Variations in World Chineses
Jingxia Lin; Dingxu Shi; Menghan Jiang; Chu-Ren Huang

Abstract
Although Mandarin Chinese is shared by Chinese communities such as Mainland China, Taiwan, Hong Kong, and Singapore, linguistic differences are frequently found among regional uses, ranging from pronunciation, orthography, vocabulary, grammar, and discourse. Along with the increasingly recognized notion of "World Chineses" in recent years, the study of the regional variations has also become more linguistically, socially, and culturally significant. Such a study facilitates more efficient communication among speakers of different varieties, reflects the social and cultural differences of the Chinese speaking communities from a linguistic perspective, and contributes to the theoretical discussion of language variation and change. With specific examples of the linguistic features exhibited in Mainland China, Taiwan, Hong Kong, and Singapore Mandarin Chinese, this chapter is an overview of the current studies, methodologies, and motivations of variation.

Introduction
The term World Chineses, though not as common as World Englishes, is becoming more and more widely used together with the growing popularity of Mandarin Chinese as a second language and the spreading of Chinese diaspora. Like World Englishes, differences are often found in the Mandarin Chinese used in different regions. For instance, the expression 看看一下 *kànkànyīxià* 'take a glimpse' is acceptable in Singapore Mandarin, but Mandarin speakers elsewhere prefer either 看看 *kànkàn* 'take a glimpse' or 看一下 *kànyīxià* 'take a glimpse'. The (dis-) preference does not imply that one expression is better than the other. There is no denying that speakers of any variant of Chinese, such as those from Singapore or different provinces of Mainland China, are *bona fide* speakers of Chinese. Hence Chinese cannot be defined by a single "variety". The preferential differences over the alternative usages of 看看一下 *kànkànyīxià* or 看看/看一下 *kànkàn/kànkànyīxià* represent a case of language variation. That is, these expressions are alternative forms of saying the same thing, even though one expression may be preferred by certain groups of speakers or communities for some reason (Labov 1972). Hence, while typical grammatical study on Chinese, such as Huang and Shi's (2016) reference grammar, focuses on the shared generalizations of all varieties, increasing awareness leads to more and more recent studies on the variation of Mandarin. Among the studies, two important questions are how to identify the variations and what are the linguistic motivations of the variations.

Methodologies to study variations in World Chineses
Corpora and the study of grammatical variations
The earliest studies on World Chineses were typically based on researchers' observations and introspection, and mainly on lexical and grammatical variations. The earliest lexical

studies often are limited to the listing of different lexical items (e.g. Tang 1995; Wang 1999). In recent years, the availability of comparable corpora (i.e. two or more corpora with similar topics and coverage) of different varieties of Chinese has enabled wider coverage and more comprehensive studies. For instance, the comparable synchronic Chinese corpus LIVAC (T'sou and Kwong 2006) is the empirical basis of T'sou and You (2010), a dictionary of region-specific neologism in Mainland China, Hong Kong, Taiwan, and Singapore from 2000-2009. With this and other similar studies (e.g. Li 2014; Li 2016) it is fair to say that lexical variations of world Chineses are well documented.

The studies of grammatical variation of World Chineses, on the other hand, are relatively less advanced. A good number of existing grammatical studies focus on observations on individual grammatical constructions, e.g., Chen (1986), Chew (2007) and Lu (2002) on Singapore Mandarin, Tin and Ma (2013) and Zhao and Shi (2012) on Hong Kong Mandarin, and Diao (2000) and Tseng (2003) on Taiwan Mandarin. Shi et al. (2006, 2014) is among the first that present the grammatical variation in Hong Kong Mandarin in a systematic and more exhaustive way. *Global Chinese Grammar* (Forthcoming) is the first that aims to comprehensively describe the grammar of Mandarin Chinese used in six regions (Taiwan, Hong Kong, Macau, Singapore, Malaysia, and USA). The existing studies of grammatical variation share two features in terms of methodology. Firstly, the majority limits the comparison to only two varieties, with one of them very often being Mainland China Mandarin, i.e. *putonghua* 'common speech'. As a consequence, the grammatical features shared among non-Mainland Chinese varieties are often neglected. For instance, the expression "directional motion verb + deictic complement 来/去 *lái/qù* 'come/go' + locative NP" (e.g. 回来新加坡 *huílái xīnjiāpō* 'come back to Singapore') is allowed in Taiwan, Hong Kong, and Singapore Mandarin, and thus not a variation between the three; but such an expression is considered special when it is only compared with Mainland China Mandarin. Secondly, although more and more recent studies rely on authentic data such as corpora for identifying possible variation, only a few exceptions (e.g. Huang et al. 2014; Lin et al. 2014) adopt computational and statistical tools to process the data. With the availability of comparable corpora, e.g. LIVAC (T'sou and Kwong 2006) and Tagged Chinese Gigaword (Huang 2009), and computational/statistical analytical tools, it is possible now to effectively identify differences among multiple Chinese varieties and to precisely describe the nature and motivation of variations. We will discuss these methodological innovations next.

Comparable corpus driven, statistic model based approaches to World Chinese Variations

A crucial development in the study of grammatical variations of different Chinese varieties is the availability of comparable corpora from different varieties as well as the new corpus driven, statistic models of variations and computational tools to implement them. Compared with lexical variation, grammatical variations are not easy to detect with the traditional approach of introspective judgment. This is because the differences among variants often lie in the presence/absence of a tendency rather than the grammaticality dichotomy (with is more typical of two different languages). Hence a more objective, data driven research methodology is needed to identify and analyze the frequency of

collocational differences and/or the preference of certain alternatives among different varieties (e.g. Huang et al. 2013; Huang et al. 2014).

This comparable corpus driven, statistical modelling based approach to variations of World Chineses has been shown to be innovative and effective. Huang et al. (2014) and Lin et al. (2014) conducted both statistical and data mining approaches to identify and classify the light verb variations, which are known to be notoriously difficult both linguistically and computationally. They studied the full set of more frequent light verbs in Mandarin Chinese: 进行/加以/做/搞/从事 *jìnxíng/jiāyǐ/zuò/gǎo/cóngshì* 'to conduct'. The aims were not only to detect grammatical differences between Mainland China and Taiwan Mandarin, but also to differentiate the grammaticality constraints of different light verbs. The account for grammatical differences among different light verbs is especially challenging given the lightness (i.e. lack of explicit grammatical constraints) of these verbs. With corpus data that was annotated with carefully selected grammatical features related to light verbs, their statistical (both univariate and multivariate analysis) and data mining (K-means clustering algorithm) methods consistently showed that the uses of different light verbs as well as differences between the two varieties could be rigorously modelled and automatically identified. These studies offered one of the strongly empirical evidence of preference based models for variations of world Chineses.

