

UNIVERZITA PALACKÉHO V OLOMOUCI

Filozofická fakulta

Katedra asijských studií

MAGISTERSKÁ DIPLOMOVÁ PRÁCE

**Cantonese-Mandarin Bilinguals from Guangdong: Accuracy and  
Response Time in Bilingual Lexical Decision**

Olomouc 2022, Kamila Liedermannová

Vedoucí diplomové práce: doc. Joanna Ut-Seong Sio, Ph.D., Gyu-Ho Shin, Ph.D.



Prohlašuji, že jsem svou diplomovou práci vypracovala samostatně a uvedla všechny použité prameny a literaturu.

V Olomouci dne ..... Podpis .....

**Name and surname:** Kamila Liedermannová

**Department and faculty:** Department of Asian Studies, Faculty of Arts

**Title:** Cantonese-Mandarin Bilinguals from Guangdong: Accuracy and Response Time in Bilingual Lexical Decision

**Supervisor:** doc. Joanna Ut-Seong Sio, Ph.D., Gyu-Ho Shin, Ph.D.

**Number of pages:** 75

**Number of characters:** 121 060 (with spaces)

**Number of bibliography titles:** 74

**Key words:** Cantonese-Mandarin bilingualism, written Cantonese, lexical decision task, Chinese varieties, Chinese languages

Cantonese is of the largest varieties of Chinese worldwide, and the second most developed one after Mandarin. It is mainly a spoken variety that is not recognized as an official language anywhere in the world and has no official standardized script. This thesis focuses on Cantonese-Mandarin bilinguals from Guangdong and their performance in lexical decision task, with an aim to find out if there is a difference between word recognition in written Cantonese and Mandarin lexemes and expressions. The results confirmed the assumption that Cantonese-Mandarin bilinguals showed higher accuracy in recognizing expression written in Mandarin than in Cantonese.



Acknowledgement:

I would like to express my sincere gratitude to my thesis supervisors,  
doc. Joanna Ut-Seong Sio, Ph.D., and Gyu-Ho Shin, Ph.D, for their help and guidance.

## Table of contents

|  |    |
|--|----|
| Terminology .....  | 4  |
| 1. 1 Introduction.....   | 6  |
| 2 Language Division of Chinese.....  | 7  |
| <b>2.1 Fangyan? Dialect? Language?</b> .....                               | 9  |
| 3 Language Policies in China .....   | 12 |
| <b>3.1 Imperial China</b> .....  | 12 |
| <b>3.2 20th century language policies</b> .....                            | 13 |
| <b>3.3 New pronunciation</b> .....   | 14 |
| 4 Creation of Putonghua .....  | 15 |
| <b>4.1 One common speech and its relationship to other varieties</b> ..... | 16 |
| <b>4.2 Standardized language and its regional varieties</b> .....          | 18 |
| 5 Cantonese.....   | 20 |
| <b>5.1 Yue 粵 Group Division</b> .....                                      | 21 |
| <b>5.2 Cantonese: Language Policy</b> .....                                | 22 |
| <b>5.3 The status of Cantonese</b> .....                                   | 23 |
| 6 Cantonese in Guangdong .....   | 26 |
| <b>6.1 School and Education: Language Policy in Classroom</b> .....        | 31 |
| Language of instruction.....   | 31 |
| Media reports about Guangdong’s language situation .....                   | 32 |
| Multilingualism in China .....   | 38 |
| Pro-Cantonese Protests 2010 .....  | 40 |
| 7 Written Cantonese .....  | 42 |
| <b>7.1 Vernacular Written Cantonese</b> .....                              | 45 |
| <b>7.2 History of Vernacular Written Cantonese</b> .....                   | 45 |
| <b>7.3 Cantonese-specific characters</b> .....                             | 45 |
| <b>7.3 Input methods and romanization systems</b> .....                    | 47 |
| <b>7.4 Typing strategies and mixing codes</b> .....                        | 49 |
| 8 Bilingual Lexical Decision Task.....                                     | 52 |
| <b>8.1 Recruiting respondents</b> .....                                    | 52 |
| <b>8.2 Questionnaire structure</b> .....                                   | 53 |

|   |    |
|---|----|
| <b>8.3 Selected respondents</b> .....             | 54 |
| <b>8.4 Lexical decision task in Chinese</b> ..... | 55 |
| <b>8.5 Cantonese words &amp; nonwords</b> .....   | 59 |
| <b>8.5 Mandarin word &amp; nonwords</b> .....     | 60 |
| <b>8.6 LTD Experiment Process</b> .....           | 61 |
| <b>Data Analysis</b> .....                        | 62 |
| <b>9 Conclusion</b> .....                         | 65 |
| Resumé.....                                       | 67 |
| <b>Bibliography</b> .....                         | 68 |

### **Editorial Note**

For romanization of Chinese, this work uses pinyin for Mandarin and jyutping for Cantonese. Simplified Characters are used for Chinese terms.

## **Terminology**

Cantonese – Yue 粵 language group

Fangyan – 方言 (*fāngyán*), synonymous to Chinese language variety

Standard Written Chinese – refers to written form of Standard Chinese, the official language of PRC

Modern Standard Chinese – 现代标准汉语 (*Xiàndài biāozhǎn Hànyǔ*), standardized language based on Mandarin varieties and official language of People's Republic of China.

Putonghua – the official spoken variety of Chinese used in the PRC as a common language, synonymous with *Mandarin*

Tuipu Policy – political campaign to promote and spread Putonghua

## **Abbreviations**

LSHK – Linguistic Society of Hong Kong

PRC – People’s Republic of China

SMC – Standard Modern Chinese

SWC – Standard Written Chinese

VWC – Vernacular Written Cantonese

SC – Simplified Chinese Characters

TC – Traditional Chinese Characters

## **1. 1 Introduction**

This thesis focuses on Cantonese-Mandarin bilinguals from Guangdong province in China. Cantonese and Mandarin are two Chinese varieties of the Sinitic group. Cantonese is mainly a spoken variety with an informal register, while Mandarin is considered the standard language of all Chinese. Both varieties are written down by Chinese script. Mandarin is represented by standardized characters; Cantonese is represented by standardized and Cantonese-specific characters. These characters are not standardized, but they are widely accepted in the community. Due to lack of standardization of written Cantonese and non-existing education in other Chinese varieties in China, native speakers acquire their knowledge of written Cantonese in informal domains, such as family, friends, internet and social media.

The aim of this thesis is to find out how Cantonese-Mandarin bilinguals recognize written Mandarin and Cantonese. We conducted a lexical decision experiment, where we analyze their performance base on response time and correct/incorrect responses in both Chinese varieties.

There has been a debate whether Cantonese and Mandarin are two separate languages or dialects. In the first part, I will list information about Chinese languages and their division to groups and subgroups to illustrate that they are rather separate languages than dialects.

The second part will explain current language policy in China and Promotion of Putonghua to clarify current China's position towards multilingual education and influence on bilingual Cantonese-Mandarin literacy.

Further, I will explain background information about Cantonese with a focus on its usage in Guangdong province in China.

In the last part, I present lexical decision experiment and its results.



## 2 Language Division of Chinese

Sino-Tibetan family (汉藏语系 *hàn zàng yǔ xì*) is the largest and the most significant language family in China. The majority of languages in China belong here and most of the population speaks of the languages belonging to Sino-Tibetan language family. Both Cantonese and Mandarin belong to this family. Other language families in China (Chappell, 2017) include: Altaic, Austroasiatic, Austronesian, Hmong-Mien, Indo-European, Koreanic, Mongolic, Tai-Kadai, Tungus-Manchu, Turkic and Creole languages (Macanese, Tang Wa Hua, Wutun etc.).

Sino-Tibetan family comprises two branches: 1) Chinese (Sinitic) languages 2) Tibeto-Burman languages

The centre of Chinese languages is China, where they spread from Manchuria in the northeast to Guangdong in the south. Majority of Chinese population speak languages that belong to this group. Chinese languages are divided into ten subgroups (Chapell, 2017). Tibeto-Burman languages are used in Qinghai, Sichuan, Tibet and Yunnan. (Chappell, 2013)

Chinese has an internal variation that is unprecedented in most other languages in the world (Kaltenegger, 2020). The standard model in language typology is not completely applicable to Chinese languages and even terms for language family (语系 *yǔxì*), language group (语族 *yǔzú*) or language branch (语支 *yǔzhī*) in Chinese are neologism based on corresponding Western terms from European languages (Mair, 1991).

Language family is presumed to derive from the same parent language, for instance: Indo-European, Ural-Altai, Sino-Tibetan. Language group is the next level below language family, for instance, within Indo-European language family: Celtic, Baltic, Germanic, Hellenic, Indian, Iranian, Romance and Slavic. This division can be also applied to Sino-Tibetan group: Chinese (Sinitic), Tibeto-Burman, Tai (Dai), Miao-Yiao, etc. Each language group consists of separate languages, for instance Belarussian, Bulgarian, Czech,

Macedonian, Polish, Russian, Serbo-Croatian, Slovak, Slovene and Ukrainian within Slavic languages (group); when applied to Sinitic group: Mandarin, Yue, Xiang, Hakka, etc. Each of these “separate languages” contains several other “sub-varieties”, so even if they are often referred to as languages, it is more accurate to describe them as “Chinese language groups”.

Language branch is another level of classification, for instance: East-Slavic (Belarussian, Russian, Ukrainian), South-Slavic (Bulgarian, Macedonian, Slovene, Serbo-Croatian) and West-Slavic (Czech, Polish, Slovak) of Slavic languages group. Branches of Sinitic (Chinese) languages have not been determined. (Mair, 1991)

Chinese languages division can be categorized as following:

Sino-Tibetan family (language family) => Sinitic (language group) => Mandarin, Yue, Wu etc. (Chinese language groups or Chinese varieties) => Gaoyang, Guangfu, Siyi (Subvarieties)

There are currently 10 recognize Chinese language groups within the Sinitic group (Chappell, 2013):

- 1) Gan 贛
- 2) Hui 徽
- 3) Hakka 客家
- 4) Jin 晋
- 5) Mandarin 北方方言
- 6) Min 闽
- 7) Pinghua 平话, Tuhua 土话
- 8) Wu 吴
- 9) Xiang 湘

## 10) Yue 粵

Not all researchers agree on this division. Some linguists consider *Jin* 晉 being a separate group, some of them include Jin within Mandarin Chinese group (Barov & Egorova, 2019). Some scholars also list *Dungan* 东干语 as a separate Chinese language group. (Mair, 1990)

### 2.1 Fangyan? Dialect? Language?

Standard models of categorization applied to other language families do not always apply to Sinitic languages. (Mair, 1991). Political, historical and cultural factors together influence attitude towards what is a language and what is a dialect. The term *fangyan* 方言 is often used to refer to separate language groups. The origin of this term is not clear, it is assumed they could be descendants 后代语 *hòudàiyǔ* of a parent language (Xing, 1982). Fangyan is a very old concept in China, but it has been contaminated by the Western concept of “dialect”, which makes the discussion about language situation in China confusing. (Mair, 1991; Kaltenecker, 2020). The term has developed and changed over history. In the late Qing Period, Western languages were described as fangyan in many texts, for instance in “Essentials of Government Zhou Ritual (周礼政要 *Zhōu lǐ zhèngyào*) by Chinese scholar and philologist Sun Yirang 孙诒让. (Mair, 1991)

This term is often translated as “dialect” in English, even though it does not fully correspond with characteristics of a dialect, because referring to fangyan as “dialect” does not precisely characterize its nature. There are too many differences among varieties or even their subvarieties on phonological, lexical, orthographical and grammatical level.

Dialects are two or more mutually intelligible varieties of a given language, distinguished by vocabulary, idioms and pronunciation. (Mair, 1991) Not all dialects are easily mutually intelligible, but at least partial mutual

intelligibility is often seen as a feature that makes two or more varieties to be considered dialects. When two or more varieties are dialects, it suggests that there is a superordinate unit, a language above them.

Language has much broader functions than a dialect does; it can be a tool of communication between two or more different dialectal groups, and it has much greater power and influence than dialects (Barov & Egorova, 2019). Every speaker of a language is a speaker of at least one dialect (Chambers and Trudgill 2004; Francis, 2016): a speaker of Mandarin speaks Beijing, Zhongyuan, Upper Yangtze, Northeastern Mandarin etc. Current division of Chinese languages is strongly influenced by the view and language policy of People's Republic of China. There's an official view, favoured by the People's Republic of China, but also an "unofficial view", based on established methods in the field (Francis, 2016).

The current language policy impacts both Chinese and Tibeto-Burman group, which many minority languages in PRC belong to. There are different criteria applied to division and categorization of these two language groups. Based on policy-neutral-procedure, if Burmese, Jingpo, Tibetan and Yi (mutually unintelligible) can be considered separate languages, so can Hakka, Mandarin, Wu and Yue. (Francis, 2016)

Meanwhile, separate Chinese language groups are often referred to as "dialects" in English, despite their mutual unintelligibility. As mentioned above, one of the reasons is an inaccurate translation of the term *fangyan*. Another reason is the influence of official PCR policy and attempts to downplay their status or possible independency on Standard Chinese Mandarin, sometimes also referred to as *Putonghua* (普通话 Pǔtōnghuà). In PCR, it is not uncommon to encounter an opinion that Standard Chinese Mandarin is a language, which other Chinese language groups like Hakka, Min or Yue are subordinated to. This attitude towards the difference between languages and dialects is often reflected in opinions of many non-linguists, as people generally confuse dialects for languages who are not completely standardized or developed (Barov & Egorova, 2019). This is mainly due to "*ideology of a standard language*", a prescriptive approach to language,

which claims there is one canonical and correct form of a given language (Milroy, 2001). According to this concept, the correct form of a language is defined externally, in grammar books (Milroy, 2001). Mutual unintelligibility between varieties means people are less likely to perceive them as “incorrect” forms of the standard variety (Liang, 2015).

The term *fangyan* is connected with many social stereotypes, such as lack of education and low-class (Guo, Ni, Wang & Zhang, 2022). However, all modern languages were once undeveloped or not standardized in their past. Languages can be developed and standardized over time. Labelling other Chinese varieties than Standard Mandarin as *nonstandard* is not accurate, because the relationship between them and Standard Mandarin is different from “standard-and-dialect” or “standard-and-nonstandard”. (Liang, 2015) If this could apply to Chinese varieties, it would imply that other Chinese varieties are subordinate to Standard Mandarin. If Standard Mandarin is a language and other varieties are its “dialects”, it assumes they belong to the same Chinese language group within the Sinitic group or to one language above them (Francis, 2019). This only applies to Mandarin language group, Modern Standard Mandarin belongs, but not for other groups such as Hakka, Min or Yue.

Sometimes, the term *fangyan* is also translated to as *language* in English, for instance: Cantonese language, Hakka language, Hokkien language, Teochew language. These Chinese varieties are often regarded as distinctive languages within the Sinitic group. (Mair, 1991) Another common term for *fangyan* is *local language* (Liu, 2016) or regional language, both correspond the direct translation of *fangyan*.

