.*

Category		Age group	Education level	Income level	Use frequency (NP)	Use frequency (Guiliu Hua)	Use frequency (Mandarin)
NP proficiency	τ	.598**	.195	-.217	.873**	.335*	-.587**
	p	<.001	.212	.163	<.001	.035	<.001

Note: NP=Northern Pinghua.

The first research questions is what is the self-rated proficiency level of heritage speakers in Northern Pinghua. The self-rated proficiency level of heritage speakers in Northern Pinghua is generally high, with most participants reporting strong competence in the language. Further breakdown by age and gender indicate a clear age-related trend in Northern Pinghua proficiency, with higher levels of proficiency more consistently reported among older participants. While younger participants (<18) exhibited only basic to moderate proficiency, individuals in older age groups, particularly those aged 36 and above, uniformly demonstrated high proficiency. Gender differences were also observed among younger adults (1835), where males showed higher proficiency levels compared to females. Overall, no participants reported minimal proficiency, suggesting that some degree of Northern Pinghua competence is maintained across all age and gender groups, although the strength of proficiency appears to be associated with age and, to a lesser extent, gender. The second research question was to examine to what extent do demographic and home language use factors influence Northern Pinghua proficiency among heritage speakers. First, the group variation analysis of Northern Pinghua proficiency revealed significant differences across several demographic variables, including age group, gender, and occupation, as well as linguistic variables such as use frequency of Northern Pinghua, Guiliu Hua, or Mandarin at home. In contrast, no significant differences were found for income level. In specific, respondents in the following categories had higher Northern Pinghua proficiency level: villagers older than 35, males, labor-oriented roles (farmers, unemployed, cleaners, service roles, etc.), those always or sometimes using Northern Pinghua at home, those less frequently using Mandarin, and those seldom or sometimes using Guiliu Hua. In contrast, villagers younger than 18, females, students, government officials, respondents who rarely or seldom used Northern Pinghua at home, and those always using Mandarin or Guiliu Hua, demonstrated lower Northern Pinghua proficiency level. Second, the correlation analysis revealed that Northern Pinghua proficiency was significantly associated with age group, use frequency of Northern Pinghua, Guiliu Hua, and Mandarin at home. No significant correlations were found with education level, and income level. These findings highlighted the significant group differences of Northern Pinghua proficiency among villagers, and the factors correlated to Northern Pinghua proficiency. While the results of the correlation and group difference analyses were largely consistent, an interesting divergence emerged in the case of educational level. Specifically, educational attainment did not show a significant correlation with Northern Pinghua proficiency, but it did yield a statistically significant difference in the group comparison analysis. This suggests that although certain educational groups may differ significantly

from others (e.g., intermediate vs. higher education), there is no clear overall linear trend between education level and Pinghua proficiency across the entire sample.

Language shift and Northern Pinghua proficiency

The findings revealed that age group was a significant correlated factor to the Northern Pinghua language proficiency. Respondents above 35 years old showed higher Northern Pinghua proficiency than those younger than 18 years of age. These suggest that as age decreases, there is a tendency for lower Northern Pinghua proficiency. The younger respondents are more likely to report lower proficiency, which could reflect less Northern Pinghua exposure and practice over time. It presents a possibility or even an established fact that the Northern Pinghua speakers are diminishing and this trend is going to continue. Recent studies indicated a similar concerning trend of declining proficiency in indigenous languages among younger generations (Wang and King, 2024; Ramonienė, 2018; Smith and Durham, 2011). In Lampung, Indonesia, there was a shift away from local languages, with few young people can speak their native tongue fluently. Also, in Malang, Java, the Krama Madya language was categorized as vulnerable due to diminishing intergenerational transmission.