In addition to light verb variations, other grammatical variations can also be detected by this comparable corpus driven approach. The transitivity variations of VO compounds (i.e. "VO+O" construction, such as 把关产品质量 *bǎguānchǎnpǐnzhìliàng* 'to safeguard the product quality') were shown to have statistical differences (Jiang et al. 2015). Jiang et al. (2016) further showed that the variation differences in distribution as well as the selection constraints for "VO+O" constructions can also be effectively identified by univariate (Chi-squares test, One-Way ANOVA and binary logistic regression) and multivariate (e.g. polytomous logistic regression) models and thus support a degree of transitivity based account.

In sum, grammatical variations of World Chineses display language change in action. The multiple varieties of Chinese as well as abundance of data offer a unique opportunity for linguists to explore and validate different theories of language variations and changes. Richly annotated language resources paired with appropriate tools can lead to effective general solutions for subtle grammatical variation classification, identification and detection.

Factors contributing to the variation
Studies on language change and variation, such as (Labov 1972 and Weinreich et al. 1968), have long established that language cannot remain static and that motivations for language variations can be largely classified as: (a) language-internal factors, i.e. the linguistic characteristics of a language; (b) language-external factors, e.g., contacts and socio-interactive factors. This chapter focuses on Mandarin varieties in different regions, so language contact, the major cause of variation, will be discussed in detail, whereas other external factors (e.g. gender and class) that can motivate variation within a single variety

will be covered when needed. The terms MCM, TWM, HKM, and SM will be adopted for the Mandarin in Mainland China, Taiwan, Hong Kong, and Singapore, respectively.

Language-internal factors
Language internal factor refers to the motivation that leads to variation by the properties of a language itself, e.g. the structural properties of a language and the tendency of a language to become more natural or less marked. In such a process of variation, no obvious external influence (e.g. social factors such as language contact) is identified. Language variation motivated by internal factors happen in all varieties of a language, because even though the varieties share the same linguistic system, they may show different effects that are triggered by the internal factors.

The well-known VO separable compounds offer a good example.SM shows a degree of lexicalization that differs from MCM. For instance, 帮忙 *bāngmáng* 'help' and 生气 *shēngqì* 'be angry' are not only used in non-separable forms but can take objects in SM, as in 帮忙他 *bāngmángtā* 'help him' and 生气老公 *shēngqìlǎogōng* 'be angry with husband'. But in MCM, the objects typically occur in between V and O, i.e. 帮他忙 *bāngtāmáng* 'help him' and 生老公气 *shēnglǎogōngqì* 'be angry with husband'. Jiang et al (2016) argue that the differences may be motivated by differences in the degree of transitivity assigned to these compounds.

Variation can also be found in the semantic and syntactic types of the complements taken by light verbs in TWM and MCM. In MCM, the light verbs 进行 *jìnxíng* 'proceed' and 从事 *cóngshì* 'undertake' typically take as complements nominal or verbal compounds that denote formal and positive events (e.g., 访问 *fǎngwèn* 'visit' and 研究 *yánjiū* 'research'). However, the use of the light verbs in TWM has been more extended (Huang et al. 2012, Huang et al. 2014; Lin et al. 2014). For instance, 进行 *jìnxíng* can take verbal phrases such as 开会 *kāihuì* 'have a meeting' and 投票 *tóupiào* 'cast a vote' as its complement and the sentimental preference of 从事 *cóngshì* has been neutralized so that it can take complements with a negative connotation, e.g. 性交易 *xìngjiāoyì* 'sex trade' and 勾当 *gòudang* 'shady business'.

Another example of internally motivated variations involves changes in grammatical categories. Variation arises when the categorical change happens to different words in different varieties. For instance, the noun 津贴 *jīntiē* 'subsidy' in HKM can be used as a verb (1a), whereas the adjective 亲爱 *qīn'ài* 'dear' in SM can be used as a transitive verb (1b).

(1) a. HKM: 为了一个迪士尼乐园，特区政府已经津贴了巨资 (*Oriental Daily*-2000-10-23, cited from Shi et al. 2006)
       wèile yī gè díshìní lèyuán, tèqū zhèngfǔ yǐjīng jīntiēle jù zī
       'For Disneyland alone, the SAR Government subsidized (it)with huge capital.'
   b. SM: 亲爱父母和兄弟姐妹，就是仁的表现 (*Ethics*- Middle school 4th Grade)
       qīn'ài fùmǔ hé xiōngdì jiěmèi, jiùshì rén de biǎoxiàn

>‘To show endearing affection to one’s parents and siblings is the instantiation of ren-humanness.’

It is possible that the internally motivated variation found in only one variety of Mandarin currently will also occur in other varieties in the future, given that all varieties of Mandarin share the same typological features. Such variations in fact offer a rare opportunity for us to observe the different stages of language change with synchronic data and allow us to test the hypothesis of how the changes are motivated based on the assumption that the early-adoption of change in one variety should be accounted for by the fact that it provides a more optimal environment for this particular change.

Language-external factors
“External factors” of language variation are often related to the speech community of the language. For Mandarin varieties, especially for MCM and TWM, many variations can be attributed to both the long separation as well as to differences in linguistic standards set for MCM and TWM after the Chinese Civil War. For instance, the two varieties adopt different writing systems, with MCM in simplified Chinese characters and TWM in traditional Chinese characters. Furthermore, variations can be frequently found in the sound of the two varieties. The comparison by Zhang (2009) based on *Contemporary Chinese Dictionary* (2005) and *New Chinese Daily* Dictionary (2000) shows that 1,080 characters are pronounced differently. For instance, 戏谑 ‘joking’ and 混淆 ‘confuse’ are *xìxuè* and *hùnxiáo* in MCM, but *xìnüè* and *hùnyáo* in TWM.

In terms of language contact, the main sources of contact to the varieties of Chinese are the other Sinitic language(s)[1] spoken locally, or other local (indigenous) non-Sinitic languages. In addition, English as the global language and as the past colonial language has also contributed to the contact. The actual contexts of contacts depend on the local linguistic environment, including the type of languages/dialects spoken locally, and the local language and education policy.