As *fangyan* does not mean “dialect” in the same sense as in it is often used in European languages; using the term “topolect” or “variety of Chinese” is much more suitable (Barov & Egorova, 2019). In this context, “topolect” stands for 方言: 方 *fāng*, the first character of *fangyan* 方言 is derived from *place* 地方 (*dìfāng*). It is common to refer to these varieties of Chinese as *fangyan* without translating the term and many researchers do so (Guo, Ni, Wang & Zhang, 2022; Liu, 2016; Mair, 1991).

## 3 Language Policies in China

### 3.1 Imperial China

China's attitude towards its language policy has not always been inclined towards monolingualism. Language diversity was not considered a problem in imperial China and there were no attempts to promote one common speech; the imperial language order relied mainly on the unified Chinese script (Zhou, 2019), which was seen as a factor that could ensure imperial law and order across all languages under Heaven (Zhou, 2018).

In Imperial China, *wenyan* 文言 (in English referred to as *Classical Chinese* or *Literary Chinese*) played an important role. *Wenyan* was a written form of Chinese which was used mainly by scholars and the bureaucratic system. It was very distant from any spoken form of Chinese languages and only an educated minority was able to use and understand it. *Wenyan* was considered supreme over another written form of Chinese, *baihua* 白话: *wenyan* was regarded as refined, elegant and ideal for higher functions, while *baihua* was deemed as vulgar, suitable for low-functions only (Chen, 2001).

At the end of 19th century, the idea of trinity of one language, one people and one state was introduced as a means to save the falling imperial Qing dynasty (Zhou, 2017). The process of creating Modern Standard Chinese first started in 19th century after Opium Wars (Kurpaska, 2019).

Imperial China fell in 1912, but the idea of “trinity of one language, one people and one state” has prevailed until now. China followed a global trend that was prevalent all over the world at that time; since the 19<sup>th</sup> century in Europe, nationalistic worldview was on the rise. Nationalism of 19<sup>th</sup> century became one of the most significant political and formative forces in history.

Before Western invasion, a Chinese national identity and a Chinese national language did not exist. With Europe's colonial expansion, the nationalistic trends expectedly came to China and language diversity was no longer desirable. Until the fall of Qing dynasty, only members of bureaucratic system were expected to use one standard Chinese variety *guanhua* 官话 but this was not expected from the rest of the population (Kurpaska, 2019; Zhou, 2019).

### 3.2 20th century language policies

The 20<sup>th</sup> century worldview stemmed from nationalistic ideas, so the imperial attitude to language policies which relied mainly on the unified Chinese script was not considered relevant. "The new era demanded a national standardized oral language; a new language order for the rejuvenation of Chinese nation". (Zhou, 2019; p. 70). *Wenyan* was assumed to be the main obstacle to a higher literacy rate, so *baihua* was chosen as the base for standardization. (Chen, 2001) One common standardized language seemed to be necessary for any country to become united and modernized.

In 1911, the Imperial Ministry of Education passed the Resolution on Methods of National Language Standardization 统一国语办法案 (*tǒngyī guóyǔ bànfǎ àn*), which included the five following major measures to standardize Chinese (Zhou, 2019):

- 1) setting up a national agency under the Ministry of Education with branches in every province to survey standard Chinese vocabulary, grammar, pronunciation, and its use,
- 2) compilation of national language textbooks, dictionaries, and comparative handbooks of Chinese dialects based on the data from the above survey,
- 3) choosing the standard pronunciation for the national language,
- 4) choosing the phonetic system for representing the standard pronunciation,

- 5) establishment of national and provincial national language training institutes to provide standard Chinese training to all teachers, use standard Chinese as the medium of instruction in schools throughout China.

Even though these measures take Standard Chinese in account, and it is mentioned in each one of them, they do not explain or define this concept. It can be assumed that the mentioned standard Chinese was *guanhua*, which was already used by the bureaucratic system, and therefore could be expectedly used as a standardized language. Only members of bureaucratic systems were able to speak *guanhua*, moreover, only a small proportion of non-Mandarin speakers spoke it on a sufficient level (Chen, 1993; Chen, 1999). The five major measures to standardize Chinese also did not define the relationship between standard Chinese and other Chinese varieties (Zhou, 2019).

The work of the language reformers started in 1911, when *General Committee for the survey of the National Language* 国语调查总会 (*guóyǔ diàochá zǒng huì*), which an aim to conduct large-scale research of the language situation in the country. The idea of national language *guoyu* 国语 was borrowed from Japanese (Kurpaska, 2019). The new modern standard language could have been based on the language used in Beijing, which would be an obvious choice. However, many scholars advocated that common language should be based on other language groups like Nanjing, Wuhan and Shanghai or multiple languages, especially the phonetics; pronunciation should include characteristics of all dialects (Kurpaska, 2019).

### 3.3 New pronunciation

The first standardized pronunciation, so called *old national pronunciation* 老国音 (*lǎo guó yīn*) included features that exist in Mandarin, Wu and southern varieties; it was “an artificial and complicated language which nobody could speak” (Chen, 2004, 18). In 1932, Beijing Mandarin was proclaimed by the



Ministry of Education to become model for the “new national pronunciation” 新国音 (*xīn guó yīn*). The same year, *Glossary of frequently used characters in National Pronunciation* 国音常用字汇 (*guó yīn chángyòng zìhuì*) was created: it contained 12,229 characters (Chen, 2004), each of them annotated in a special phonetic alphabet 注音符号 (*zhùyīn fúhào*) and *National Language Romanization* 国语罗马字 (*guóyǔ luómǎ zì*).

Many reforms were focused at Chinese characters, as “Chinese writing system developed many variations of usage over time” (Chen, 2004, p. 15). One of the major goals of written language reform between 1910 and 1920 was to reduce the gap between writing and speech, represented by a popular slogan “My hand writes down what I say with my mouth” 我手写我口 (*wǒ shǒu xiě wǒ kǒu*) (Chen, 1993).

#### **4 Creation of Putonghua**

After the establishment of People’s Republic of China in 1949, the government continued the reforms and worked on one common language. A new system for transliteration, pinyin, was designed and promoted. Pinyin has been used until now and it is used worldwide, for instance for learning Chinese as a second language.

In 1956, the name of the *Guoyu* 国语 was changed to *Putonghua* 普通话 (*pǔtōnghuà*, literally *common speech* or *common tongue*), to emphasize its closeness to all nationalities in China (Kurpaska, 2019). The definition of the standardized common speech said “Putonghua is the standard form of Modern Chinese with the Beijing phonological system as its norm of pronunciation, and Northern dialects as its base dialect, and looking to exemplary modern works in *baihua* for its grammatical norms”. (Chen, 2004, p. 24; Kurpaska, 2019).

Since that time, usage of Putonghua is widely promoted in a long-time campaign called “Promotion of Putonghua” (推广普通话 *tuīguǎng pǔtōnghuà*, shortened as 推普 *tuīpǔ*, literally “spread Putonghua”). Putonghua is regulated

by State Language Commission 国家语言文字工作委员会 (*Guójiā yǔyán wénzì gōngzuò wěiyuánhui*). The results of codification are published in The Contemporary Chinese Dictionary 现代汉语词典 (*xiàndài hànyǔ cídiǎn*); this dictionary was first published in 1978, the last edition is from 2016 (Kurpaska, 2019).

#### **4.1 One common speech and its relationship to other varieties**

Following the Soviet Union, China adopted the communist evolutionary language ideology in 1950s Zhou (2019). Evolutionary language ideology is based on the ideology of social evolution and believes in a parallel between social and linguistic evolutions: “societies gradually evolve from clans to tribes in feudalism, from tribes to nations in capitalism, and finally to one people in communism” (Zhou, 2019, p. 48). This ideology believes, that languages are anticipated to follow the same paths: “from many clan and tribal languages to fewer national languages, from national languages to single one spoken by one people in the communist society” (Zhou, 2019, p. 48).

According to Guo (2004), in 1958, Zhou Enlai 周恩来, the then prime minister of China, stated in his report *Current tasks of script reform* 当前文字改革的任务 (*Dāngqián wénzì gǎigé de rènwù*):

“The promotion of Putonghua has as its goal the removal of the barrier of dialects, not the prohibition or abolition of dialects. Does the promotion of Putonghua mean to prohibit or abolish the dialects? Of course, not. Dialects will exist for a long time. They cannot be prohibited by administrative order, nor can they be abolished by artificial measures. In the promotion of Putonghua, distinctions should be made between old and young people; between activities on a national scale and those of a local nature; between the present and the future. There should be no overgeneralization. On the other hand, those who can speak only Putonghua should learn local dialects so that they will be able to make close contact with the working people in dialect communities.”

This attitude of Zhou Enlai shows much more permissive approach to other varieties of Chinese, but this did not represent the common stance of that time. Mandarin was expected to wipe out other Chinese varieties for the sake of unification (Barov & Egorova, 2019; Guo, 2004; Zhou, 2019), but this view changed in 1980s. In 1986, State Language Commission came under the authority of State Educational Commission 国家教育委员会 (*guójiā jiàoyù wěiyuánhui*); it was decided that pinyin would be used as a tool for learning Chinese and Putonghua was to become language of instruction in all school. (Rohsenow, 2004; Spolsky, 2014). In the same year, Putonghua Proficiency Test was proposed: it was divided into three levels, and in 2000, it was included in the National Law of Standard Spoken and Written Language 中华人民共和国国家通用语言文字法 (*Zhōnghuá rénmín gònghéguó guójiā tōngyòng yǔyán wénzì fǎ*).

In the 1990s, after the collapse of the Soviet Union, the PRC abandoned the communist evolutionary language ideology and replaced it with an “integrationist language ideology”: the goal of this ideology is to “integrate all Chinese citizens into the Han mainstream” (Zhou, 2019, p. 48).

The integrationist language ideology includes so-called “Three Cannots” (三个离不开 *sān gè lì bù kāi*) minorities cannot live without the Han, the Han cannot live without the minorities, and the minorities cannot live without each other. (Zhou, 2019, 48) It also includes the idea of linguistic citizenship. (Zhou, 2016) Linguistic citizenship emphasizes the unifying role of the common speech in Chinese citizens’ identification with the PRC as a state, the inclusive Chinese nation and Chinese culture: learning the common speech is an obligation connected to citizenship (Zhou, 2015) Chinese citizens are expected to have a command of Standard Chinese 现代标准汉语 (*xiàndài biāozhǔn hànyǔ*), the official language of PRC, and use it in both spoken and written forms. Every September since 1998, *Putonghua Promotion Week* 推广普通话宣传周 (*tuīguǎng pǔtōnghuà xuānchuán zhōu*) is launched (Wang & Yuan, 2013).

Putonghua is often presented as a language of progress, language of 21<sup>st</sup> century. Speaking it fluently is necessary to become a proper Chinese citizen. This type of promotion downplays role of other Chinese varieties and suggests those who cannot speak Putonghua are uncivilized. From government's point of view, "Putonghua must be predominantly used in all public domains while minority languages and Chinese varieties are considered transitional between home and school, and are used at home and in few assigned domains" (Zhou, 2019; 49).

#### 4.2 Standardized language and its regional varieties

In a broader sense, *Mandarin* refers to the largest Chinese language group in China and worldwide. Mandarin Chinese language group contains 7 varieties:

- 1) Beijing Mandarin 北京官话 (*Běijīng guānhuà*)
- 2) Central Plains Mandarin 中原官话 (*Zhōngyuán guānhuà*)
- 3) Jilu Mandarin 冀鲁官话 (*Jì lǔ guānhuà*)
- 4) Jiaoliao Mandarin 胶辽官话 (*Jiāo liáo guānhuà*)
- 5) Lanyin Mandarin 兰银官话 (*Lán yín guānhuà*)
- 6) Lower Yangtze Mandarin 下江官话 (*Xià jiāng guānhuà*)
- 7) Southwestern Mandarin 西南官话 (*Xīnán guānhuà*)
- 8) Modern Standard Chinese 现代标准汉语 (*Xiàndài biāozhǔn hànǚ*)

In a narrow sense, Mandarin refers to Modern Standard Chinese, the official language of PRC and a lingua franca among speakers of all Chinese varieties. It is based mainly on grammar, lexicon and phonetics of Mandarin Chinese group, especially the Beijing variety. The terms *Putonghua* and *Mandarin* are considered synonymous; they usually refer to the same variety, the common

standardized speech. While *Mandarin* rather refers to its spoken form, *Putonghua* is used in a relation to government language policies and in formal context.

Mandarin is often considered superordinate to the whole due to its position in Chinese society as a common tongue, although it is one of Mandarin group varieties. There are also terms *Huayu* 华语 (*Huáyǔ*, literally *language of ethnic Chinese*) or *Guoyu* 国语 which are used as a synonym for Modern Standard Chinese. The terms 中文 (*Zhōngwén*) and 汉语 (*Hànyǔ*), which refer to Sinitic (Chinese) language group as whole, are often used as another term for the common tongue.

Mandarin (Putonghua), as an artificial language, did not have any native speakers in the 1950s (Zhou, 2019). Even though the common tongue is based on Mandarin varieties, these varieties have many specific features which do not exist in the common standardized language. At the beginning of Promotion of Putonghua, even speakers of Mandarin varieties did not speak the “proper Putonghua”.

This situation has changed with the younger generation. Nowadays, there are many speakers in the northern China who do not speak any other Mandarin varieties, only Putonghua. Due to its proximity to Mandarin varieties, it is often the chosen code in various environments in northern cities and speakers of Mandarin varieties easily become fluent speakers (He, 2006).

This is different in southern China, as southern varieties are very distant from Mandarin in terms of lexicon, phonology and syntax. Especially Cantonese speakers find it difficult to master the standardized variety (Pan, 2000). Speakers of Mandarin varieties learn the common tongue as a received pronunciation, while speakers of other varieties learn it as a second language (He, 2006). Therefore, Mandarin is less likely to become the chosen code in southern China, which was the reason for the government to promote Mandarin more intensively in these regions.

Due to differences between Chinese language groups and diverse background of Chinese population, there is no one kind of Mandarin, but rather diverse

regional versions influenced by local varieties called *regional Putonghua* 地方普通话 *Difāng pǔtōnghuà* (Zhao & Liu, 2020). Recently, these regional varieties have risen extensively as its promotion intensified (Saillard, 2004).

There are varieties like Beijing Putonghua 北京普通话, Tianjin Putonghua 天津普通话, Shanghai Putonghua 上海普通话 and Chengdu Putonghua 成都普通话. Southern regional varieties have been rather notional than real (He, 2006), due to lack of usage in the public space. As *tuipu policy* became more intense and more migrants started to arrive in southern provinces, especially Guangdong, local Mandarin varieties started to emerge.

As of September 2020, Ministry of Education states that 80.72% of population speak Mandarin. According to Department of Language Application and Administration, only 7% can speak the standard variety (Zhao & Liu, 2020). This means that approximately 72.28% of Mandarin speakers use regional varieties. We can estimate that the 7% are more likely younger generation speakers of Mandarin varieties, who easily acquire the standard form.