Men with higher Northern Pinghua proficiency than women

The results revealed that male respondents demonstrated higher levels of Northern Pinghua proficiency compared to their female counterparts. Taking the 18-35 age group for example, a total of 9 females and 8 males participated in the survey. Among them, 4 females rated their language proficiency as moderate, while the remaining 5 females and all 8 males rated their proficiency as high proficiency. This result contrasts with the findings of previous studies which suggest women generally tend to have better heritage language competence than men (Guevara et al., 2024) or no gender difference in heritage language proficiency of the children (Vorobyeva and Bel, 2021). The pattern that woman outperformed men is possibly attributed to womens stronger social networks promoting regular use of the ethnic language and their greater emphasis on its social and emotional value (Holmes, 1993). However, the finding of this study reveals that men showing a higher Northern Pinghua proficiency level than women. One possible explanation may be that men traditionally play the critical role in dialect maintenance in China. In Rural China, Fei et al. (1992) argued that the traditional Chinese social structure is centered around patrilineal kinship, with men playing a crucial role in the continuation of the family and clan. Due to this patrilineal system, the transmission of culture and customs, especially language, primarily depends on men. Men not only bear the responsibility for bloodline inheritance within the family but also play a central role in maintaining clan culture. For instance, activities such as ancestral rites and the compilation of genealogies are typically led by men, and these cultural practices serve as key contexts for the use and transmission of dialects. As a result, men play a dominant role in the inheritance of dialects.

Domain functions and Northern Pinghua proficiency

Occupation was another significant factor influencing Northern Pinghua proficiency. The lower mean ranks of students and government officials means that they were less proficient in Northern Pinghua. This suggests that these groups may have less exposure to or reliance on Northern Pinghua, potentially due to differing occupational demands or

social environments. The finding echoes previous research that official and globalized sectors rely on mainstream languages to ensure standardization and accessibility (Mei et al., 2016; Ayeomoni, 2012). In China, in official and educational settings, Mandarin shall be used as the official and instructional language (Ministry of Education of the People's Republic of China, 2001). On contrast, fishermen, security guards, the unemployed and alike were with higher Northern Pinghua level. This suggests that these occupations are usually concentrated in dialect areas, and their clients are mainly local people. These findings are consistent with previous research showing that roles involving direct engagement with specific linguistic communities tend to accommodate local language use (Leo and Abdullah, 2013; Adams et al., 2012). In case of Northern Pinghua heritage speakers, labor-related occupations such as farming are deeply connected to Northern Pinghua-speaking communities, which in turn further reinforces their use of Northern Pinghua.

Educational level, income level and Northern Pinghua proficiency

The findings of this study indicated that education level and income level did not show a significant impact on Northern Pinghua proficiency in this sample size. This suggests that the Northern Pinghua proficiency may be more influenced by other factors like age, attitudes, and domain functions, rather than being directly determined by income or education level. This is contrary to the normal beliefs that the education in dominant or more advantaged languages negatively impact the heritage language maintenance (Bromham et al., 2020). Previous research also shows that while heritage language proficiency correlates with improved economic outcomes with bilingual competence (De La Fuente Stevens and Pelkonen, 2023; Henley and Jones, 2005) or in specialized industries, some studies state that monolingual in heritage or minority languages may have limited opportunities to access to high-paying jobs (Chiswick et al., 2007). However, individual language choice is inherently complex, and regional diversity in both language policy and economic frameworks must be considered. In the case of Northern Pinghua, the functional roles of Mandarin and the regional dialect Guiliu Hua already meet the communicative needs of both formal and private domains. Respondents in this study exhibit a multilingual linguistic repertoire, with the majority demonstrating at least moderate proficiency in all three languages: Northern Pinghua, Guiliu Hua, and Mandarin. This linguistic flexibility allows them to switch languages as required by context. Consequently, the level of proficiency in Northern Pinghua may not be significantly associated with income level.

Language use frequency and Northern Pinghua proficiency

The most significant finding was the effect of language use frequency at home on Northern Pinghua proficiency. Respondents who frequently used Northern Pinghua at home-while using Mandarin and Guiliu Hua less often demonstrated higher proficiency in Northern Pinghua. This shows that increased use of Northern Pinghua within the home domain appears to be a crucial factor in improving or sustaining its proficiency across generations. This supports prior findings that the frequency and context of language use are central to heritage language development (Mulík et al., 2021; Vorobyeva and Bel, 2021). In contrast, those who often Mandarin or Guiliu Hua at home showed relatively lower Northern Pinghua proficiency level. This suggests that as the use of Mandarin increases, the proficiency in Northern Pinghua decreases,

indicating a potential language shift or interference where the dominance of Mandarin may hinder the maintenance and use of Northern Pinghua. These findings are consistent with Fishman (1991) theory of language shift, where the dominance of a majority language contributes to the decline of minority languages. This is also echoed by Chui (2023) finding which is the adaption of dominant languages often leads to the abandonment of mother tongue and the loss of cultural tradition.