Mandarin is the sole official language and the language of education in Mainland China and Taiwan. English is learned as a second language in these two regions. According to the Three Circles Model on the diversity in the spread of World Englishes (Kachru 1985), Mainland China and Taiwan belong to the Expanding Circle, where English is used only as a foreign language. Other than English, foreign languages are not widely taught in Mainland China and Taiwan. However, it is worth noting that Japanese, as a past colonial language, has contact influence on older people in Taiwan, and it influences the younger generations everywhere through media/cultural borrowing. More recently, Korean contact is prevalent in Taiwan through media and cultural influences.[2]

There are six major non-Mandarin Sinitic languages in China: Wu, Xiang, Cantonese, Min, Gan, and Hakka. MCM has borrowed words from these dialects, e.g. 炒鱿鱼 *chǎoyóuyú* ‘fire’ and 拍拖 *pāituō* ‘date’ from Cantonese, 发嗲 *fādiǎ* ‘acting in a cutsey-pootsy way’ and 瘪三 *biēsān* ‘a wretched-looking tramp without proper employment’ from Wu (Su 2001). However, generally speaking, MCM is not heavily influenced by any single Sinitic language.

Southern Min and Hakka have been dominant local languages in different parts of Taiwan since the first wave of immigration of Han Chinese people in the 17th century from Southern China (mostly from Fujian and Guangdong). Mandarin was introduced to Taiwan in 17th century by Ming and Qing dynasty Mandarins but only became the functional official language and had to be learnt compulsorily after Taiwan's restoration to the Republic China in 1945. A 2010 survey of the relative frequency of languages used at home in Taiwan shows that Mandarin, Taiwanese (i.e. Southern Min), Hakka, and indigenous languages are at a rate of 83.5, 81.9, 6.6, and 1.4, respectively per hundred resident nationals (6 years and above).[3] The data suggests that while Mandarin is the official language, Southern Min is the dominant local dialect with close contact to TWM, and thus being the major source of the variations in TWM.

English and Cantonese have been the dominant language and dialect in Hong Kong before the handover of Hong Kong to China in 1997. Since then, the SAR Government promotes the policy of biliteracy (English and Chinese in writing) and trilingualism (English, Putonghua, and Cantonese in speech). Putonghua was introduced to the curriculum in Hong Kong primary and middle schools, although English and Cantonese remain as the main media of instruction. In the meantime, Putonghua becomes more and more popular due to increasing interactions with Mainland China (Shi et al. 2014). As such, HKM arises as a variety of Mandarin with influence from English and Cantonese.

Singapore has four official languages: English (also the language of administration), Malay (also the national language), Mandarin Chinese, and Tamil. In addition to the official languages, several Chinese dialects are spoken by the ethnic Chinese Singaporeans, including Min Chinese (e.g. Southern Min, Teochew, Hainanese, etc.), Cantonese, and Hakka. These dialects have been the major language of communication among Chinese Singaporeans before the government started to promote the Speak Mandarin Campaign in 1979. Starting from 1987, the bilingualism policy of Singapore requires Singaporeans to acquire English and their mother-tongue. The mother-tongue is defined by each Singaporean's ethnic identification rather than the first language that a Singaporean acquires or the major language that a Singaporean uses in family. For example, an ethnic Chinese should learn Mandarin, and an ethnic Malay should learn Malay. Within the few decades from 1980 to 2010, the most frequently spoken language at home in Singapore underwent a remarkable change: English increased from 10.2% to 32.6%, Mandarin from 13.1% to 47.7%, whereas Chinese dialects decreased from 76.2% to 19.2%.[4] As for the Chinese dialects, Min has been the dominant dialect since 1881, followed by Cantonese and Hakka. In sum, SM has been in close contact to multiple languages and dialects, including English, Malay, Tamil, and Chinese dialects, especially Min, Cantonese, and Hakka.

English as the global language
English as a global language has had a significant impact on local languages, particularly in Singapore and Hong Kong, which belong to the Outer Circle of the Three Circles Model (Kachru 1985). The influence of English can be identified in almost every level of variation, including syntax, discourse, and vocabulary.

Influenced by the English conjunction *and*, the use of the conjunction 和 *hé* shows variations in HKM and SM. When connecting verbal elements, 和 *hé* 'and' traditionally is typically constrained to verbs that share the same adverbials or objects in MCM, e.g., the adverb 进一步 *jìnyībù* 'a further step' and the object 明年的财务预算 *míngnián de cáiwùyùsuàn* 'the financial budget of the next year' in (2) (Lü 1999[1980]).

(2) a. 事情还要进一步调查和了解。(Lü 1999[1980]: 266)
　　　shìqing hái yào jìnyībù diàochá hé liǎojiě
　　　'The issue needs further investigation and understanding.'
　　b. 会议讨论和通过了明年的财务预算。(Lü 1999[1980]: 266)
　　　huìyì tǎolùn hé tōngguò le míngnián de cáiwù yùsuàn
　　　'The meeting discussed and approved the financial budget for next year.'

和 *hé* 'and' in SM and HKM has been extended to additional contexts. For instance, the two verbal phrases in (3) are modified by different adverbials (依时 *yīshí* 'on time' and 一起 *yīqǐ* 'together'), whereas the two verbs in (3b) take different objects (一名建筑工人 *yīmíngjiànzhúgōngrén* 'a construction worker' and 3 个人 *sāngèrén* 'three people').

(3)a. HKM: 请各学员依时出席上堂和一起学习。(www.keichun.org, cited from Shi et al. 2014: 205)
　　　　qǐng gè xuéyuán yīshí chūxí shǎngtáng hé yīqǐ xuéxí
　　　　'All learners are expected to attend the class on time and learn together.'
　　b. SM: 嫌犯星期一在一家建筑公司内，枪杀一名建筑工人，和射伤 3 个人。
　　(*Channel 8-Sina Weibo News*-2016-9-1)
　　　　xiánfàn xīngqīyī zài yī jiā jiànzhù gōngsī nèi, qiāngshā yī míng jiànzhù gōngrén hé shèshāng sān gè rén
　　　　'The suspect killed a construction worker and injured three others by shooting in a construction company on Monday. '

相信 *xiāngxìn* 'believe' is another example that is not as constrained in SM and HKM as in MCM. In MCM, 相信 *xiāngxìn* 'believe' is a psych verb and takes nominal or clausal objects, with the subject being the person who thinks that something or some event is true.