Despite huge number of Regional Putonghua speakers, the emergence of this phenomena is seriously under-researched; it is often treated as a transitional form and from nonstandard to standard or a by-product of language contact (Zhao & Liu, 2020).

## 5 Cantonese

Among Chinese varieties, Cantonese has the most attention from linguists and researchers. Cantonese is spoken in overseas diaspora and Guangdong, which has been one of the most economically dominating provinces since the 1980s. Besides that, Guangdong province has also geographical, cultural and historical ties to Hong Kong. Cantonese has a strong position among Chinese varieties (Barov & Egorova, 2019) and it is one of the few varieties with

developed script. It is considered the most influential regional variety in China (Xiong, 2018).

The area of Guangdong, Guangxi, Hong Kong and Macau is referred to as *Lingnan* 岭南. The term *Lingnan* culture is often used for Cantonese culture specifically. In a broader sense, it refers to culture of the whole Lignan region and also Hakka, Hainanese, Taishanese and Teochew cultures. The two regions *Guangdong* 广东 and *Guangxi* 广西 are called *Liangguang* 两广. While around 90% Cantonese speakers in Guangzhou, the capital of Guangdong are bilingual (Li & Shun, 2018), Nanning 南宁, the capital of Guanxi, has already lost a part of its Cantonese heritage.

### 5.1 Yue 粤 Group Division

Cantonese (Yue 粤), as most Chinese varieties, is not a single language, but a language group within Sinitic languages. According to Kwok, Chin & Tsou (2016), it is divided into seven subgroups:

- 1) Guangfu 广府
- 2) Siyi 四邑
- 3) Gaoyang 高阳
- 4) Wuhua 吴化
- 5) Goulou 勾漏
- 6) Yongxun 邕潯
- 7) Qinlian 钦廉

Guangfu subgroup is considered the most prestige variety. It stretches over a wide area from the Pearl River Delta to Wuzhou 梧州 in Guangxi 广西

province. Yue dialects used in Guangzhou 广州, Zhongshan 中山, Dongguan 东莞, Zhaoqing 肇庆, Yunfu 云浮, Shaoguan 韶关, Hong Kong and Macau are all considered to be varieties of this subgroup.

Siyi subgroup is spoken in the western part of the Pearl River Delta, such as Taishan 台山, Xinhui 新会, Kaiping 开平 and Enping 恩平. Siyi subgroup is distinct from other subgroups due to its grammatical peculiarities.

Gaoyang subgroup covers a huge area in southwestern coastal Guangdong; Yangjiang 阳江, Maoming 茂名, Xinyi 信宜, Zhanjiang 湛江 and Lianjiang 廉江. This subgroup is heavily influenced by the surrounding Min and Hakka at phonological, lexical and syntactic level.

Wuhua subgroup remains poorly known and occupies a relatively small area in southwestern Guangdong including Huazhou 化州 and Wuchuan 吴川. It is surrounded by Gaoyang subgroup.

Goulou subgroup is spoken in northwestern Guangdong; Fengkai 封开, Huaiji 怀集, Yangshan 阳山 and southeastern Guanxi; Yulin 玉林, Bobai 博白 and Cenxi 岑溪. Especially Yulin dialect has a historical and typological significance due to its specific negation forms and the diminutive tone sandhi.

Yongxun subgroup is located in the Yong Yu Xun 邕郁潯 valley in the central and western parts of Guangxi, including Nanning 南宁, Baise 百色, Chongzuo 崇左 and Longzhou 龙州. Yongxun subgroup shares the same origin with Guangfu. As a result of contact with Zhuang, it has acquired many non-Han features.

Qinlian 钦廉 subgroup is used in the southern coastal area in Guangxi, the Qin-Lian subgroup includes the Yue dialects spoken in Beihai 北海, Lianzhou 连州 and Qinzhou 钦州.

## 5.2 Cantonese: Language Policy



According to Sautman, Barry & Xie (2021), Cantonese has over 80 million native speakers. Despite that is not an official or state language anywhere in the world (Cheng & Tang, 2014). However, it has a strong position in Hong Kong, which is considered contemporary Cantonese's capital; 90% of ethnic Chinese population of about 6.5 million speak it daily (Bauer, 2016). Cantonese in Hong Kong has acquired an extraordinary status due to its daily usage and existing written form. Although Cantonese is widely used in both spoken and written form in Hong Kong, there is no officially recognized written standard form of Cantonese (Yan, 2008).

### 5.3 The status of Cantonese

According to PRC's official policy, Cantonese is recognized as a *fangyan* 方言; when it comes to ethnic minorities, their right to learn and speak their languages are protected by law (Feng & Adamson, 2015), while varieties of Chinese do not have the same level of protection. This is mainly because they are considered to be subordinate to Modern Standard Chinese 现代标准汉语.

In 2013, Song Xinqiao 宋欣桥, a consultant at Chinese University of Hong Kong's Centre for Research and Development of Putonghua Education, wrote an article related to promotion of Modern Standard Mandarin. The article was titled "*On the Nature and Development of Standard Mandarin Education in Hong Kong*"<sup>1</sup> and was published in a set of 25 articles the website of Hong Kong Education Bureau. The aim of these articles was to promote Putonghua education in Hongkong.

In the article, Song Xinqiao claims that Cantonese cannot be a "mother tongue", it can be only a "mother topolect". Chinese linguistics indeed distinguishes these two different terms for mother tongue 母语 (*mǔyǔ*) and mother topolect 母言 (*mǔyán*) (Barov & Egorova, 2019).

---

<sup>1</sup> In original SWC 淺論香港普通話教育的性質與發展 *Qiǎn lùn xiānggǎng pǔtōnghuà jiàoyù dì xìngzhì yǔ fāzhǎn*

He further maintains “*We should emphasize that Standard Mandarin and Cantonese are not two languages, they belong to the same language, that is Chinese.*”<sup>2</sup> He uses the term hanyu 汉语, which refers to the Sinitic group (Mair, 1991). Some Chinese linguists consider Cantonese to be a variety of Modern Standard Chinese 现代标准汉语. (Mair, 1991) If hanyu refers to a language group, then Cantonese and Putonghua must be languages 语言, not varieties 方言. It is not clear if he refers to “Sinitic” as a branch of Sino-Tibetan languages or Modern Standard Chinese. Sinitic branch contains 10 separate Chinese language groups, so referring to it as *language* is the same as calling Tibeto-Burman *a language*. They both contain various language groups. The same applies for the term “mother tongue” 母语: if Sinitic and Tibeto-Burman are branches of Sino-Tibetan family, it makes no sense to refer to them as “mother tongue”.

UN defines mother tongue as "the language usually spoken in the individual's home in his or her early childhood". It would make sense to regard Cantonese as mother tongue, if a person speaks it at home from early childhood. However, the term “mother tongue” is sometimes not used for a person’s first language, but for the language of their ethnic group (Love & Ansaldo (2010); Davies (2003)). When Song Xinqiao refers to hanyu as a mother tongue, he does so based on an argument that Chinese have their own mother tongue as an ethnic group, not as individuals.

Is it possible for Modern Standard Chinese 现代标准汉语 to be a mother tongue of whole ethnic group? Standard Chinese is Standard Chinese is not a parent language that would naturally unify all Chinese, but a language created by with an aim to unify China, which has started to acquire speakers first in the 20<sup>th</sup> century (Liang, 2014). Chinese varieties are unified by Chinese script to a certain extent, even though many varieties have their own specific characters and differ in written form.

---

<sup>2</sup> In original SWC 我们应该强调普通话和粤语不是两种语言，两者同属于一种语言，那就是汉语。Wǒmen yīnggāi qiáng diào pǔtōnghuà hé yuèyǔ bùshì liǎng zhǒng yǔyán, liǎng zhě tóng shǔyú yīzhǒng yǔyán, nà jiùshì hànyǔ.

The article was removed after sparking criticism online. In fact, author's opinion is very widespread among Chinese linguists and it represents the current PRC policy towards other varieties of Chinese. According to Xiong (2018), "Cantonese being mistaken for a dialect is due to a Stalinist language ideology that an independent nation must construct a unified national language and other languages must be subordinate to it".

In 2014, Hong Kong Education Bureau 香港政府教育局 published a similar statement in English on their website: "Cantonese is a Chinese dialect not considered an official language". As in the previous case, it was removed later and Education Bureau had to apologize (Cheng, 2020; Tam, 2018).

These statements raised the two questions that were debated over following months: "Is Cantonese (just) a Chinese dialect? Is Cantonese *not* an official language?" (Cheng & Tang, 2014) According to Xing (1982), there are three conditions for a Chinese variety to be a *fangyan* rather than a separate *language*:

- 1) Sharing a common standard language
- 2) Sharing the same script
- 3) Being able to converse directly or with a bit of effort

The first condition, "sharing a common standard", indicates there should be a common standard language that unites other varieties. Mandarin (Putonghua) is considered to be the common standard language for all Chinese speakers. However, the common standard is not above other varieties, it is rather one of the varieties. Some participants who filled in my research questionnaires listed *Guoyu* 国语 and *Putonghua* 普通话 as *fangyan* 方言, not as a language; this is a notion which is often shared by Cantonese speakers. Most linguists also list the standardized common speech as one of Chinese varieties, not as a language above other Chinese varieties. Cantonese group is at the same level as Mandarin group, both are language groups subordinate to Sinitic languages. This would indicate that Cantonese is rather an independent language than a *fangyan*.

The second condition, “sharing the same script”, also does not completely apply to Cantonese. Although Cantonese uses standardized Chinese script, it has also developed its own written vernacular variety, *Vernacular Written Cantonese*. This script variety has its own history and contains specific characters that do not appear in standardized script. Even though Cantonese can be written in *Standard Written Chinese characters*, some expressions can be written only in Vernacular Written Cantonese.

The third condition, “being able to converse directly or with a bit of effort” hardly applies to Cantonese. Cantonese speakers and speakers of other varieties, especially speakers of northern varieties, are not able to converse directly or with a bit of effort, unless they already know each other’s language.

## **6 Cantonese in Guangdong**

Guangdong is now the most populous province in PRC; its capital Guangzhou is one of the biggest cities in China. Guangzhou variety of Yue 粵 belongs to the Guangfu 广府 subgroup. It is considered the most prestigious and standard form of all varieties; the prestige of a linguistic variety is influenced by its degree of development and the ideological significance of such development (Liang, 2015). Guangfu 广府 subgroup is the most developed variety due to its usage in Hong Kong, where it developed significantly in both spoken and written form. Guangdong is considered a stronghold of Cantonese as well, although the language situation in the province is far more complex. Beside Cantonese, Guangdong’s population also speaks Hakka, Hokkien or Teochew.

Due to fast development and job opportunities, Guangdong has attracted a lot of migrants from other provinces, who bring their own varieties of Chinese, which has an impact on local language situation. Even though the usage of Cantonese has declined, and there are less and less people who are able to speak it, Cantonese has maintained its significance within Guangdong society. When migrant workers cannot speak Cantonese, it limits their job

opportunities, and this applies mainly for unskilled job workers with lower qualifications; they might be even paid less than locals (Wang, 2017). Without Cantonese, they cannot fully adjust to the local life and take part in social activities. Even with local ID, they cannot receive respect from local people, so many of them choose to learn Cantonese (Wang, 2017). Being able to speak Cantonese is not necessary to live in Guangdong, but it plays an important role as a marker of identity. According to Li, Li & Gao (2019), 81,9% of their respondents would like to maintain or learn Cantonese despite having a different mother tongue.

In Guangzhou, as a big city, the necessity to learn Cantonese might be lower; there are many workers hailing from other provinces, who do not speak Cantonese, even though Cantonese still maintains very strong position among locals, who attach symbolic power to it. They see it as not just as a tool of communication, but as a tool of evaluation and a symbol of social identity, which helps them classify people (Wang, 2017). Due to Cantonese domination over other dialects and its identity-shaping influence, migrants from other provinces are motivated to learn Cantonese. Being able to speak Cantonese well grants them privileges over those who speak little or no Cantonese. Even though nowadays most Guangdong locals can speak Putonghua, for a long time, there was an undercurrent of resistance to learn Putonghua (Pan, 2000).

Wang Wei (2017) divides people in Guangdong province to these demographic classes, based on their connection to Cantonese:

- 1) Native speakers
- 2) CSL speakers
- 3) Mandarin speakers
- 4) Silent class

- 1) Native speakers usually have the highest knowledge of Cantonese and they are at the top of the social ranking in Guangdong. They can participate in all social activities, and the fact they are bilingual makes

it easy for them to have social connections in both Cantonese and Mandarin speaking environments.

- 2) CSL (Cantonese as a second language) speakers represent a large middle class in Guangdong. They have the second highest knowledge of Cantonese. Most of them are migrants who married local people, workers from specialized industry or students. They have an opportunity to speak Cantonese and improve their fluency within their social circle. They have social connections in both Cantonese and Mandarin speaking environment and can participate in local life.
- 3) Mandarin speakers are migrants who try to make their living in Guangdong, most of them belong to working class. Some of them learn Cantonese successfully and become CSL speakers. In case they fail to learn Cantonese and communicate fluently in it, they will not be able to have the same social connections as native speakers and CSL speakers. Their participation in local activities is limited and they might even experience unequal treatment.
- 4) Silent class represent a class of people who have little or no contact with Cantonese. Most of them are rural migrants who speak little or no Standard Mandarin, they only speak their mother tongue fluently. They usually work in hard, dirty and risky environments with a low salary, and take jobs the locals refuse. Local people and other social classes often look down on them and they can easily experience discrimination. Not being able to speak Cantonese or Standard Mandarin makes their social connections very constrained, as they are able to communicate fluently only with rural migrants who speak the same mother tongue.

Among adult locals, there's a notion of "pure Guangzhou local people", which refers to those whose family had lived in Guangzhou for generations and spoke Cantonese as a native language (Liang, 2015). Nowadays, with the population boom of immigrants from other provinces, Mandarin has become lingua franca not only among speakers of other languages, but also among Cantonese speakers (Chen, 2011), but due to the association of Mandarin with poor migrants, Mandarin is used only as a communication tool. It does not

have the same prestige as Cantonese among natives of Guangdong. As Cantonese is strongly associated with class and social status, affluent families from Guangdong are more inclined to preserve Cantonese culture and support their children in learning Cantonese (Wang & Ladegaard, 2008). People from “interdialectal” families are reluctant to share information mixed identities to avoid being judged as *outcomers* or being judged for their linguistic competence in Cantonese (Liang, 2015). Some people only see the dichotomy of the two varieties, Mandarin and Cantonese, so natives from interdialectal marriages are often seen as Mandarin speakers. This is not unique to Guangdong; communities who identify with Lignan 岭南 culture are more likely to distinguish themselves from other communities in China (Yan, 2008). Mandarin is on the rise in Guangdong, and more people speak it now than ever before, it is clear that Cantonese still maintains its significance in society. Cantonese also becomes more popular in other provinces in China, thanks to the economic success of south China regions (Zhu & Chen, 1991).