Conclusion

The following limitations should be recognized. First, the studys sample size was relatively small (n=34), which may limit the generalizability of the findings. The sample primarily consists of individuals from a specific geographic area, so future research should consider larger, more diverse samples from different regions to ensure more representative results. Second, the study relied on self-reported language proficiency, which may introduce subjective bias. Objective assessments of language proficiency are suggested if a more accurate measurement of Northern Pinghua language skills is required. This study contributes to the field of heritage language research in the following aspects. First, the study in this multilingual context reveals age and gender differences in Northern Pinghua proficiency. Older speakers reported higher proficiency than younger generations, and males showed higher proficiency than females. Second, the study finds that home language use is significantly correlated with heritage language proficiency. The strong correlation between frequent use of Northern Pinghua at home and higher self-rated proficiency reveals that regular family language practices is important to sustain the heritage languages. Lastly, contrary to the common assumptions about the influence of socioeconomic factors, the study finds that, there is no significant relationship between education or income level and heritage language proficiency in this sample size. The research findings provide several directions for future studies. First, further investigation can investigate the intergenerational transmission patterns of Northern Pinghua, such as family interactions and the parental influence on childrens language practices. Second, the findings of gender differences in Northern Pinghua proficiency among young adults suggest that gender, with different social roles and expectations, may impact language use patterns. Future research could explore how gender intersects with linguistic ideologies and its role in heritage language maintenance. Finally, while this study focuses on home language use, other domains in the immediate environment, like peers, community, or digital communication, also need further investigation. Expanding the scope beyond family domain will contribute to a more comprehensive understanding of the heritage language maintenance like Northern Pinghua.

Acknowledgement

We are deeply grateful to the villagers who generously participated in this study. Their willingness to share their time and experiences made this research possible and meaningful. Our appreciation also goes to Universiti Malaysia Sabah (UMS) and Guiliu University for their invaluable support throughout the research journey. Without their assistance and encouragement, this study would not have come to fruition.

Conflict of interest

The authors confirm that there is no conflict of interest involve with any parties in this research study.

REFERENCES

[1] Adams, Y., Matu, P.M., Ongarora, D.O. (2012): Language use and choice: A case study of Kinubi in Kibera, Kenya. – International Journal of Humanities and Social Science 2(4): 99-104.
[2] Ajepe, I., Ademowo, A.J. (2016): English language dominance and the fate of indigenous languages in Nigeria. – International Journal of History and Cultural Studies 2(4): 10-17.
[3] Ayeomoni, M.O. (2012): The languages in Nigerian socio-political domains: Features and functions. – English Language Teaching 5(10): 12-19.
[4] Bromham, L., Hua, X., Algy, C., Meakins, F. (2020): Language endangerment: a multidimensional analysis of risk factors. – Journal of Language Evolution 5(1): 75-91.
[5] Cao, Z. (2014): Dialect endangerment, cultural fragments, and the mission of dialect scholars. – Zhongguo Yuyan Xuebao (Journal of Chinese Linguistics) 16(1): 207-214.
[6] Chiswick, B.R., Miller, P.W. (2007): The economics of language: International analyses. – London: Routledge 622p.
[7] Chui, J. (2023): The effects of cultural assimilation on the loss of the Tuvan and Seri languages. – Journal of Student Research 11(3): 1-4.
[8] De La Fuente Stevens, D., Pelkonen, P. (2023): Economics of minority groups: Labour-market returns and transmission of indigenous languages in Mexico. – World Development 162: 33p.
[9] Fei, X., Hamilton, G.G., Zheng, W. (1992): From the soil: The foundations of Chinese society. – Univ. of California Press 160p.
[10] Fishman, J.A. (1991): Reversing language shift: Theoretical and empirical foundations of assistance to threatened languages. – Clevedon: Multilingual Matters 76: 431p.
[11] Gao, K.B. (2017): Dynamics of language contact in China: Ethnolinguistic diversity and variation in Yunnan. – University of Hawai'I 153p.
[12] Guevara, I., Cabrera, G., Suleiman, N., Sanjur, N. (2024): Role of gender in the conservation of the Ngabere language in Sitio Prado, Indigenous Area. – Revista Colegiada de Ciencia 5(2): 41-49.
[13] Haugen, E. (1972): The ecology of language. – Stanford, CA: Stanford University Press 10p.
[14] Henley, A., Jones, R.E. (2005): Earnings and linguistic proficiency in a bilingual economy. – The Manchester School 73(3): 300-320.
[15] Holmes, J. (1993): Immigrant women and language maintenance in Australia and New Zealand. – International Journal of Applied Linguistics 3(2): 159-179.
[16] De Houwer, A. (2017): Bilingual language input environments, intake, maturity and practice. – Bilingualism: Language and Cognition 20(1): 19-20.
[17] Leo, A.R., Abdullah, A.N. (2013): Language choice and use of Malaysian Tamil Christian youths: A survey. – Frontiers of Language and Teaching 4: 149-166.
[18] Li, G., Lin, Z., Zhen, F., Gunderson, L., Ji, R. X. (2023): Home literacy environment and early biliteracy engagement and attainment: A gendered perspective. – Bilingual Research Journal 46(3-4): 258-274.
[19] Li, L. (2007): The distribution, internal division, and genetic issues of Pinghua. – Fangyan (Dialect) 1: 71-78.
[20] Lin, Y., Yu, J. (2012): The status of Guangxi Pinghua. – Journal of Guilin Teachers College 2: 1-5.