(3) a. 我相信他。
　　　wǒ xiāngxìn tā
　　　'I believe in him.'
　　b. 我相信他会成功的。
　　　wǒ xiāngxìn tā huì chénggōng de
　　　'I believe that he will succeed.'

In SM and HKM, 相信 *xiāngxìn* 'believe' is commonly used in texts where the subject is raised from the clause expressing the belief but often not the experiencer of believing. For instance, although 日本游客 *rìběnyóukè* 'Japanese tourist' and 这些症状 *zhèxiē*

zhèngzhuàng 'these syndromes' are the syntactic subject of 相信 *xiāngxìn* 'believe'in (5a-b), they are not the experiencer of believing. Such an extended use of 相信 *xiāngxìn* 'believe' corresponds to English *It's believed…* and thus is very likely a transfer from English.

(5) a. SM: 一名日本游客相信是因为驾驶的水上摩托艇失控，撞上另一艘客轮，当场丧命。(*Channel 8-Sina Weibo News*- 2016-9-23)
   yī míng rìběn yóukè xiāngxìn shì yīnwèi jiàshǐ de shuǐshàng mótuōtǐng shīkòng zhuàngshang lìng yī sōu kèlún dāngcháng diūmìng
   'It is believed that a Japanese tourist lost life because the water jet he drove ran out of control and hit another passenger boat.'
   b. HKM: 这些症状相信是源于脑部神经中多种化学物质活动失调。
   (http://www3.ha.org.hk/cph/imh/mhi/article_02_03_01_chi.asp)
   zhè xiē zhèngzhuàng xiāngxìn shì yuányú nǎobù shénjīng zhōng duō zhǒng huàxué wùzhì huódòng shītiáo
   'It is believed that these symptoms are caused by the disorders of a variety of chemical activities in the brain nerve.'

In terms of discourse, code-mixing and code-switching with English are frequently observed in SM and HKM. For code-mixing, i.e. the use of English items in Mandarin context, content words are often replaced with English words following the grammar of Mandarin (Shi et al. 2006). For instance, when the adjective *raw* modifies the head noun 感觉 *gǎnjué* 'feeling' in (6a), the particle 的 *de* that functions as a modifier marker is used; the verb *keep* in (6b) is followed by a resultative complement 住 *zhù* (lit.) 'hold' as a Chinese verb does; and the noun *console* is modified by a classifier 个 *gè* in (6c).

(6) a. HKM: 正如牛仔裤一样，穿得久了，更能为梳化添上一点 raw 的感觉。(*Sing Pao*-2004-9-22, cited from Shi et al. 2006)
   zhèngrú niúzǎikù yīyàng, chuān de jiǔ le, gèng néng wèi shūhuà tiānshang yīdiǎn raw de gǎnjué
   'Just like jeans that have been worn for a long time, it [using for a long time] can add some raw feeling for sofas.'
   b. HKM: 因为我妈妈认同弹性活肤精华的概念，就算皮肤几好，都有衰老的一天，所以一路保鲜，keep 住最佳状态。(*Sing Tao Daily*-2004-9-21, cited from Shi et al. 2006)
   yīnwèi wǒ māma rèntóng tánxìng huófū jīnghuá de gàiniàn, jiù suàn pífū jī hǎo, dū yǒu shuāilǎo de yītiān, suǒyǐ yīlù bǎoxiān keep zhù zuì jiā zhuàngtài
   'My mom buys the rationale behind elastic skin essence. Even for (a person with) very good skin, her skin will become aged some day. So (she) should keep it refreshed all the time in order to keep (her skin) in the best condition.'
   c. SM: 我觉得很 interesting 是因为那幅墙画了一个 console (*Home Décor Survivor*-S5-ep2)
   wǒ juéde hěn interesting shì yīnwèi nā fú qiáng huà le yī gè console
   'I thought it very interesting because a console was drawn on that wall.'

English speech act words, such as greeting and apology, are frequently adopted in conversation, and can be used even more often than their Mandarin counterparts. For example, in the first eight episodes of *Home Décor Survivor (Season 5)*, a Singapore Mandarin variety TV show in 2016, *sorry* (22 instances) was used twice more often than 对不起/抱歉/不好意思 *duìbùqǐ/bàoqiàn/bùhǎoyìsi* 'sorry' (10 instances).

Code-switching, where a speaker switches to English for full clauses and sentences, is also very common in HKM and SM, as in (7).

(7) SM: 我们现在终于可以摆家具了。So ladies and gentlemen are you ready? (*Home Décor Survivor-S5-ep2*)
  wǒmen xiànzài zhōngyú kěyǐ bǎi jiājù le
 'Now we finally can display the furniture. So ladies and gentlemen are you ready? '

The most significant influence of English loan words on variations of World Chinese involves translation and transliteration variants, and especially those of proper names (Huang et al. 2007; Šimon et al. 2008). Such variations can be used as evidence of the time and context of borrowing as well as a tool for investigating social and cultural motivations for the differences. For instance, the fact that Jesus is translated as 耶稣 *yēsū* shows that it was translated from a Romance language to a southern Sinitic language (and definitely not from English to Mandarin). However, these variations also post serious challenges in web search and information retrieval as the same entity will have different names in different contexts. A small sample of examples based on words of English origin is given in Table 1. Note that some variety may choose semantic translation over transliteration, such as the translation of *bus*.

Table 1 Words of English origins in Mandarin varieties

English	MCM	TWM	HKM	SM
Bus	公共汽车 *gōnggòngqìchē*	公车 *Gōngchē*	巴士 *bāshì*	巴士 *bāshì*
Cherry	樱桃 *yīngtáo*	樱桃 *yīngtáo*	车厘子 *chēlízi*	樱桃 *yīngtáo*
Julia Roberts	朱莉娅罗伯茨 *zhūlìyàluōbócí*	茱莉亚萝勃兹 *zhūlìyàluóbózī*	茱莉亚罗拔斯 *zhūlìyàluōbásī*	茱莉亚罗拔丝 *zhūlìyàluōbásī*
Zika	寨卡 *zhàikǎ*	兹卡 *zīkǎ*	寨卡 *zhàikǎ*	兹卡 *zīkǎ*

Other languages
Lexical variations can arise from contact with different languages. The most significant examples are the borrowings from Japanese in TWM and from Malay in SM. There are also loanwords from other contact languages (Austronesian and Portuguese in TWN, Tamil and other South Asian languages in SM and HKM). A few examples of TWM specific lexical items borrowed from Japanese are given in Table 2.