Prior 1990s, Mandarin speakers from other provinces often complained about unfriendliness and even hostilities they encountered in Guangzhou service industry, e.g. ignoring Mandarin-speaking customers (Pan, 2000). Guangzhou service staff was known for their “excluding outsiders” (排外 páiwài) attitude (Pan, 2000). In 1990s and early 2000s, it was possible to come across Mandarin speakers, but usage of Mandarin was limited to as school and workplace (Pan, 2000), even though the government wanted Mandarin to be used in all public domains (Zhou, 2019, 49). Cantonese still prevailed in public space and was used in all domains (Wang & Ladegaard, 2008). The anti-Mandarin attitude has changed due to increase of migrants from other provinces, who not only work in service industry, but also come as customers. Nowadays, staff in Guangzhou service industry is expected to speak Mandarin. Businesses are motivated to attract customers; language is one of the means to achieve that goal (Pan, 2000)

The linguistic landscape and usage of languages in Guangdong has changed rapidly during the last 30 years. In 1990s, even staff in state-run stores did not speak Mandarin (Pan, 2000). Nowadays, it is rare to encounter staff who speaks only Cantonese. Negative attitudes towards “northerners” have almost

disappeared, even though the upward trend of using Mandarin more than Cantonese is still unwelcome among some locals.

Negative attitude towards Putonghua has several historical and cultural reasons. In the early 1950's, leading posts were taken by military people, who were mostly speakers of northern Chinese variety that forms the basic pronunciation and syntactic structure for Putonghua. (Pan, 2000) So to this degree, Cantonese is associated with political power, but also with northerness and backwardness (Pan, 2000), as the north is less economically developed than south. Until now, northerners might be called *baklou* 北佬, *bakgu* 北姑 or *bakmui* 北妹, especially on internet forums and social networks (Gao, 2012). Depending on context, all of these terms might have slightly derogatory meaning. They are not always necessarily used for *northerners*, but for *outsiders* in general, such as people, who promote Putonghua instead of Cantonese; they are often blamed for the decline of Cantonese (Gao, 2012). Cantonese speakers also use the mocking term *baodonggua* 煲冬瓜, literally *claypot winter melon*, which is the Cantonese homonym for Putonghua 普通话.

Compared to other provinces or cities, Guangdong had much softer implementation of *tuipu* policy for economic reason (Cheung, 2002); due to the strong local sense of identity, Cantonese is often seen as a stronghold against *tuipu* policy. That might be one of the main reasons why softer *tuipu* policy was implemented, and the implementation was not very successful for a long time. With new waves of migrants coming from other provinces, it has been under debate whether Guangzhou is still Cantonese-dominant (Liang, 2015).



## 6.1 School and Education: Language Policy in Classroom

### Language of instruction

Until 1980s in Guangdong, proficiency in Mandarin was extremely limited (Zhang, 2001). In reality, Cantonese was even used as a language of instruction in classes in 1980s (Li, Li & Gao, 2019), even though the government was already enforcing *tuipu* policy at that time. Usage of Cantonese as a language of instruction in classroom became more limited in 1990s and 2000s, and teachers below the age of 40 at all levels have to pass a compulsory test in Putonghua (Wang & Ladegaard, 2008). In February 1992, *Decision to vigorously promote Mandarin* 关于大力推广普通话的决定 (*Guānyú dàlì tuīguǎng pǔtōnghuà de juédìng*) was issued by the provincial government in Guangdong: it stipulated, that by the end of 1995, Mandarin should become the only language used in kindergartens, primary schools and middle schools. (Ho & Lu, 2019). Despite these measures, Cantonese have been used as language of instruction at some schools even in the 2000s. Even if teachers used Mandarin as a language of instruction, students were still allowed to use Cantonese in their free time.

Since the mid-2000s, *tuipu* policy have become more restrictive (Ho & Lu, 2019). Cantonese have become unwelcome or even Some of the recent measures regarding language policy and promotion of Mandarin have caused uproar among natives of Guangdong. In the 21<sup>st</sup> century, Mandarin is considered to be the only acceptable language of instruction and all teachers and students are expected to be proficient speakers.

Decisions regarding language of instruction in classroom have had a significant impact on language use within families. Many parents are worried about their children's prospects in school. School results play an important role in highly competitive Chinese society, and families are willing to take all necessary measures to help their children succeed in school and later in life.

Language policies regarding public space also have a serious impact on private and family life. Some families choose to minimalize the usage of local

languages at home and use Mandarin instead. This is not an issue that would be specific in China, a trend of giving up local languages and using state languages prevails all over the world. Local languages and dialects are also considered to be little economic and social value, so families rather abandon local dialects, because they believe it could hurt their child's future prospects. Cantonese might have a little economic value in eyes of many families, but it definitely has a strong social value among local, so many parents still try to speak Cantonese with their children and they want their children to preserve culture connected to their family heritage.

Children from Cantonese-speaking families do not always speak Mandarin fluently from the very beginning, which is often not taken into consideration by the school system. Sometimes, the consequences are very traumatic, because teachers treat their students as if they all had the same ability in Mandarin and place the responsibility on children's parents Liang (2015). Parents who come across this issue avoid speaking Cantonese or other languages with their children.

### **Media reports about Guangdong's language situation**

In 2008, Chinese reporter He Xuehua 何雪华 wrote an article called *Many Guangzhou children cannot speak Cantonese* 好多广州细路唔识讲白话 (*hou2 do1 gwong2 zau1 sai3 lou6 m4 sik1 gong2 baak6 waa2*), which was published in Guangzhou Daily 广州日报. The title itself was typed in Cantonese, the article was written in Mandarin. Other follow-up articles also used titles written in Cantonese (Liang, 2015). He Xuehua talked to 32 people in Guangzhou, including students, economically active people and elderly people.

Only 9 of them used Cantonese on daily basis, the rest 23 used Mandarin and rarely spoke Cantonese, some of them claimed they can understand Cantonese, but they cannot speak it or do not have an opportunity to speak it. Some respondents moved to Guangzhou because of their job, but they did not feel the necessity to learn Cantonese, as Mandarin was used as a language of

instruction in class and also in their workplace. The article also discusses the issue mentioned above – many children do not speak Cantonese at all or their knowledge is very limited.

He Xuehua visited a class in a primary school in Guangzhou, and found out of 43 pupils in class, 10 children could comprehend Cantonese and spoke it fluently, 20 children did not have any knowledge of Cantonese. This was also mentioned by some respondents: many classmates do not speak Cantonese, and use Mandarin together, so children do not even have an opportunity to learn or practice Cantonese. With an upward trend of migration from other provinces, there are less and less children from Cantonese-speaking families, which makes it more complicated to encounter Cantonese in school environment. This also applies for children from multilingual families. One elderly citizen called mentioned difficulties when talking to his grandson, whose mother is not a local Guangdonger and speaks Mandarin only. From his comment, it was apparent he struggled to use Mandarin with his grandson, which is common for many older residents who predominantly use Cantonese.

Another respondent mentioned his wife was a native from Guangdong, but she used Mandarin with him to make their communication more comfortable for him. Therefore, their children were not able to learn and speak Cantonese as well. Tuipu policy is another crucial factor which makes it difficult for children to encounter Cantonese; even if they had an opportunity to hear Cantonese from classmates, stricter language policies demand everyone to use Mandarin in classroom.

With an attempt to reverse this tendency, one Guangzhou school organized a campaign called Cantonese Day Pilot Class 广州话日试点 (*Guǎngzhōu huà rì shìdiǎn bā*). According to He Xuehua's report, this class was set up every Friday, with an aim to help pupils learn more Cantonese. All lessons were taught in Standard Chinese Mandarin, which is the only officially allowed language of instruction, but between classes and after class, pupils used Cantonese. Their teacher also took part in this activity, so children could practice Cantonese with each other and their teacher. As many classmates did not speak Cantonese well or at all, they were allowed to use Mandarin, but the

maximum was 20 Mandarin sentences per day. After several weeks, their Cantonese improved and they were able to use their classmates' names in Cantonese. "Cantonese Day" seemed very successful. Both children and teachers interviewed in He Xuehua's survey expressed positive attitude towards this campaign.

Three months later, Liang Sihua 梁斯华 wanted to conduct a survey related to Cantonese Day at the same school, so she contacted the school headmistress. However, their attitude was completely different than it appeared before. The school officials accused the report about Cantonese Day by He Xuehua to be dishonest and exaggerated (Liang, 2015). They even had to explain this matter to the municipal officials and local education authority under Municipal Bureau of Education expressed negative attitude towards Cantonese Day activities (Liang, 2015). During one meeting of Liang Sihua with the school officials, the whole activity was downplayed as a small part of activities to introduce Guangzhou culture for support of 2010 Asian Games (Liang, 2015). In the end, the headmistress refused to provide any possibility for a survey and explained it would be against the law if Cantonese was promoted at school instead of Putonghua, as national laws require that Putonghua must be promoted at school (Liang, 2015).

The school's contradicting attitude could have been expected, given the pressure from local education authority. The fact that the school even organized a "pro-Cantonese" activity is rather surprising, because there has been an opposite trend among schools in Guangdong. Several schools have banned usage of Cantonese in classrooms and campus area. This happened also in Zhixin South Road Elementary School 执信南路小学 (*Zhí xìn nánlù xiǎoxué*) located in Yuexiu District 越秀区 (*Yuèxiù qū*).

In 2010, Guangzhou newspaper Yangcheng Evening News 羊城晚报 (*Yángchéng Wǎnbào*), also known as Ram City Evening Post, published an article titled "*Zhixin Nan Lu Primary School in Yuexiu District requires students to speak Putonghua whether it is during or after class*".

According to the article, Zhixin South Road Elementary School applies very strict language policy. The school has their own motto "*Putonghua is the*

*language of the campus!*” Although this is not a written rule, the school demands regarding this matter are uncompromising. Children are required to speak only Standard Mandarin both in class and after class. Once a pupil is found to speak Cantonese, his name is handed over to a teacher and your credits will be deducted. This can influence their results and the pupils is not allowed to become a class leader. School staff is encouraged to supervise that children abide by these language policy.

The article describes a story of Mr. Zeng 曾, whose daughter Xiaowen 小雯 is a pupil of Zhixin South Road Elementary School. He and his family are natives of Guangdong and his mother, the child’s grandmother, does not speak and understand Mandarin. Mr. Zeng’s daughter used to speak Cantonese, but due to the strict school regulations, she is afraid it would negatively affect her results and she could be subjected to criticism at her school.

This situation causes many disputes within the family. Grandmother often complains of many misunderstandings that happen with her granddaughter. According to the article, she even accused her son and daughter-in-law of not teaching their child Cantonese and forgetting their roots. Grandmother’s lack of knowledge of the common speech is very common for older Guangdong citizens. Even when all family members are able to speak Cantonese, Mr. Zeng’s daughter proficiency in Cantonese decreased substantially and she has no opportunities to practice it due to school regulations. As children in China spend most of their time at school, school policies influence their language policies significantly. Even if Mr. Zeng and his family would encourage Xiaowen to speak and practice Cantonese, she still would not be able to reach a high level of proficiency.

Another story in the article mentions Mr. Lin 林 and his family, who are not natives of Guangdong. They claimed that due to no knowledge of Cantonese, they have missed a lot of job opportunities, so they hoped their children would be able to speak Cantonese fluently from early age to avoid possible difficulties with job seeking. According to Mr. Lin, school rules are too strict. He maintained language policy should not be applied in children’s spare time. A lot of parents quoted in the article have similar thoughts: the best for their

children is to be bilingual and speak both Cantonese and Mandarin. They agreed that Mandarin-only-policy should be demanded during class, but not after class.

One teacher from Zhixin South Road Elementary School claimed that to master a language, a little time in class is not enough and it must be used in daily communication. The teacher also labelled parents' reactions as unreasonable and said they must understand the school language policy is for the good of the children.

When journalists Hu Jun 胡军 and Zi Yongting 资勇庭 from Yangcheng Evening News attempted to interview the school officials to discuss this matter, they were stopped by security guard at the school gate. They were told that there were no written regulations and the school officials would not give them an interview unless they had an approval from local education authority.

The reporter later interviewed Yuexiu Bureau of Education 越秀区教育局 (*Yuèxiù qū jiàoyù jú*) and he received a simple response regarding this matter, which said that “Bureau of Education does not expressly ban students from speaking Cantonese, but recommends to strongly promote Mandarin”.

This events actually represents the state's contradictory attitude towards language policy. At first, Cantonese Day was organized with a lot of enthusiasm for local Cantonese culture and it was presented in the media as an event with a very positive impact. When the event attracted attention of local education authorities and the school was subjected to pressure from them, their attitude changed. The headmistress never acknowledged the fact she was under a pressure from the authorities, but rather downplayed the whole event as unimportant, and even accused the newspaper's article as dishonest, exaggerated and attention-seeking. She did not dare to admit the school's attitude towards this activity changed significantly and said the event's goal was not to promote Cantonese, because only Standard Mandarin is allowed at school ground.

The policies published by authorities are transparent, as we can see from the Yuexue Bureau of Education's response. This type of response is very

ambiguous, it does not express any attitude towards the school's policy and does not give any instructions how to handle language policy in classroom, In Mainland China, Standard Written Chinese is the only legitimate written language to be used in public domains, the only written language taught and used in school (Liang, 2015). As this only concerns written form, it does not refer to usage of spoken language in classroom or after class. It is apparent that this ambiguity and lack of transparency in language policy leads to uncertainty among everyone involved.

The education system also has a slightly paternalistic approach towards the application of language policy, which manifests itself when the Zhixin South Road Elementary School's teacher says that the school language policy is for the good of the children; if it's for the common good, something must be sacrificed, whether people appreciate it or not. When people express disagreement with this type of policy, it's marginalized as unimportant and against the common good of everyone. This strict policy goes against the wishes of most families in Guangdong. According to a survey conducted Li, Li & Gao (2019), 91,5% of parents approve of children speaking Mandarin at young age, but 81,5% also support the idea of speaking Cantonese with them.

Some inhabitants of Guangzhou recounted that their son was able to speak both Cantonese and Mandarin proficiently, because the teachers in kindergarten spoke both Chinese varieties, but once he and all his classmates enrolled in primary school, they all switched to Mandarin. (Liang, 2015) It's important for children to have an opportunity to speak both Chinese varieties at school if they want to keep sufficient proficiency. Even though the only official language is Standard Modern Chinese, children still have the opportunity to speak Cantonese with their classmates. When language policy is too restrictive, such as in the case of Zhixin South Road Elementary School, children will not be able to speak both languages for everyday communication, as is shown in the story of Mr. Ceng and his family. This is a common issue in Guangdong; children grow up in a Cantonese speaking family, but when they enroll a school where speaking Cantonese is unwelcome, they level of Cantonese knowledge declines. According to Li, Li & Gao (2019), 85,4%

respondents agree that it is crucial for young children to have a family environment which supports them speaking Cantonese.