[21] Mei, T.L., Abdullah, A.N., Heng, C.S., Kasim, Z.B. (2016): Language choice and use of Malaysian public university lecturers in the education domain. – Advances in Language and Literary Studies 7(1): 21-32.
[22] Ministry of Education of the People's Republic of China (2020): Response to Proposal No. 8137 at the Third Session of the 13th National People's Congress. – Ministry of Education of the People's Republic of China 4p.
[23] Ministry of Education of the People's Republic of China (2001): Law on the Standard Spoken and Written Chinese Language of the People's Republic of China. – Ministry of Education of the People's Republic of China 6p.
[24] Montrul, S. (2013): How "native" are heritage speakers? – Heritage Language Journal 10(2): 153-177.
[25] Mulík, S., Corona-Dzul, B., Amengual, M., Carrasco-Ortíz, H. (2021): Psycholinguistic profile of Otomi (Hñäñho)–Spanish bilinguals, migrants from Santiago Mexquititlán to Querétaro, Mexico [Perfil psicolingüístico de los bilingües otomí (hñäñho)–español, migrantes de Santiago Mexquititlán a Querétaro, México]. – Cuadernos de Lingüística de El Colegio de México 8: 1-50.
[26] Park, K.S., Kim, S.H. (1997): Tools for interactive multiattribute decision making with incompletely identified information. – European Journal of Operational Research 98(1): 111-123.
[27] Ramonienė, M. (2018): Regional dialects in the Lithuanian urban space: Skills, practices and attitudes. – In M. Ramonienė & L. Zabarskaitė (Eds.), Multilingualism in the Baltic States: Societal Discourses and Contact Phenomena, London: Palgrave Macmillan UK 29p.
[28] Ricento, T. (2000): Historical and theoretical perspectives in language policy and planning. – Journal of Sociolinguistics 4(2): 196-213.
[29] Smith, J., Durham, M. (2011): A tipping point in dialect obsolescence? Change across the generations in Lerwick, Shetland. – Journal of Sociolinguistics 15(2): 197-225.
[30] Spolsky, B. (2021): Rethinking language policy. – Edinburgh: Edinburgh University Press 276p.
[31] Spolsky, B. (2009): Language management. – Cambridge: Cambridge University Press 308p.
[32] Spolsky, B. (2004): Language policy. – Cambridge: Cambridge University Press 250p.
[33] United Nations Educational, Scientific and Cultural Organization (UNESCO) (2024): World Atlas of Languages. – UNESCO 4p.
[34] Vorobyeva, T., Bel, A. (2021): Factors affecting language proficiency in heritage language: The case of young Russian heritage speakers in Spain. – Journal of Language Contact 14(2): 304-330.
[35] Wang, L., King, K. (2024): Language ideologies, language policies, and shifting regional dialect proficiencies in three Chinese cities. – Journal of Multilingual and Multicultural Development 45(6): 2166-2182.
[36] Yang, H., Liang, Z., Li, P., Liu, C. (1985): Chinese dialects in Guangxi (Draft). – Fangyan (Dialect) 3: 181-190.
[37] Yazici, B., Yolacan, S. (2007): A comparison of various tests of normality. – Journal of Statistical Computation and Simulation 77(2): 175-183.