Table 2 TWM words of Japanese origin

TWM	Japanese origin
便当 *biàndang* 'boxed lunch'	弁当 *bento*
撒西米 *sāxīmǐ* 'sashimi'	刺身 *sashimi*
观光客 *guānguāngkè* 'tourist'	観光客 *kankokyaku*
阿莎力 *āshālì* 'without hesitation, decisive'	あっさり *assari*
坪 *píng* 'unit of land measurement'	坪 *tsubo*

In Singapore, Malay is the second largest mother tongue following Mandarin, as Malay Singaporeans compose the second largest ethnic group of the population (13.4% in 2010).[5] In addition, Malay is familiar to Singaporeans of other ethnic groups, because it is the national language, as well as the language of the National Anthem and the language of command in the armed forces. Tamil (the largest Indian language spoken by ethnic Indian Singaporeans) is less familiar to Chinese Singaporeans, but the effects of its contact to SM are still visible. Table 3 lists a few examples of SM words that originated from Malay and Tamil.

Table 3 SM words of Malay/Tamil origin

SM	Origin
巴冷刀 *bālěngdāo* 'a kind of machete or cleaver'	*parang*(Malay)
叻沙 *lèshā* 'a kind of spicy noodle soup'	*laksa*(Malay)
甘榜 *gānbǎng* 'village'	*kampong*(Malay)
巴刹 *bāshā* 'market'	*pasar*(Malay)
苏东 *sūdōng* 'blur, muddle-headed' (lit.) 'squid'	*sotong*(Malay)
罗地 *luódì* 'a kind of a flat bread'	*roti*(Tamil)
纱丽 *shālì* 'a kind of female garment'	*sari*(Tamil)
屠妖节 *túyāojié* 'a Hindu festival'	*Deepavali*(Tamil)

Grammatical variations induced by contact from a language other than English is less significant in Mandarin varieties currently, which is probably due to the fact that no other languages are in extended contact with Mandarin Chinese now. One example discussed in literature is the sentence final (的)样子 *(de)yàngzi* 'looks' in TWM, which is probably resulted from earlier contact. According to Diao (2000), 样子 *yàngzi* 'appearance' probably was borrowed from Japanese 様子 *yosu* 'appearance, state of things'. They are cognate and behave very similarly: when (的)样子 *(de)yàngzi* 'appearance' is attached to the end of a sentence, it expresses the speaker's judgment of a situation (8) like Japanese 様子 *yosu* 'appearance, state of things' does, which thus functions as a sentence final particle in TWM.

(8) a. 我喜欢你这套衣服，挺有品味的样子。(Diao 2000: 140)
　　　 wǒ xǐhuan nǐ zhè tào yīfu, tǐng yǒu pǐnwèi de yàngzi
　　　 'I like your clothes; it looks stylish.'
　　b. 用自己的牛车赶运趟别人的货件，三十块钱的样子。(Diao 2000: 140)

yòng zìjǐ de niúchē gǎnyùn tàng biéren de huòjiàn, sānshí kuài qián de yàngzi
'It earns about 30 dollars to deliver goods for others using our own oxcart.'

Local dialects
The influence of local dialects is most obvious in the pronunciation of Mandarin by speakers in Taiwan, Hong Kong, and Singapore. For instance, speakers of these three regions usually do not pronounce retroflex consonants such as *zh, ch, sh,* and *r* and seldom have the neutral tone (e.g., *jīngshén* for 精神 *jīngshen* 'energetic') because Chinese southern dialects usually do not possess these features. Also, due to the fact that the phoneme /h/ does not exist in Southern Min dialect, it often sounds like /fu/ by TWM and SM speakers. Furthermore, HKM speakers often do not distinguish the prepalatal (*j,q,x*), alveolar (*z,c,s*), and retroflex (*zh, ch, sh*) consonants, because of the lack of prepalatal and retroflex sounds in Cantonese.

All Mandarin varieties have borrowed a large amount of words from the local dominant dialects, e.g., TWM words from Southern Min and Hakka (9), and SM words from Southern Min and Cantonese (10). It is also not surprising that TWM and SM share words with Southern Min origin, e.g., 头家 *tóujiā* 'boss', 奥客 *àokè* 'unwelcomed guest', 烧 *shāo* 'hot', and 猫 *māo* 'stingy'.

(9) TWM words borrowed from Chinese dialects (Tang1999)
   a. from Southern Min: 头路 *tóulù* 'career', 白贼 *báizéi* 'to lie; liar', 法度 *fǎdù* 'idea', 黑白讲 *hēibáijiǎng* 'to talk carelessly', 活跳跳 *huótiàotiào* 'lively'
   b. from Hakka: 不四鬼 *bùsìguǐ* 'a shameless person', 讲古 *jiǎnggǔ* 'fable telling', 硬颈 *yìngjǐng* 'stubborn; indomitable'

(10) SM words borrowed from Chinese dialects
   a. from Southern Min：怕输 *pàshū* 'grasping', 三层肉 *sāncéngròu* 'pork belly', 角头 *jiǎotóu* 'corner', 散钱 *sǎnqián* 'changes', 家婆 *jiāpó* 'busy body', 做工 *zuògōng* 'work'
   b. From Cantonese: 大耳窿 *dà'ěrlóng* 'loan shark', 摆乌龙 *bǎiwūlóng* 'goof', 堂费 *tángfèi* 'litigation costs', 好彩 *hǎocǎi* 'lucky'

Grammatical variation is another product of dialect contact. For instance, both 才 *cái* 'just, only' and 再 *zài* 'then' can be used as temporal adverbs in MCM, but 才 *cái* usually emphasizes something that just happened or happens later than expected, whereas 再 *zài* is used to express that an event will be repeated, or an event will happen after another event or at a certain time. However, in SM, 再 *zài* is usually replaced by 才 *cái*, as in (11). According to Xing (2005), the mixed use of 才 *cái* and 再 *zài* is ascribable to Southern Min which does not distinguish 才 *cái* and 再 *zài*.

(11) SM: 走，吃饭！谢谢。我们吃了饭才来找你们。(*Home Décor Survivor-S5*-ep7)
       Zǒu, chīfàn! Xièxie. wǒmen chī le fàn cái lái zhǎo nǐmen
       'Go. Eat! Thank you. We will look for you after we eat.'