This tendency is not unique for Guangdong, children in other provinces are often discouraged or even prohibited from speaking other Chinese varieties than Mandarin; this is a consequence of *Promotion of Putonghua* policy (Barov & Egorova, 2019). According to Xiong (2018), it is a common situation, that middle aged are able to speak Mandarin and vernacular, their children speak Mandarin and little to no vernacular, and children's grandparents speak vernacular and little to no Mandarin.”

From previously mentioned stories, we can see that family environment is not enough if children do not have the opportunity to speak their local languages in different environments. Children spend a substantial amount of time at school, so being able to use their local varieties between classes and communicate with each other is essential to keep sufficient proficiency. According to Liang (2015), Cantonese is not the primary language of Guangzhou anymore, due to Mandarin interference.

### **Multilingualism in China**

Why do schools apply such strict measures? One of the reasons is the idea of unifying China as a country through one common national language; an idea that has prevailed since the 19<sup>th</sup> and 20<sup>th</sup> century reforms in China. Even though Zhou Enlai stated in 1956, that promotion of Mandarin should not lead to prohibition of other varieties (Guo, 2004)), in reality, speaking other varieties is often prohibited even between classes.

Other reasons for these strict policies are common beliefs and biases related to monolingualism and monolingualism. Specifically, *standard language ideology* and *monolingual ideology*.

*Standard language ideology* believes there is one standardized, homogenous and model form of a language, which should be used in public space. This concept is an important part of *linguistic prescriptivism* and *linguistic purism*. SLI is biased against varieties which are not standardized or do not have a



standardized script, because the core belief of this concept is that *standardized written language* drawn primarily from the spoken language of the upper middle class should be the only acceptable language. (Pérez-Quiñones & Salas, 2021; Green, 1994)

*Ideology of multilingualism* is the idea that communication happens by only one language at a time (Pérez-Quiñones & Salas, 2021). Multilinguals do not necessarily switch a language, they can use all of the languages available fluidly: mixing words, phrases, idioms (Pérez-Quiñones & Salas, 2021). This is often ignored by education system in PRC, as teachers expect children to have the same ability in Mandarin and place the responsibility on children's parents (Liang, 2015); this approach ignores bilingual individuals as a group (Pérez-Quiñones & Salas, 2021). According to ideology of multilingualism, one single dominant language, a national language, should be used in a public life of a multilingual society. (Zhou, 2019; 36) This idea has prevailed in many education systems worldwide for a long time, as prohibiting children speaking other than the official standardized variety is not an issue that is unique to Guangdong or China.

*Standard language ideology* and *ideology of multilingualism* both contribute to negative stereotypes towards multilingual speakers. Due to the mistaken notion that being monolingual and speaking one standardized language is the desirable norm, speakers of other languages or varieties are often seen as less intelligent, especially when they do not speak the standardized language or speak with a heavy accent (Pérez-Quiñones & Salas, 2021). Bilinguals and multilinguals also do not develop the same language registers in all languages. This applies also in Guangdong's diglossia, where Cantonese is mainly spoken in informal setting, while Mandarin is used in formal setting. Using different registers often leads to a false conclusion, that bilingualism and multilingualism have a negative impact on learning languages properly. Speaking Cantonese still maintains prestige among locals, it is often deemed as "uncivilized" and "unstandardized" by the education system. There is a gap between locals' and the government's attitude towards multilingualism in Guangzhou; according to Li, Li & Gao (2019), 95,2% of research participants think children should speak their local variety, 69,2% of them disagree with

the statement that speaking Cantonese will affect children's Mandarin negatively.

Mandarin already became widespread in Guangzhou, and with stricter *tuipu* policy, “the language situation is changing and the functional principle of diglossia is violated” (Landry & Allard, 1994). Guangdong cannot expect bigger changes or policies inclined towards multilingualism; the education system is still in favour of standard language ideology and ideology of monolingualism. When Cantonese speakers from Guangdong want to gain better Cantonese literacy, they do so in WeChat groups or on other social media. Chinese varieties that are defined as *fangyan*, such as Yue Group, cannot be studied in the education system and this is presumably not going to change any time soon.

### **Pro-Cantonese Protests 2010**

On 5<sup>th</sup> of July 2010, Chinese People's Political Consultative Conference 中国人民政治协商会议 (*Zhōngguó rénmin zhèngzhì xiéshāng huìyì*), a political advisory body and a key part of Chinese communist party, wrote a proposal to the then mayor of Guangzhou, Wan Qingliang 万庆良 (Wang, 2019). The proposal suggested increase of Mandarin broadcasting on Guangzhou Broadcasting Network 广州广播电视台 (*Guǎngzhōu Guǎngbò Diànshìtái*). According to the proposal, submitted by Ji Keguang 纪可光, a deputy director of People's Political Consultative Conference in Guangzhou, Mandarin should be made primary broadcasting language all prime-time programs should be converted into Mandarin.

Another controversial comment was made by Wang Yang 汪洋, the then Party Secretary of Guangdong province, who is not a native of Guangdong. He stated “The Cantonese, having gotten prosperous, ought to become educated” (Ho & Lu, 2019). This comment implied that learning Mandarin would raise their cultural level, which represents the typical government narrative towards other varieties of Chinese.

Broadcasting in Cantonese is very important for local community. Modern media, including television, strongly influence Cantonese literacy in Guangzhou (Yan, 2008). Even though 80% of people were against the proposal to increase Mandarin broadcasting, the author of the proposal did not take it into consideration and said “we must guide and teach the people who oppose the proposal” (Ho & Lu, 2019).

In July and August 2010, “Pro-Cantonese Protests” 广州撑粤语行动 (*Guǎngzhōu chēng yuèyǔ xíngdòng*) took place in Guangzhou. These events started on July 11, when a gathering called “*For Cantonese, I Sing from the Top of My Voice*” 我为粤语大声唱 (*Wǒ wèi yuèyǔ dà shēng chàng*) was organized. Another event called “*Big Movement to support Cantonese*” 齐撑粤语大行动 (*Qí chēng yuèyǔ dà xíngdòng*) was organized on July 25; even though the authorities forced police to cancel the event, 10 000 out of 20 000 people gathered to protest. On August 1, an event called “*Big Move to Protect Cantonese: A Synchronized Joint Action*” 保护粤语大行动同步联动 (*Bǎohù yuèyǔ dà xíngdòng tóngbù liándòng*) was organized. Most of the participants were young, born in the 1990s (Ho & Lu, 2019), which is typical for Guangdong’s pro-Cantonese movements and activities. The protests received a lot of attention online, the especially on Weibo: the majority of Weibo entries were typed in Cantonese (Wang, 2019). During these events, Cantonese acted as a unifying factor among protesters in Guangzhou. Many people argued for uniqueness of Cantonese: “Cantonese is a much more sophisticated language with 9 pitched tones while Mandarin only has 4” or linguistic and identity rights “I can endure the government misusing my tax, demolishing my house and social inequalities, but how dare they to take away my right to speak my mother tongue! I can never endure that because Cantonese is who I am!” (Wang, 2019).

These series of protests were seen as a manifestation against restrictive tuipu policy. In Guangdong’s diglossia, the language situation is seen differently by authorities and by locals. The typical position promoted by the government says that the common standardized language is a high variety, while the local Chinese variety is the low variety. In some regions, local people have

complied with this language policy. In Guangdong, the situation is reversed: locals see Cantonese as the high, more prestige variety.

## 7 Written Cantonese

While *Standard Written Chinese* is considered universal for all Chinese varieties, some of them, including Cantonese, maintain their own writing tradition. Cantonese can be represented in:

- 1) Standard Written Chinese ( 中文书面语 *Zhōngwén shūmiànyǔ*)
- 2) Vernacular Written Cantonese ( 粵文 *kyut6man4*)
- 3) Mixed codes

Standard Written Chinese is used in formal communication and official documents. Even in Hong Kong, where most population speak Cantonese, the officially used standardized language is Standard Written Chinese. It is a norm to read SWC with Cantonese pronunciation, especially at school. (Cheng & Tang, 2014) Although SWC can be read in Cantonese, it does not correspond to distinctive Cantonese features, such as its syntax or specific lexemes.

Standard Written Chinese Characters are divided into six categories 六书 (*liù shū*):

- 1) Pictograms 象形字 (*xiàngxíngzì*)
- 2) Simple ideograms 指事字 (*zhǐshì*)
- 3) Compound ideographs 会意字 (*huìyìzì*)
- 4) Phono-semantic compound characters 形声字
- 5) Transfer characters 转注字 (*zhuǎnzhùzì*)

6) Loan characters 假借字 (jiǎjièzì)

Pictograms 象形字 are the oldest type of characters and depict the object they represent, for instance:

- a) mountain 山
- b) wood 木
- c) mouth 口

Simple ideographs 指事字 express the meaning through to an iconic form, for instance:

- a) up 上
- b) down 下
- c) one, two, three 一, 二, 三

Compound ideographs 会意字 contain two or more ideographs, for instance:

- a) forest 森, compound of three characters for wood 木
- b) burn 焚, compound of characters for fire 火 and forest 林
- c) face 脸, compound of characters for meat 肉 and moon 月

Phono-semantic compound characters 形声字 contain one phonetic component, which determines the pronunciation, and one semantic component, for instance:

- a) silk 绸 consists of semantic component for silk 丝 and phonetic component 周
- b) to stare 瞪 consists of semantic component for eye 目 and phonetic component 登
- c) mountain 岭 contains semantic component for mountain 山 and phonetic component 令

Transfer characters 转注字, also referred to as derivative cognates, are characters that used to be semantically interchangeable. They have the same etymological roots. It is the smallest category of Chinese characters.

- a) space 空 and hole 窍
- b) old 老 and exam 考

Loan characters 假借字, homophonous morphemes, that were borrowed to create another morpheme with the same pronunciation:

- a) to call 呼 was used to create breathe 呼
- b) north 北 was used to create back 背
- c) few 少 was used to create sand 沙

## 7.1 Vernacular Written Cantonese

Cantonese has maintained and developed its own writing tradition: Vernacular Written Cantonese, which is estimated to be the most complete form of written Chinese after Mandarin. Vernacular Written Cantonese respects specifics and peculiarities of spoken language that cannot be expressed in SWC.

Written Cantonese is difficult to find in public space due to restrictive government policy. In private lives, written Cantonese is much more common, especially in private communication, online chat groups and social network.

## 7.2 History of Vernacular Written Cantonese

Written Vernacular Cantonese has its own history; it had already developed by the 1930 (Snow, 2004:98). The first written Cantonese texts are *mukjyu* 木魚, literally “*wooden fish*”. *Mukjyu* 木魚 are rhymed texts with themes drawn from history, myths, legends, folk tales and religion. (Zheng, 1992) The first dated *mukjyu* songbook, “*The Flowery Paper*” 花箋記 (*faalzinlgei3*), was composed during Ming dynasty and dates from 1713. (Zheng, 1992) One of the earliest publications in colloquial Cantonese is also *Jyutau* 粵謳, written by Zhao Ziyong 招子庸, first published in 1828. (Chen & Joch, 2015) Zhao Ziyong used a mixture of standard Chinese characters and additional characters created specifically for Cantonese to capture spoken Cantonese of that time. (Chen & Joch, 2015). This corresponds to the situation nowadays, when Cantonese is often represented by mixed codes. Written Vernacular Cantonese can be also found in scripts for Cantonese opera 粵劇 (*jyut6kek6*).

## 7.3 Cantonese-specific characters

Written Vernacular Cantonese is almost unintelligible to non-Cantonese readers. (Cheung & Bauer, 2002) A text written purely in Cantonese without interference from Mandarin, only 30-40% of it is comprehensible by non-

Cantonese speakers. (Mair, 2003) Standard Written Chinese characters and also considerable number of Cantonese-specific characters (Yan, 2012), as many Cantonese-specific words do not have any equivalent in Mandarin, for instance:

- 1) 唔 *m4 not*
- 2) 佢 *keoi5 he, she, it*
- 3) 哋 *dei2 marker for plural forms (佢哋 keoi5dei2 they)*
- 4) 咁 *gam2 so*
- 5) 俾 *tau2 to take a rest*
- 6) 咩 *me1 final particle for questions which indicate surprise*
- 7) 睇 *tai2 to see*
- 8) 㗎 *zo2 particle of perfective aspect*

Many of these Cantonese characters contain the mouth radical 口 (jyutping *hau2*), which is typical for Cantonese-specific characters. In VWC, the phonetic norm comes first to represent the phonetic value as closed as possible; the semantic norm is secondary. (Yan, 2008) This principle was already found in *mukjyu* 木鱼; according to Snow (1991; 23), there was a tendency for Cantonese characters to be chosen on the basis of the sound. This trend has prevailed until now; as Vernacular Written Cantonese is not used in the domain of education, literacy in VWC can be gained only in informal domains. Cantonese-specific characters are not taught at school, so readers have to rely mainly on pronunciation, which leads to stronger tendency to phonetic borrowing than in Mandarin. Phonetic borrowing often leads to ambiguity and a discrepancy between the semantic component and the target meaning. Li (2000) gives this example: 一舊鷄 *jat1 gau6 gai1*, “a lump of chicken” in Cantonese. The character 舊 *gau6* means “old” in Standard Written Chinese, but it is borrowed based on its pronunciation to express the homophonous



Cantonese classifier for “a lump of”. Readers who are not familiar with Cantonese might misread 一舊鷄 as “one old chicken”.

Due to lack of standardization, many Cantonese speakers use mix codes. However, written Hong Kong Cantonese is highly conventionalized and accepted, so it is slowly becoming standardized (Snow, 2008). Consensus about character usage differ depending on the character. Some characters are conventionalized, for instance:

- 1) 冇 *mou5* not
- 2) 佢 *keoi5* he, she, it

Some characters are less conventionalized

- 1) 啲 or 的 *di1* for the general comparative marker
- 2) 俾 or 畀 *bei2* for the verb “to give”

### **7.3 Input methods and romanization systems**

Using and typing Cantonese-specific characters encounters another issue: suitable input method. Due to no official standardization, there is a lack of input method for Cantonese characters. Most input methods are fan-made and non-official. There are Cantonese input methods which use jyutping or other romanization systems, such as Canton Easy Input, CPIME 2009, LHSK Jyutping, Online Jyutping Input Method, Red Dragonly or RIME. There’s also Hanyu-Pinyin based software which enables users to input other Chinese varieties, or Siri, the speech recognition software, which is able to understand, transcribe and speak Cantonese.