11

In HKM, because the Cantonese aspectual marker system does not fully corresponds to that in MCM, the use of aspectual markers shows several variations (Shi et al. 2014). Take the continuous marker 着 *zhe* as an instance. In HKM, 着 *zhe* is very often either omitted (12a), or replaced by verbal complements such as 上 *shàng* (12b) or other aspect markers such as 了 *le* (12c) (Shi et al. 2014).

(12) a. HKM: 加拿大有一面巨型美国旗，由三千多面小型美国旗和其他国家的国旗组成，象征(着)来自各国的 "911" 罹难者. (*Sing Tao Daily*- 1998-9-6/A22, cited from Shi et al. 2014: 326)
　　　jiānádà yǒu yī miàn jùxíng měiguóqí, yóu sānqiān duō miàn xiǎoxíng měiguóqí hé qítā guójiā de guóqí zǔchéng, xiàngzhēng (zhe) láizì gèguó de 911 línànzhě
　　　'There is a huge flag in Canada. It consists of more than three thousand flags from the United States and other countries, representing the 911 victims from each country.'

　　c. HKM: 直到一名穿上古怪警员制服的男子上台，那人才跳下台返回嘉宾席. (*Apple Daily* 1998-9-22/A21, cited from Shi et al. 2014: 328)
　　　zhídào yī míng chuān shang gǔguài jǐngyuán zhìfú de nánzǐ shàngtái, nā rén cái tiàoxià tái fǎnhuí jiābīnxí
　　　'Only till a man with a strange police uniform got up to the stage, did that person jump off the stage and return to the guest seat.'

　　c. 福泽谕吉是庆应大学的创办人，庆应大学的老校园内，树立了他的半身铜像 (*Ming Pao* 1998-9-1/D7, cited from Shi et al. 2014: 328)
　　　fúzéyùjí shì qìngyìng dàxué de chuàngbànrén, qìngyìng dàxué de lǎo xiàoyuán nèi, shùlì le tā de bànshēn tóngxiàng
　　　'Fukuzawa Yukichi is the founder of Keio University. His half-length statue was set up in the university's old campus.'

Given contacts to similar and related Sinitic languages in Taiwan, Hong Kong, and Singapore, the Mandarin varieties in these three regions share some grammatical features that are distinct from MCM. For instance, 有 *yǒu* 'have' mainly functions as a verb in MCM, but it can be used as an adverb in TWM, HKM, and SM, just like the corresponding 有 *yǒu* (lit.) 'have' in Southern Min and Cantonese, as in (13) from SM. In addition, the three Mandarin varieties share the comparative form marked by 过 *guò* (lit.) 'pass' (14), which is probably a feature of Classical Chinese that is retained in southern Chinese dialects, but rarely used in Northern varieties such as MCM.

(13) a. 我有输过吗？ (*Home Décor Survivor-S5*-ep2)
　　　wǒ yǒu shū guò ma
　　　'Did I ever lose?'
　　b. 你有做工啊？ (*Home Décor Survivor-S5*-ep2)
　　　nǐ yǒu zuògōng a
　　　'Do you work?'

(14) a. TWM: 音速可不可以快过光速?
　　　yīnsù kě bu kěyǐ kuàiguò guāngsù
　　　'Can sound speed travel faster than light speed?'
　　　(http://www.phy.ntnu.edu.tw/demolab/demolab/phpBB/viewtopic.php?topic=119
　　　37)
　　b. HKM: 内地二月份通胀 2.3% 高过市场预期
　　　nèidì èryuèfèn tōngzhàng gāoguò shìchǎng yùqī
　　　'The inflation rate in Mainland is 2.3%, which is higher than market expectation.'
　　　(http://www.metroradio.com.hk/news/default.aspx?NewsId=20160310093248)
　　c. SM: 他对朋友好过对老婆跟情人 (*Mars vs. Venus*-ep3)
　　　tā duì péngyou hǎoguò duì lǎopó gēn qíngrén
　　　'He treats his friends better than his wife and lover.'

The interaction of internal and external factors

The discussion above showed that the (dis-)similarities among different Chinese varieties can be predicted according to whether a particular change is internally or externally motivated. While all Mandarin varieties are expected to share internally motivated changes because of the same linguistic system, different outcomes arise if the internal factors work differently, e.g., in terms of scope, speed, and degree, in the different varieties of Mandarin. On the other hand, externally motivated changes such as contact-induced changes depend on the local language environment of each variety and are expected to vary. For instance, English has been a dominant language in Singapore and Hong Kong, so SM and HKM share features that are derived from English, e.g., the use of 和 *hé* 'and' and 相信 *xiāngxìn* 'believe'. Meanwhile, some features common to SM and TWM are traceable to Southern Min that has been a dominant local language in Singapore and Taiwan. Furthermore, SM is also special in possessing vocabularies borrowed from Malay and Tamil.

It should be noted that sometimes the two types of factors interact with each other to motivate a variation and no simplistic division can be drawn to distinguish their role in the process (Hickey 2012). One example is the progressive construction "VP+中 *zhōng*" in Mandarin varieties such as MCM, TWM, and HKM. From a language internal perspective, Yao (1997) argues that "VP+中" is a shortened form of the progressive construction "在 *zài* + VP + 中 *zhōng*", and Zhang (2002) proposes that 中 *zhōng* 'middle' underwent an extension from a spatial localizer to a temporal localizer, and finally to a verbal aspectual marker as a metaphor. However, based on a cross-variety comparison, Diao (2013) finds that "VP+中 *zhōng*" first appeared in TWM and is found more frequently used in TWM than in the other varieties. Therefore, Diao (2013) proposes that, in addition to the internal factors proposed by previous studies, it is possible that contact with Japanese gave rise to "VP+中 *zhōng*" in TWM, which was then spread to the other Mandarin varieties. In other words, both internal and external factors may have contributed to the emergence of "VP+中 *zhōng*" in Mandarin varieties, but it is unclear how the two factors interact.

The second example is the VP 做一个 X 的动作 *zuòyīgè X de dòngzuò* 'do an action of X' that has become increasingly popular in TWM. The construction was influenced by western

languages and is a bad usage for being redundant according to Taiwanese media (Her et al. 2016). Some frequently used examples include 做一个刷卡的动作 *zuòyīgèshuākǎ de dòngzuò* 'do an action of swiping card', 做一个点菜的动作 *zuòyīgèdiǎncài de dòngzuò* 'do an action of ordering', and 做一个了解的动作 *zuòyīgèliǎojiě de dòngzuò* 'do an action of understanding', despite the existence of the simple and short forms such as 刷卡 *shuākǎ* 'swipe card', 点菜 *diǎncài* 'order food', and 了解 *liǎojiě* 'understand'. However, according to Her (2016), the VP 做一个 X 的动作 *zuòyīgè X de dòngzuò* is not only grammatical but also properly used if given the right context. For example, (15) is fully acceptable when a dance instructor asks the students to do a turn around. This is likely due to the fact that the focus is on the execution of the turn-around and on how to improve it (i.e. on 'how to do' the event rather than the event itself).