Using Cantonese input methods also requires knowledge of Cantonese Romanization systems and phonetic symbols. These schemes include:

- 1) Cantonese Bopomofo 粤语注音符号 (*Jyut6jyu5 zyu3jam1 fu4hou6*), a system of phonetic characters used to transcribe Cantonese pronunciation

粤语 → ㄐㄨㄥ

- 2) Cantonese Pinyin 教院式拼音方案 (*Gaa3jyun3 ping3jam1 fong1on3*) uses latin alphabet and numbers to represent the pronunciation and 9 tones

粤语 → *jyt9 jy5*

- 3) Guangdong Romanization 广东拼音方案 (*Gwong2dung1 ping3jam1 fong1on3*), published by an education department in Guangdong, along with romanizations for Hakka, Hainanese and Teochew

粤语 → *yud6 yu5*

- 4) Jyutping 粤拼 (*Jyut6ping3*), developed by LHSK in 1993 and used mainly for teaching Cantonese

粤语 → *jyut6jyu5*

- 5) Sidney Lau Romanisation 刘錫祥拼音 (*Lau4 Sek3 Coeng4 ping3jam1*), scheme created by Sidney Lau 刘錫祥 to teach Cantonese to Westerners living in Hong Kong

粤语 → *yuet5yue6*

- 6) Standard Romanization of Cantonese 粤语标准罗马字 (*jyut6ju5 biu1zeon2 lo4 maa5 zi6*), developed at the end of 19<sup>th</sup> century by Christian missionaries in South China

粤语 → *yûet-yǔe*

- 7) Yale Romanization of Cantonese 耶鲁拼音 (*Je4 lou5 ping3jam1*), created at Yale University and widely used in Cantonese textbooks for English speaker

粤语 → *Yuhtyúh*

Most Cantonese speakers in Guangdong are not familiar with any of these romanization systems, as they only learn pinyin at school. As there is no widespread system for typing, every existing input method uses different romanization; input methods are often created by users themselves, so they do not use one standardized system and contain many errors. Input methods of Cantonese are not always accessible in PRC. Many of input methods are not supported by computer operated system or there is no possibility to download them, unless users install VPN.

#### 7.4 Typing strategies and mixing codes

Yan (2008) introduced four strategies of typing Vernacular Written Cantonese:

- 1) Writing in SWC characters
- 2) Writing in VWC characters
- 3) Writing in false characters (pseudo-characters)
- 4) Writing in Romanized letters

Many users combine two or more typing strategies; it depends on input methods that are accessible for them. Cantonese speakers in Guangdong often use pinyin, which influences the choice of Cantonese characters. Many characters that are frequent in Cantonese are very rare in Mandarin or they do not exist in Mandarin at all. Some users avoid this situation by mixing Mandarin and Cantonese. When they want to type Cantonese-specific characters, they often have to choose alternative variants, for instance:

- 1) 抖 *to tremble* (pinyin *dǒu*, jyutping *dau2*) instead of 啱 *to rest* (jyutping *tau2*) when typing the expression 早啱 (jyutping *zou2 tau2*), which means “*good night*”
- 2) 野 *wild* (pinyin *yě*, jyutping *je5*) instead of 嘢 *thing* (jyutping *je5*)
- 3) 晒 *shine, dry in the sun* (pinyin *shài*, jyutping *saai3*) instead of 晒 *verbal aspect marker for full extent* (jyutping *saai3*) in 唔该晒 (jyutping *m4 goi1 saai3*), which means “*thank you very much*”

- 4) D or 的 *possessive particle* (pinyin *de*, jyutping *dik1*) instead of 啲 *the general comparative marker* (jyutping *di1*)
- 5) 呀 *particle expressing surprise* (pinyin *yā*, jyutping *aa1*) instead of 㗎 *final particle* (jyutping *aa1*)
- 6) 来 to arrive (pinyin *lái*, jyutping *loi4*) or 黎 (pinyin *lí*, jyutping *lai4*) instead of 嚟 (jyutping *lai2*)

Due to lack of standardization and input methods many expressions above exist in two or more orthographic forms. Although there are less Cantonese speakers among younger generation compared to the generation of their parents and grandparents, young internet users use much more Written Cantonese online and they are also form chat groups which advocate for speaking and writing Cantonese correctly. (Liang, 2015) Internet and social media offer much more possibilities for a further development of Written Cantonese; More than 90% of Guangdong locals type vernacular Cantonese characters in text messages. (Yan, 2008) Most speakers from Guangdong who are literate in Vernacular Written Cantonese (VWC) are also literate in SWC, which makes them logographic-logographic bilinguals (Ma et al., 2020), or more precisely, logographic-logographic biliterates. Not all Cantonese speakers from Guangdong are literate in VWC, but all of them are literate in SWC. For learning VWC, being literate in SWC is necessary; due to lack of learning materials, VWC literacy is obtained mainly from other Cantonese speakers through mutual communication and sharing information about VWC. Cantonese-Mandarin bilinguals who are literate in SWC but not in VWC type Cantonese by Simplified standardized characters.

## 7.5 Written Cantonese in media

In Hong Kong, written Cantonese is frequently used in media, while in Guangdong usage of Cantonese is limited due to restrictions regarding language policies. Vernacular Written Cantonese is not visible in public space in Guangdong due to strict *tuipu* policy, which restricts usage of other written variants than SWC. One famous slogan, a part of *tuipu* campaign, says “Speak Mandarin, use standardized characters, be a civilized person” 说普通话, 用规范字, 做文明人 (*Shuō pǔtōnghuà, yòng guīfàn zì, zuò wénmíng rén*). A part of Cantonese-speaking community is also influenced by this attitude towards written vernaculars, and deems VWC uncivilized. (Yan, 2008)

Despite that, VWC can be seen in Guangzhou media. The article “*Many Guangzhou children cannot speak Cantonese*” 好多广州细路唔识讲白话 is an example of VWC usage in Guangdong media. Chinese Newspaper *Southern Metropolis Daily* (南方都市报 *Nánfāng dūshì bào*) dedicates one whole page to articles written in Cantonese. (Liang, 2015) *Guangzhou Daily* regularly includes elements of vernacular Cantonese in their titles and articles (Yan, 2008):

- 1) 餐厅不明码标价顾客埋单算三次。

Cāntīng bù míngmǎ biāojià gùkè **máidān** suàn sāncì

“Not clearly marking the price, the restaurant charged the customer three times.”

- a) 埋单 (pinyin: *máidān*; jyutping: *maai4daan1*) – pay the bill

- 2) 大热天尝完啫啫饮靚汤。

Dàrè tiān chángwán zhězhě yǐn liàng tāng

“In the hot day, let’s first try zeze, then drink good soup.”

- a) 啫啫 (pinyin: *zhězhě*; jyutping: *ze1ze1*) – type of Cantonese food

b) 饮靚汤 (pinyin: *yǐn liàng tāng*; jyutping: *jam2leng3tong1*) – drink good soup

3) 广州市区“限摩”知多 D。

Guǎngzhōu shì qū xiàn mó zhīduō D

“Know more about the motorcycle restriction in Guangzhou city.”

a) 知多 D (pinyin: *zhī duō D*, jyutping: *zi1do1 D*) – know more

Unless readers have some literacy in Cantonese, they cannot understand these titles completely. Mandarin speakers with no literacy in Cantonese will be only able to understand part of it. In some cases, Mandarin speakers are only able to understand separate characters, but not meaning of the whole sentence

## 8 Bilingual Lexical Decision Task

### 8.1 Recruiting respondents

Research participants were recruited online, on social networks WeChat and Facebook, additionally on Research Gate. Most participants were recruited with the help of my friends and teachers on Guangdong, who created chat groups on WeChat, where all participants were invited. Respondents who took part in the survey were asked to share the link with their friends. I also shared the link for the first part in WeChat groups, Facebook and other social media. The research constituted of two parts: 1) questionnaire 2) lexical decision task. The first part was conducted on a free website 腾讯问卷 (*téngxùn wènjuàn*) <https://wj.qq.com>.

Suitable participants were required to be natives of Guangdong province and speak Cantonese. The most important part of the research was aimed at

participants' usage of language. Research respondents were asked to fill in information about year of birth, province and city, education and of study. The respondents could use a pseudonym nickname for the study.

## 8.2 Questionnaire structure

Respondents were asked these questions:

- 1) 列出所有你会说的语言（例如：中文/英语/法语等等）

List all languages you can speak (for instance: Chinese, English, French, etc.)

- 2) 列出所有你会说的方言（例如：广州话/客家话/潮汕话/闽南语/吴语等等）

List all varieties of Chinese you can speak (for instance: Cantonese, Hakka, Teochew, Hokkien, Wu)

- 3) 请你把方言按照熟练水平列出

List Chinese varieties you speak according to your proficiency

- 4) 当你与所用方言同样流利的人交谈时，你会首选哪一种方言？请你列出

When speaking to someone who is equally fluent in all Chinese varieties you can speak, which one would you choose to speak with them according to your preference?

- 5) 你最佳首选的方言是哪一种

What is your most preferred variety of Chinese?

The following questions were related to the usage of Chinese variety they chose in question 5). They had to answer information about what age they started to acquire and what age they started to speak fluently. Further, they had to answer questions about environment where they use the particular language: *family, friends, university, professional environment*. They had to mark frequency how often they use it in each environment: *daily, three times a week, once a week, occasionally, not applicable*.

Another set of questions was related to exposure to the particular Chinese variety during free time activities: reading books, watching films, scrolling through social media, listening to radio/music).

The second part of the research was the lexical decision task. The requirements for the participants to attend the second part of the research were following:

- 1) to fill “Cantonese” (粤语/广东话/广州话) as an answer in 4) and 5)
- 2) use it daily
- 3) being exposed to it daily

### **8.3 Selected respondents**

In total, 110 participants filled in the questionnaire. Out of 110 people, 70 were invited to attend the second round, which contained lexical decision experiment.

All 30 participants who attended the second round were born between 1979 and 2001 (mean 1997). All of them were natives of Guangdong province: 15 from Guangzhou 广州, 4 from Dongguan 东莞, 2 from Foshan 佛山, 2 from Meizhou 梅州, 2 from Shenzhen 深圳, and the rest 4 from Huizhou 惠州, Qingyuan 庆元, Yunfu 云浮 and Zhanjiang 湛江.

Except for one respondent, the rest 29 all either enrolled in bachelor degree program or graduated with bachelor degree already.



Beside Cantonese, respondents were also speakers of other Chinese varieties: Hakka 客家话 (5 respondents), Teochew 潮州话 (4 respondents), Kaipinghua 开平话 (2 respondents) and Hokkien 闽南语 (1 respondent).

#### **8.4 Lexical decision task in Chinese**

Chinese script 汉字 (*hànzì*) is a logographic script. It can be classified as *deep orthography* (Cheng & Tang, 2014). While *shallow orthography* has one-to-one relationship between words and sounds, in *deep orthography*, the relationship between words and sounds are more complex.

One basic orthographic unit is one Chinese character called 字. Each character represents a morpheme 语素 *yǔsù*, one single unit of meaning. One Chinese morpheme can also stand for a single lexeme. In contemporary Chinese languages, most lexemes are compounds of two or more characters, but characters with more than three or four characters are less frequent. Two character-compounds represent 73,6% of Chinese lexeme (Lim, 2020).

When creating the lexical decision task, it is especially important to pay attention to the parts where Chinese languages with logographic script differ from European languages, which mostly use alphabet. In a *deep orthography* like Chinese script, the information density can be packed within one character. Over 70% of Chinese characters are phonograms. (Tsang, 2017) Chinese is considered a morphosyllabic language, as each character corresponds to both syllable and morpheme (Tsang, 2017). This also plays role in recognizing characters. While word recognition in alphabetic languages is the key to rely heavily on phonological encoding, this is not completely applicable to Chinese. Lexical decision task in Chinese has many specifics; readers rely predominantly on orthographic and semantic information when processing visually presented Chinese words (Lim, 2020). Measures based on self-reported language proficiency and academic ability are poor predictors of word recognition performance; there is a weak correlation between fluency and self-rated language proficiency (Lim, 2020), so I did not include any

questions about readers proficiency, which would be also difficult to estimate for the participants.

This experiment was based on lexical recognition exclusively, without audio recordings. Only Simplified characters were used for the study. Even though Cantonese traditionally uses Traditional characters, people from Guangdong are exposed to Simplified characters for most of their life, as they learn Simplified characters at school, and they also use it to type Cantonese. They are not exposed to Traditional characters for the same amount of time during they lives, so using Traditional characters could presumably increase the response time and number of incorrect responses. Most of the characters are standardized, although Cantonese-specific characters were included in some expressions:

- 1) 唔 *m4* (not)
- 2) 𨀐 *nau1* (angry)
- 3) 瞓 *fan3* (sleep)
- 4) 𨀐 *gui6* (tired)
- 5) 𦉳 *ngap1* (to babble)
- 6) 𨀐 *goe1* (stupid)

Sometimes characters can be associated with more than one meaning (花, 曾, 了) Because the boundary between a logograph (character) and a word is often unclear, I wanted to avoid this situation by omitting single-character words (惊, 累, 𨀐). Most characters carry their own meaning. The only exception are characters that do not carry any meaning as separate morphemes 语素, for instance:

- 1) in 垃圾 *lājī* (trash), separate characters 垃 *lā* and 圾 *jī* do not mean anything
- 2) 蜘 *zhī* from 蜘蛛 *zhīzhū* (spider) does not carry any meaning

- 3) 蝴 *hú* in 蝴蝶 *húdié* (butterfly) does not have any meaning as a separate character

In most cases, characters act as morphemes 语素 and they carry certain meaning, so participants would recognize them as real words. Characters do not act as an orthographic code for accessing the representation of words. The definition of a “word” in Chinese is very ambiguous: there is a term 词 *cí*, which can refer to a word, term or a speech. There are also terms 词语 *cíyǔ* and 语词 *yǔcí*. The term 词语 can refer to words, terms and expressions; the term 语词 can refer to words, expressions and phrases. Chinese languages do not have a clear definition of a *word* due to the nature of Chinese script. In LDT, it is a common usage to use terms *words* and *nonwords* in English, so I adhere to this practice in my experiment. Instructions for the experiment were written in Chinese; terms 单词 *dāncí* (individual word) and 短语 *duǎnyǔ* (phrase, expressions) were used instead of *words*.

In total 80 compounds were used for the study: 20 Cantonese words, 20 Mandarin words, 20 Cantonese non-words and 20 Mandarin non-words. All compounds were checked and approved by native-speaker from Guangzhou. In bilingual studies, Mandarin and Cantonese are usually considered to be two distinct languages; Cantonese-Mandarin bilinguals have separate lemma representations for Mandarin and Cantonese translation equivalents. (Cai et al., 2011)

Lexical decision task was designed in Psychopy, a software used for experiments in psycholinguistics and neuroscience. The experiment was uploaded to pavlovia.org, a website for online experiments in behavioural sciences. The compounds used in the experiment were divided into 5 categories: Mandarin words (MW), Mandarin nonwords (MN), Cantonese words (CW) and Cantonese nonwords (CN). In total, each participant had to: 20 MW, 20 MN, 20 CW and 20 CN. No lexical priming was chosen for the experiment.

Compounds chosen for the lexical decision task are related to emotions, feelings and behaviour. Most of them are colloquial expressions that are not a part of formal vocabulary. Cantonese has a rich informal register, especially emo-

tion lexicon, but lacks formal register; formal registers of Cantonese and Mandarin largely overlap.

Previous researches proved that Mandarin and Cantonese have separate lemma representations, such as distinct languages English and Dutch, and from a psycholinguistic view, they are rather separate languages than dialects. (Cai et al, 2011)

Words in MW category are translation-equivalents of words in CW category. Cantonese is considered as a vivid and lively language by its speakers due to the absence of formal lexicon and complexity of expressions related to emotions (Cheng & Tang, 2014), so many words and expressions do not have their exact translation equivalent in Mandarin. The translation equivalents are in the same order in both tables.

All words and nonwords were checked, corrected and approved by a bilingual native speaker from Guangzhou, who is biliterate in both Cantonese and Mandarin.

Nonwords were created by several strategies:

- 1) switching the order of characters: 喉擒 => 擒喉
- 2) replacing one character: 唔开心 => 唔哈心
- 3) replacing two characters: 唔抵得 => 唔底特
- 4) replacing with visually similar characters: 曳曳 => 拽拽
- 5) creating a nonword with a same or similar pronunciation: 丢架 => 叻  
啲

## 8.5 Cantonese words & nonwords

| Number | Words | Jyutping              | Translation                 | Nonwords | Jyutping without tones |
|--------|-------|-----------------------|-----------------------------|----------|------------------------|
| 1      | 唔开心   | <i>m4hoi1sam1</i>     | happy                       | 唔哈心      | <i>mhoisam</i>         |
| 2      | 老定    | <i>lou5ding6</i>      | calm                        | 姥啖       | <i>louding</i>         |
| 3      | 𩙱爆爆   | <i>nau1baau3baau3</i> | angry                       | 纽爆爆      | <i>naubaaubaa</i>      |
| 4      | 激𩙱    | <i>gik1nau1</i>       | very angry/to make angry    | 搵𩙱       | <i>giknau</i>          |
| 5      | 喉擒    | <i>hau4kam4</i>       | can't wait for/desperate    | 擒喉       | <i>kamhau</i>          |
| 6      | 惊青    | <i>ging1ceng1</i>     | frightened                  | 憬青       | <i>gingceng</i>        |
| 7      | 慌失失   | <i>fong1sat1sat1</i>  | frantic, flustered          | 慌失失      | <i>fongsatsat</i>      |
| 8      | 搵搵震   | <i>tan4tan2zan3</i>   | shivering, very dizzy       | 搵搵振      | <i>tantanzan</i>       |
| 9      | 得戚    | <i>dak1cik1</i>       | smug, arrogant              | 戚得       | <i>cikdai</i>          |
| 10     | 曳曳    | <i>jai1jai1</i>       | mischievous, naughty        | 拽拽       | <i>jaijai</i>          |
| 11     | 闭翳    | <i>bai3ai3</i>        | worried, depressed          | 翳闭       | <i>aibai</i>           |
| 12     | 眼冤    | <i>ngaan5jyun1</i>    | offensive to the eye        | 眼宛       | <i>ngaanjyun</i>       |
| 13     | 眼瞓    | <i>ngaan5fan3</i>     | sleepy                      | 瞓眼       | <i>fanngaan</i>        |
| 14     | 好边    | <i>hou2gui6</i>       | very tired                  | 边好       | <i>guihou</i>          |
| 15     | 心喻    | <i>sam1ngap1</i>      | distressed, grieved         | 心吸       | <i>samngap</i>         |
| 16     | 唔喺    | <i>m4goe4</i>         | disgruntled                 | 唔锯       | <i>mgoe</i>            |
| 17     | 顶唔顺   | <i>ding2m4seon6</i>   | cannot stand something      | 酇唔顺      | <i>dingmseon</i>       |
| 18     | 唔抵得   | <i>m4dai2dak1</i>     | cannot restrain one's anger | 唔底特      | <i>mdaidak</i>         |
| 19     | 丢架    | <i>diu1gaa2</i>       | lose face                   | 叨架       | <i>diugaa</i>          |
| 20     | 冇面    | <i>mou5min2</i>       | be embarrassed              | 面冇       | <i>minmou</i>          |

### 8.5 Mandarin word & nonwords

| Number | Words | Pinyin           | Translation                 | Nonwords | Pinyin without tones |
|--------|-------|------------------|-----------------------------|----------|----------------------|
| 1      | 不开心   | <i>bùkāixīn</i>  | happy                       | 不開心      | bukaixin             |
| 2      | 镇定    | <i>zhèndìng</i>  | calm                        | 震定       | zhending             |
| 3      | 愤怒    | <i>fènnù</i>     | angry                       | 努愤       | nufen                |
| 4      | 激怒    | <i>jīnù</i>      | very angry/to make angry    | 怒激       | nuji                 |
| 5      | 急切    | <i>jíqiè</i>     | can't wait for/desperate    | 急窃       | jiqie                |
| 6      | 惊慌    | <i>jīnghuāng</i> | frightened                  | 憬慌       | jinghuang            |
| 7      | 心慌    | <i>xīnhuāng</i>  | frantic, flustered          | 心荒       | xinhuang             |
| 8      | 好晕    | <i>hǎo yūn</i>   | shivering, very dizzy       | 晕好       | yunhao               |
| 9      | 得瑟    | <i>dése</i>      | smug, arrogant              | 瑟得       | sede                 |
| 10     | 淘气    | <i>táoqì</i>     | mischievous, naughty        | 掏汽       | taoqi                |
| 11     | 闷气    | <i>mēnqì</i>     | worried, depressed          | 懣汽       | menqi                |
| 12     | 碍眼    | <i>àiyǎn</i>     | offensive to the eye        | 眼碍       | yanai                |
| 13     | 困了    | <i>kùnle</i>     | sleepy                      | 了困       | lekun                |
| 14     | 好累    | <i>hǎolèi</i>    | very tired                  | 累好       | leihao               |
| 15     | 心痛    | <i>xīntòng</i>   | distressed, grieved         | 心痛       | xintong              |
| 16     | 不爽    | <i>bùshuǎng</i>  | disgruntled                 | 补爽       | bushuang             |
| 17     | 忍不住   | <i>rěnbuzhù</i>  | cannot stand something      | 忍不猪      | renbuzhu             |
| 18     | 气不过   | <i>qìbùguò</i>   | cannot restrain one's anger | 汽不挝      | qibuwo               |
| 19     | 失面子   | <i>shīmiànzi</i> | lose face                   | 面失子      | mianshizi            |
| 20     | 没面子   | <i>méimiànzi</i> | be embarrassed              | 没子面      | meizimian            |

## 8.6 LTD Experiment Process

The experiment was conducted in Psychopy and uploaded online to pavlovia.org.. Respondents, whose answers were suitable were invited to attend the second round. Before sharing the link, I informed them about the whole experiment procedure: they were asked to use computer with an external keyboard and possibly VPN.

After opening the link for the experiment, participants saw a screen with instruction for the trial test. The trial test contained 10 items, 5 words and 5 nonwords in a randomized sequence.

After completing the trial experiment, the participants received instructions for the experiment itself. The experiment contains 80 items: 20 Cantonese words, 20 Cantonese nonwords, 20 Mandarin words and 20 Mandarin nonwords. When the item which appeared on a screen was a real word, they were supposed to press **a** key. In case it was not a real word, they pressed **l** key. Each item appeared for three seconds, and then automatically disappeared, even if the respondent did not press any key. A + symbol appeared after each item for two seconds.

Running the second round encountered technical difficulties, presumably due to firewall in China. Some respondents attempted to run the session several times due to technical problems or reported that their screen froze at the beginning or in the middle of the experiment. Due to firewall issues, most of the respondents were not able to finish the experiment, so the recruiting process had to be repeated several times until there was a sufficient number of completed experiment attempts. In total, 85 attempts (73.9%) were aborted and 30 attempts (26.1%) were successful.

## Data Analysis

Overall data analysis was conducted in R, a software for statistical analysis.

- 1) *One-way ANOVA* was performed to analyse the effect of different categories (MW, MN, CW, CN) on participants' response time.

*Response Time ~ Category (CW, CN, MW, MN)*

Results of *One-way ANOVA* revealed that there was statistically significant difference between the four different categories (( $F= 46.21$ ,  $p = < 0.05$ )).

*ANOVA* was followed by *Levene's test for homogeneity of variance* and *Tukey's HSD* test for post-hoc analysis.

- 2) *Levene's test for homogeneity of variance* indicated unequal variances (( $F= 15.191$ ,  $p = < 0.05$ )).
- 3) *Tukey HSD* test revealed there were significant differences between categories ( $p = < 0.05$ )

## Chi-Square test

*Chi-Square Test* was used to determine whether some categories had higher number of accuracy than other. It was expected Mandarin words and nonwords had the highest accuracy. It was also used for determining the difference between the response time of correct and incorrect answers and response time between Mandarin and Cantonese categories.



## Correct Response

The results revealed following:

- 1) There was a significant difference between:
  - a) Cantonese and Mandarin words (X-squared = 6.0549, df = 1, p-value = 0.01387)
  - b) Cantonese nonwords and Mandarin nonwords (X-squared = 160.4, df = 1, p-value < 0.05)
  - c) correct answers in Cantonese and Mandarin categories overall (X-squared = 13.617, df = 1, p-value = < 0.05)
  
- 2) There was no significant difference between:
  - a) Mandarin words and Mandarin nonwords (X-squared = 0.25098, df = 1, p-value = 0.6164)
  - b) Cantonese words and Cantonese nonwords (X-squared = 0.04186, df = 1, p-value = 0.8379)

## Incorrect Response

The results were following:

- 1) There was a significant difference between:
  - a) Cantonese and Mandarin words (X-squared = 20.093, df = 1, p-value = < 0.05)
  - b) Cantonese nonwords and Mandarin nonwords (X-squared = 25.412, df = 1, p-value = < 0.05)

c) Mandarin words and Mandarin nonwords (X-squared = 4.5849, df = 1, p-value = 0.03225)

d) incorrect answers in Cantonese and Mandarin categories overall (X-squared = 47.004, df = 1, p-value = < 0.05)

2) There was no significant difference between:

a) Cantonese words and Cantonese nonwords (X-squared = 2.4059, df = 1, p-value = 0.1209)

### Correct Response

| Category           | Response Time (mean) | Number | Percentage |
|--------------------|----------------------|--------|------------|
| Mandarin words     | 1.027                | 502    | 83,7 %     |
| Mandarin nonwords  | 1.332                | 518    | 86,3 %     |
| Cantonese words    | 1.159                | 427    | 71,2 %     |
| Cantonese nonwords | 1.487                | 433    | 72,2 %     |

### Incorrect Response

| Category           | Response Time (mean) | Number | Percentage |
|--------------------|----------------------|--------|------------|
| Mandarin words     | 1.291                | 93     | 15,5 %     |
| Mandarin nonwords  | 1.446                | 66     | 11 %       |
| Cantonese words    | 1.418                | 165    | 27,5 %     |
| Cantonese nonwords | 1.552                | 138    | 23 %       |

## 9 Conclusion

The main focus of this thesis were Cantonese-Mandarin bilinguals and their performance in lexical decision task. My aim was to find out if bilingual native speakers who use both languages daily will have different results in lexical decision task aimed at written representation of words. I expected that Mandarin compounds would have a higher rate of correct answers. This confirmed the assumption that Cantonese-Mandarin bilinguals in Guangdong are exposed to Cantonese mainly as a spoken variety, not so much to written Cantonese.

In total, 600 units in each category were analysed. Due to technical issues and firewall in China, some responses were not recorded (in both graphs as *NA* – not answered). In total, there were 58 missing answers. The results showed difference in response time between accurate and inaccurate answers. Participants took the longest time for Cantonese nonwords (mean 1.4), followed by Mandarin nonwords (mean 1.3), Cantonese words (mean 1.2) and Mandarin words (mean 1). According to Chi-square test results, the differences between response time were not significant.

The participants were most accurate with Mandarin words (502 correct answers). Mandarin nonwords had the second highest accuracy (518 correct answers), followed by Cantonese nonwords (433) and Cantonese words (427). Cantonese words had the highest number of incorrect answers (165), followed by Cantonese nonwords (138), Mandarin words (93) and Mandarin nonwords (66).

In general, the results show much accuracy rate in Mandarin than in Cantonese, which was expected. Cantonese is mainly spoken variety and this applies for people in Guangdong especially. While Cantonese speakers from Hong Kong are exposed to Vernacular Written Cantonese to a certain extent, this does not apply completely to Guangdong locals. Due to restrictions regarding usage of Chinese script in public space, Cantonese-Mandarin bilinguals in Guangdong are not exposed to VWC as much as Cantonese-Mandarin bilinguals from Hong Kong. When respondents claimed they used Cantonese daily, they assumably thought of spoken variety. All respondents

said they were also exposed to some kind of media in Cantonese, so some of them might be exposed to written Cantonese as well. Presumably, the results would be different in phonological lexical decision task; there could be a higher percentage of correct answers in both Cantonese words and nonwords.

Low accuracy in Cantonese words and nonwords probably stems from ambiguity of Cantonese lexemes, as I discussed in chapter *Written Cantonese*. Due to lack of input methods and no standardization, many compounds exist in two or more orthographic norms. In Cantonese, phonetic norm comes before the semantic norm: some items contained the mouth radical 口 *hau2*, for instance 嘢 *m2/m6*, 唔 *m4*, 㗎 *diu1* or 㗎 *gaa3*. The mouth radical indicates pronunciation; it is a common component of Cantonese-specific characters. Many nonwords were homophonous with real words in the task, which could have led participants to assumption that they are real lexemes with a different orthographic norm. This is different from Mandarin with a standardized norm, so respondents were able to recognize Mandarin words and nonwords accurately.

## Resumé

Tato práce se zabývala rodilými bilingvními mluvčími standardní čínštiny a kantonštiny z provincie Guangdong. Cílem práce bylo zjistit, jak rozumějí psaným slovům v obou jazycích. Kantonština je v Číně označována za 方言 *fangyan*, což bývá do evropských jazyků překládáno jako *dialekt*. Na začátku jsem proto uvedla jazykové rozdělení čínštiny s cílem podpořit tvrzení, že kantonština je vůči standardní čínštině v rovnocenné pozici druhého jazyka, nikoliv dialektu. V další části jsem uvedla informace o historii a současnosti jazykové politiky v Číně, což je důležité pro pochopení bilingvismu a multilingvismu v Číně. Čínská jazyková politika v současné době neumožňuje vzdělávání v různých čínských jazycích, pouze ve standardní čínštině. To má negativní vliv na znalost těchto jazyků, což je popsáno v kapitole věnující se vzdělávání. Další kapitola popsala psanou kantonštinu a její používání. V poslední části popisuji svůj experiment, *lexical decision task*, ve kterém měli respondenti určit, jestli dané slovo je nebo není reálné. Výsledek podle předpokladů ukázal mnohem vyšší přesnost ve standardní čínštině než v kantonštině. Bilingvní mluvčí z Guangdongu jsou častěji vystaveni psané standardní čínštině než kantonštině, proto slova v psané standardní čínštině rychleji rozeznají.