(15) 你再做一个转身的动作给我看 (Her 2016: 46)
    nǐ zài zuò yī gè zhuǎnshēn de dòngzuò gěi wǒ kàn
    'Please show me another action of turn around.'

Xiong and Huang (2015) argue that the example of 做一个 X 的动作 *zuòyīgè X de dòngzuò* in TWM actually can instantiate the mechanism of *de*-verbalization. In this case, the addition of the light noun 动作 *dòngzuò* 'action' to a verb or a VP, with the possible assistance of 的 *de*, converts a verbal category into a nominal one. In fact, a similar usage has also been attested in MCM (e.g., 扑萤的动作 *pūyíng de dòngzuò* 'action of catching fireflies'. But *dòngzuò*-induced *de*-verbalization actually differs between MCM and TWM in the sense that the addition of 动作 *dòngzuò* in TWM is applicable to various kinds of actions (both concrete and abstract), but it is often restricted to bodily actions in MCM (Xiong and Huang 2015). It is noteworthy that beside 动作 *dòngzuò*, other light nouns such as 行为 *xíngwéi* 'behavior' (e.g. 财产让渡行为 *cáichǎn ràngdùxíngwéi* 'the behavior of transferring properties') and 步骤 *bùzhòu* 'procedure' (e.g. 一个过滤的步骤 *yīgèguòlù de bùzhòu* 'the procedure of filtering') can have the similar usages. In other words, what seemed to be an anomaly in variations in language use may in fact be a change in process, with strong internal motivation (*de*-deverbalization) and speakers' need to underline different aspects of meaning.

Summary
This chapter presents an overview of the current study on World Chinese variations and the factors contributing to the variation, with examples mainly from MCM, TWM, HKM, and SM. Each variety of Mandarin is active and changing, not only because of the linguistic features of Mandarin, but also due to external factors, especially language contact in the competition between the global language (i.e. English), local languages and dialects along with the increasing globalization and the rise of China. We also introduced comparable corpus driven, statistical model based computational approaches as a unique opportunity for the discovery of new linguistic facts and the empirical proof of theories of language variations. In addition, the language variations could also be understood as the reflection

of individual differences. Then one question to ask is how such language variations could be seen as relating to social, psychological or biological parameters (Fillmore et al. 1979). Hence it would also be meaningful to reveal what such variations can mean in the context of language evolution, psychological as well as socio-linguistic patterns. It should also be noted that the idea of 'individual variation' is also related to the topics we discussed in some other chapters (e.g. Chapters on aphasia, communication disorder, amusics, as well as heritage language and language of new immigrants).

References

Chen, Chung-Yu 陈重瑜. 1986. Grammatical features of Singapore Chinese 新加坡华语语法特征. *Language Studies 语言研究* 1: 138–152.

Chew, Cheng Hai 周清海. 2007. The language policy of Chinese in globalization 论全球化环境下华语的规范问题. *Language Teaching and Study 语言教学与研究* 4: 91–96.

Chinese Daily Publication Center 国语日报出版中心. 2000. *New Chinese Daily dictionary 新编国语日报词典*. Taipei: Chinese Daily.

Diao, Yanbin 刁晏斌. 2000. *Divergence and convergence: a cross-strait language comparison 差异与融合: 海峡两岸语言应用对比*. Nanchang: Jiangxi Education Press.

Diao, Yanbin 刁晏斌. 2013. An investigation of "(zai+) VP + zhong in two shores and four places 两岸四地"（在＋）VP+中"形式考察. In *New achievements of the comparison studies in two shores and four places 两岸四地现代汉语对比研究新收获*, ed. Yanbin Diao 刁晏斌, XXX–XXX. Beijing: Language and Culture Press.

Fillmore, Charles J., Daniel Kempler, and William S-Y Wang (Eds). 1979. *Individual differences in language ability and language behavior*. New York: Academic Press.

Her, One-Soon 何万顺. 2016. Does the language get cancer? 语言癌不癌? In *Does the language get cancer? A perspective from linguists 语言癌不癌? 语言学家的看法*, ed. One-Soon Her, Wei-Tien Tsai, Jung-Hsing Chang, Kawei Chui, Jennifer M. Wei, and Victoria Rau 何万顺, 蔡维天, 张荣兴, 徐佳慧, 魏美瑶, 何德华, 37–52. Taipei: Linking Publishing.

Her, One-Soon, Wei-Tien Tsai, Jung-Hsing Chang, Kawei Chui, Jennifer M. Wei and Victoria Rau 何万顺, 蔡维天, 张荣兴, 徐佳慧, 魏美瑶, 何德华. 2016. *Does the language get cancer? A perspective from linguists 语言癌不癌? 语言学家的看法*. Taipei: Linking Publishing.

Hickey, Raymond. 2012. Internally and externally motivated language change. In *The Handbook of Historical Sociolinguistics*, ed. Juan Manuel Hernandez-Compoy and Juan Camilo Conde-Silvestre, 401–421. Malden: Wiley-Blackwell.

Huang, Chu-Ren, Petr Simon, and Shu-Kai Hsieh. 2007. Automatic Discovery of Named Entity Variants. Paper presented at P*roceedings of the Association of Computational Linguistics Annual Meeting,* Prague-Czech.

Huang, Chu-Ren. 2009. Tagged Chinese Gigaword Version 2.0. Philadelphia: Lexical Data Consortium, University of Pennsylvania.

Huang, Chu-Ren, Jingxia Lin, Menghan Jiang and Hongzhi Xu. 2014. Corpus-based Study and Identification of Mandarin Chinese Light Verb Variations. *Proceedings of the*

First Workshop on Applying NLP Tools to Similar Languages, Varieties and Dialects. 1–10.

Huang, Chu-ren, Jingxia Lin, and Huarui Zhang. 2012. World Chineses based on comparable corpus: The case of grammatical variations jinxing. *Proceedings of Macau studies on language and culture* 澳门语言文化研究, 397–414.