## Bibliography

- Barov, S. A., & Egorova, M. A. (2019). CANTONESE DIALECT IN MODERN CHINA: THE PROBLEM OF CONSERVATION. *RUDN Journal of Language Studies, Semiotics and Semantics*, 10(1), 152–166. <https://doi.org/10.22363/2313-2299-2019-10-1-152-166> (in Russian)
- Cai, Z. G., Pickering, M. J., Yan, H., & Branigan, H. P. (2011). *Lexical and syntactic representations in closely related languages: Evidence from Cantonese–Mandarin bilinguals*. *Journal of Memory and Language*, 65(4), 431–445. doi:10.1016/j.jml.2011.05.003
- Chambers, J. K., & Trudgill, P. (2004). *Dialectology* (2nd ed.). Cambridge: Cambridge University Press.
- Chappell, H. (2017). Languages of China in their East and Southeast Asian Context. In R. Hickey (Ed.), *The Cambridge Handbook of Areal Linguistics* (Cambridge Handbooks in Language and Linguistics, pp. 651-676). Cambridge: Cambridge University Press. doi:10.1017/9781107279872.025
- Chen, L. (2011). The Imposition of Cantonese on Mandarin in the City of Guangzhou.
- Chen, L., & Joch, N. (2015). Cantonese Love Songs. *Asian Literature and Translation*, 1. <https://doi.org/10.18573/j.2015.10214>
- Chen, P. (2001). Development and standardization of lexicon in Modern Written Chinese. *Language Planning and Language Policy: East Asian Perspectives*. Routledge.
- Chen, P. (1993). Modern Written Chinese in development. *Language in Society*, 22(4), 505-537. doi:10.1017/S0047404500017450
- Cheng, S.P., & Tang, S.W. (2014) Languagehood of Cantonese: A Renewed Front in an Old Debate. *Open Journal of Modern Linguistics*, 04(03), 389-398. doi: [10.4236/ojml.2014.43032](https://doi.org/10.4236/ojml.2014.43032).
- Cheung, K.H., & Bauer, R. S. (2002). The Representation of Cantonese with Chinese Characters. *Journal of Chinese Linguistics Monograph Series*, 18, i–489. <http://www.jstor.org/stable/23826053>
- Cheung, P. (2002). “Guangdong under reform: Social and political trends and challenges,” in

- Davies, A. (2003). *The Native Speaker: Myth and Reality*. Bristol, Blue Ridge Summit: Multilingual Matters. <https://doi.org/10.21832/9781853596247>
- Feng, A., & Adamson, B. (2015). Researching Trilingualism and Trilingual Education in China. In: Feng, A., Adamson, B. (eds) *Trilingualism in Education in China: Models and Challenges*. *Multilingual Education*, vol 12. Springer, Dordrecht. [https://doi.org/10.1007/978-94-017-9352-0\\_1](https://doi.org/10.1007/978-94-017-9352-0_1)
- Francis, N. (2016). Language and dialect in China. *Chinese Language and Discourse*. 7. 136-149. 10.1075/cld.7.1.05fra.
- Gao, X. (2012). ‘Cantonese is not a dialect’: Chinese netizens’ defence of Cantonese as a regional lingua franca. *Journal of Multilingual and Multicultural Development - J MULTILING MULTICULT DEVELOP.* 33. 449-464. 10.1080/01434632.2012.680461.
- Green, R.L. (1994) Accent, standard language ideology, and discriminatory pretexts in the courts. *Language in Society* 23, 2, 163–198; <https://www.jstor.org/stable/4168513>
- Guo, H. & Ni, Z., Wang, Z., Zhang, Y. & Li, J. (2022). Revitalization of Fangyan Through Social Media Promotion in China. *Asian Social Science*. 18. 1. 10.5539/ass.v18n3p1.
- Guo, L. (2004). The Relationship between Putonghua and Chinese Dialects. *Language Policy in the People’s Republic of China*, 45–54. doi:10.1007/1-4020-8039-5\_3
- He, Y. (2006). Some observations on dialectal lexis interference in Putonghua: With a specific reference to question words and interrogative particles in wh-questions of Chengdu Putonghua. *Journal of Asian Pacific Communication*. 16. 279-298. 10.1075/japc.16.2.08he.
- Ho, W-Ch., & Lu, J. (2018). Culture versus the State? The “Defend-My-Mother-Tongue” Protests in Guangzhou. *The China Journal*. 81. 000-000. 10.1086/699253.
- Hyltenstam, K.; Bylund, E.; Abrahamsson, N.; Park, H. S. (2009). "Dominant-language replacement: The case of international adoptees". *Bilingualism: Language and Cognition*. **12** (2): 121–140.
- Kaltenegger, S. (2020). Modelling Chinese as a pluricentric language. *Journal of Multilingual and Multicultural Development*. 1-12. 10.1080/01434632.2020.1810256.

- Kurpaska, M. (2019). How to Create a Language for the Whole People? On the Codification of Putonghua Pronunciation. *Roczniki Humanistyczne*. 67. 89. 10.18290/rh.2019.67.9-4.
- Kwok, B-C., Chin, A. & Tsou, B. (2016). Grammatical diversity across the Yue dialects. *Journal of Chinese Linguistics*. 44. 109-152. 10.1353/jcl.2016.0002.
- Landry, R., & Allard, R. (1994). Diglossia, ethnolinguistic vitality, and language behavior. *International Journal of the Sociology of Language*, 108, 15–24.
- Li, D. C. (2000). Phonetic Borrowing: Key to the vitality of written Cantonese in Hong Kong. *Written Language and Literacy*, 3(2), 199–233. <https://doi.org/10.1075/wll.3.2.02li>
- Liang, S. (2014). Problematizing monolingual identities and competence in Guangzhou in the era of multilingualism and superdiversity. In E. Esch & M. Solly (Eds.), *Language education and the challenges of globalisation: Sociolinguistic issues*. Cambridge: Cambridge Scholars Publishing.
- Liang, S. (2015). *Language Attitudes and Identities in Multilingual China*. doi:10.1007/978-3-319-12619-7
- Liu, J. (2016). A Historical Review of the Discourse of Fangyan in Modern China. *Twentieth-Century China*. 41. 1-17. 10.1080/15215385.2016.1205788.
- Liu, Y., Shu, H., & Li, P. (2007). Word naming and psycholinguistic norms: Chinese. *Behavior Research Methods*, 39(2), 192–198. <https://doi.org/10.3758/bf03193147>
- Lim, R. Y., Yap, M. J., & Tse, C.-S. (2020). Individual differences in Cantonese Chinese word recognition: Insights from the Chinese Lexicon Project. *Quarterly Journal of Experimental Psychology*, 73(4), 504–518. <https://doi.org/10.1177/1747021820906566>
- Li, Y. & Li, D. & Gao, X. (2019). The complexity of family language policy decisions: The case of Cantonese and other regional Chinese varieties. *Círculo de Lingüística Aplicada a la Comunicación*. 79. 63-78. 10.5209/clac.65648.



Love, N., & Ansaldo, U. (2010) The native speaker and the mother tongue, *Language Sciences*, Volume 32, Issue 6, 2010, Pages 589-593, ISSN 0388-0001, <https://doi.org/10.1016/j.langsci.2010.09.003>.

Ma, J., Wu, Y., Sun, T., Cai, L., Fan, X., & Li, X. (2020). Neural substrates of bilingual processing in a logographic writing system: An fMRI study in Chinese Cantonese-Mandarin bilinguals. *Brain Research*, 1738, 146794. doi:10.1016/j.brainres.2020.14679

Mair, V.H. (1990). Two Non-Tetragraphic Northern Sinitic Languages a . Implications of the Soviet Dungan Script for Chinese Language Reform b . Who Were the Gyámi ?

Mair, V. H. (1991). What Is a Chinese “Dialect/Topolect”? Reflections on Some Key Sino-English Linguistic Terms. *Sino-Platonic Papers*, 29, 1-31.

Milroy, J. (2001). Language Ideologies and the Consequence of Standardization. *Journal of Sociolinguistics*. 5. 530 - 555. 10.1111/1467-9481.00163.

Pan, Y. (2000). Code-switching and social change in Guangzhou and Hong Kong. *International Journal of the Sociology of Language*, 146(1). doi:10.1515/ijsl.2000.146.21

Pérez-Quiñones, M., & Salas, C. C. (2021). How the ideology of monolingualism drives us to monolingual interaction. *Interactions*, 28(3), 66–69.

<https://doi.org/10.1145/3457871>

Rohsenow, J.S. (2004). Fifty Years of Script and Written Language Reform in the P.R.C.. In: Zhou, M., Sun, H. (eds) *Language Policy in the People’s Republic of China*. Language Policy, vol 4. Springer, Dordrecht. [https://doi.org/10.1007/1-4020-8039-5\\_2](https://doi.org/10.1007/1-4020-8039-5_2)

Saillard., C. (2004). On the Promotion of *Putonghua* in China: How a Standard Language Becomes a Vernacular. In: Zhou, M., Sun, H. (eds) *Language Policy in the People’s Republic of China*. Language Policy, vol 4. Springer, Dordrecht. [https://doi.org/10.1007/1-4020-8039-5\\_9](https://doi.org/10.1007/1-4020-8039-5_9)

Sautman, B. & Xie, X. (2021). Today in Guangzhou, Tomorrow in Hong Kong? A Comparative Study of the Language Situation in Two Cities. *Journal of Current Chinese Affairs*. 49. 186810262098393. 10.1177/1868102620983939.

- Snow, D. (2004). *Cantonese as Written Language: The Growth of a Written Chinese Vernacular*. Hong Kong University Press. [ISBN 978-962-209-709-4](https://doi.org/10.1017/japc.18.2.05sno).
- Snow, D. (2008). Cantonese as written standard?. *Journal of Asian Pacific Communication*. 18. 190-208. 10.1075/japc.18.2.05sno.
- Song X. (2018). On the Nature and Development of Standard Mandarin Education in Hong Kong
- Spolsky, B. (2014). Language management in the People's Republic of China. *Language*, 90, e165 - e179.
- Tsang, Y-K., & Huang, J., Lui, M., Xue, M., Chan, F., Wang, S. & Chen, H-C. (2017). MELD-SCH: A megastudy of lexical decision in simplified Chinese. *Behavior Research Methods*. 50. 10.3758/s13428-017-0944-0.
- Wang, H., & Yuan, Z. (2013). 2 The promotion of Putonghua (Mandarin Chinese): An overview. 10.1515/9781614512530.27.
- Wang, L., & Ladegaard, H. J. (2008). *Language Attitudes and Gender in China: Perceptions and Reported Use of Putonghua and Cantonese in the Southern Province of Guangdong*. *Language Awareness*, 17(1), 57–77. doi:10.2167/la425.0
- Wang, W. (2017). Cantonese with an Accent Identity and Symbolic Power in Guangdong. 10.2991/ichssr-17.2017.12.
- Wang, W. (2019). Mapping Cantonese: The Pro-Cantonese Protest and Sina Weibo in Guangzhou. 10.1007/978-3-319-73400-2\_100-1.
- Xing G. (1982) *Hanyu fangyan diaocha jichu zhishi [Fundamental Knowledge for the Investigation of Sinitic Topolects]*. Wuchang, Hupeh: Huazhong gongxue yuan shubanshe
- Xiong, T. (2018). Xiong 2018 Media Representation of Attitudes to Multilingual education.
- Yan, J. (2008). Social Variation of Vernacular Written Cantonese in Guangzhou (Canton City), China.
- Yan, J. (2012). WRITING CANTONESE AS EVERYDAY LIFESTYLE IN GUANGZHOU (CANTON CITY). *Chinese Under Globalization*, 171–202. doi:10.1142/9789814350709\_0009

- Yan, J. (2022). Linguistic Convergence and Divergence in Guangzhou (Canton City): Social Variation of Vernacular Written Cantonese.
- Zhao, Hui & Liu, Hong. (2020). (Standard) language ideology and regional Putonghua in Chinese social media: a view from Weibo. *Journal of Multilingual and Multicultural Development*, 42. 1-15. 10.1080/01434632.2020.1814310.
- Zheng, S. D. S. (1992). *From Toison to New York: Muk'yu Songs in Folk Tradition*. *CHINOPERL*, 16(1), 165–205. doi:10.1179/chi.1992.16.1.165
- Zhou, M. (2010). China: The Mainland, Hong Kong, and Taiwan. In J.A. Fishman & O. García (Eds.), *Handbook of language and ethnic identity: Disciplinary and regional perspectives* (pp. 470–485). Oxford: Oxford University Press.
- Zhou, M. (2015). Nation-state building and rising China: PRC's discourse on the Chinese language since the turn of the 21st century. In L. Tsung & W. Wang (Eds.), *Contemporary Chinese discourse and social practice in China* (pp. 59–80). London: Continuum.
- Zhou, M. (2016). Nation-state building and multiculturalism in China. In X. W. Zang (Ed.), *Handbook on ethnic minorities in China* (pp. 111–137). Cheltenham: Edward Elgar Publishing
- Zhou, M. (2017). Language ideology and language order: Conflicts and compromises in colonial and postcolonial Asia. *International Journal of the Sociology of Language*, 243, 97–118
- Zhou, M. (2018). Language classification: Nation-state building and globalization. *Chinese Journal of Language Policy and Planning*, 3(2), 6–15.
- Zhou, M. (2019). *Language Ideology and Order in Rising China*. doi:10.1007/978-981-13-3483-2
- Zhu, Q. (2021). Multilingualism in urban China. 10.4324/9780429286056-3.
- Zhu, W.J.; & Chen, J.M. (1991). Some aspects of the language Situation in China. *Journal of Asian Pacific Communication* 2, 91-101.

## Newspaper articles

Cheng, K. (2020, March 31). *Cantonese a dialect, not a mother tongue, says Hong Kong Education Bureau supporting material on Mandarin*. Hong Kong Free Press HKFP. <https://hongkongfp.com/2018/05/02/cantonese-dialect-not-mother-tongue-says-hong-kong-education-bureau-supporting-material-mandarin/>

Tam, J. (2018, June 25). *Education Bureau rapped over Cantonese “not an official language” gaffe*. South China Morning Post.

<https://www.scmp.com/news/hong-kong/article/1419237/education-bureau-rapped-over-cantonese-not-official-language-gaffe>

He, X. (2008). 好多广州细路唔识讲白话 (Many Guangzhou children cannot speak Cantonese). Guangzhou Daily, A19.

学校要求学生讲普通话祖孙俩竟变“鸡同鸭讲” [由于学校规定，奶奶和孙女无法互相交流]。羊城晚报（用中文（表达））。2010-07-09。归档自原本的于2010-07-12。已检索 2010-07-27.

Hu, J., & Zi, Y. (2010). 越秀区执信南路小学“要求”学生不管是上课还是课余时间都要讲普通话 (Zhixin Nan Lu Primary School in Yuexiu District requires students to speak Putonghua whether it is during or after class). Yangcheng Evening Post, A17.

## Internet resources

Mair, V., & Mair, V. (2016). *Language Log » Language vs. script*. Retrieved April 27, 2022, from <https://languagelog.ldc.upenn.edu/nll/?p=29443>

*Taiwanese, Mandarin, and Taiwan’s language situation*. (2003). Retrieved April 27, 2022, from <http://www.pinyin.info/readings/mair/taiwanese.html>