Institute of Linguistics, Chinese Academy of Social Sciences 中国社会科学院语言研究所. 2005. *Contemporary Chinese dictionary* 现代汉语词典. Beijing: Commercial Press.

Jiang, Menghan, Jingxia Lin and Chu-Ren Huang. 2015. A comparable corpus-based study of VO compound variations between Mainland and Taiwan Mandarin. *The 23rd annual conference of the International Association of Chinese Linguistics (IACL)*. Hanyang University, Seoul, Korea.

Jiang Menghan, Dingxu Shi and Chu-Ren Huang. 2016. Transitivity in light verb variations in Mandarin Chinese- A comparable corpus-based statistical approach. Paper presented at *Proceedings of the 30th Pacific Asia Conference on language, information and computing.* Seoul, Korea.

Kachru, Braj B. 1985. Standards, codification, and sociolinguistic realism: The English language in the outer circle. In *English in the World: Teaching and Learning the Language and the Literature*, ed. Randolph Quirk and Henry G. Widdowson, 11–30. Cambridge: Cambridge University Press.

Lau, KakEn. 1993. *Singapore Census of Population 1990: Literacy, Languages Spoken and Education.* Singapore: Department of Statistics.

Lin, Jingxia, Hongzhi Xu, Menghan Jiang, and Chu-Ren Huang. 2014. Annotation and Classification of Light Verbs and Light Verb Variations in Mandarin Chinese. *The Workshop on Lexical and Grammatical Resources for Language Processing*, 75–82.

Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.

Lu, Jianming 陆俭明. 2002. The grammatical features of Singapore Chinese 新加坡华语语法的特点. In *Lexicon and grammar of Singapore Chinese 新加坡华语词汇与语法*, ed. Chew Cheng Hai 周清海. Singapore: Lingzi Media.

Lü, Shuxiang 吕叔湘. 1999. *Eight hundred words in Modern Chinese现代汉语八百词.* Beijing: Commercial Press.

Shi, Dingxu, Jingmin Shao, and Zhiyu Zhu 石定栩, 邵敬敏, 朱志瑜. 2006. *A comparison of Hong Kong Chinese and Standard Chinese港式中文与标准中文的比较.* Hong Kong: Hong Kong Educational Publishing Company.

Shi, Dingxu, Jingmin Shao, and Zhiyu Zhu 石定栩, 邵敬敏, 朱志瑜. 2014. *A comparison of Hong Kong Chinese and Standard Chinese (2nd edition)港式中文与标准中文的比较(第二版)*. Hong Kong: Hong Kong Educational Publishing Company.

Šimon, Petr, Chu-Ren Huang, Shu-Kai Hsieh, and Jia-Fei Hong. 2008. Transliterated named entity recognition based on Chinese word sketch. *International Journal of Computer Processing of Languages* 21(1): 19–30.

Su, Xinchun 苏新春. 2001. The assimilation and renewal of dialectal words in Putonghua vocabulary: Study of dialectal words in Contemporary Chinese Dictionary 普通话词汇系统对方言词的吸收与更新———〈现汉〉方言词研究. *Language 语言*2 .

T'sou, benjamin, and Rujie You 邹嘉彦, 游汝杰. 2010. *Global Chinese new words dictionary*全球华语新词语词典. Beijing: Commercial Press.

T'sou, Benjamin, and Oi Yee Kwong. 2006. Toward a Pan-Chinese thesaurus. *The Fifth International Conference on Language Resources and Evaluation (LREC 2006)*, 2391–2394.

Tang, Zhixiang 汤志祥. 1995. Examples of lexical difference in Mainland China, Taiwan, Hong Kong, Singapore Chinese中国大陆, 台湾, 香港，新加坡汉语词汇方面若干差异举例. *Journal of Xuzhou Normal College 徐州师范学院学报* 1:103–109.

Tim, Siu-lam, and Paul M. Ma 田小琳, 马毛朋. 2013. A brief study on the word order of Hong Kong Chinese 港式中文语序问题略论. *Chinese Linguistics汉语学报* 41(1): 69–75.

Tsao, Feng-fu 曹逢甫. 1999. The language spoken by the four major ethnic groups in Taiwan: their history and present state 台湾语言的历史及其目前的状态与地位. *Chinese Studies汉学研究* 17(2): 313–343.

Tseng, Hsin-I 曾心怡. 2003. *The Syntax Structures of Contemporary Taiwanese Mandarin 当代台湾国语的句法结构*. MA thesis, National Taiwan Normal University, Taipei.

Wang, Huidi 汪惠迪. 1999. Unique words in Singapore Mandarin: A social portrait of Singapore 华语特有词语：新加坡社会写真. *Journal of Yangzhou University (Humanities and Social Sciences)扬州大学学报人文社会科学版* 4: 44–51.

Weinreich, Uriel, William Labov, and Marvin I. Herzog. 1968. Empirical foundations for a theory of language change. In *Directions for Historical Linguistics: A Symposium*, ed. Winfred P. Lehmann and Yakov Malkiel, 95–188. Austin: University of Texas Press.

Xing, Fuyi 邢福义. 2005. Potential influences of the mother tongue dialect in using Chinese language in Singapore 新加坡华语使用中原方言的潜性影响. *Dialect方言* 2: 175–181.

Xing, Fuyi 邢福义. To appear. *Global Chinese Grammar全球华语语法研究*. Beijing: Commercial Press.

Xiong, Jiajuan, and Chu-Ren Huang. 2015. De-verbalization and Nominal Categories in Mandarin Chinese: A corpus-driven study in both Mainland Mandarin and Taiwan Mandarin. Paper presented at *The 29th Pacific Asia Conference on Language, Information and Computation*. Shanghai.

Yao, Te Hwai 姚德怀. 1997. The use of *yingye-zhong* and verb+*zhong* "营业中"和"动词＋中"的用法. *Chinese Language Review(Hong Kong) 语文建设通讯(香港)* 54: 50–52.

Zhang, Lan 张岚. 2009. *A Study on the Difference of Pronunciation Between Mainland China and Taiwan海峡两岸现代汉语通用语语音差异对比研究*. MA thesis, East China Normal University, Shanghai.

Zhang, Yisheng 张谊生. 2002. The function of VP-zhong and the grammaticalization of zhong "VP中"的功能特征及"中"的虚化历程. *In Grammar research and exploration (11) 语法研究和探索(11)*. Beijing: Commercial Press.

Zhao, Chunli, and Dingxu Shi 赵春利, 石定栩. 2012. Types and features of the Chinese comparative sentences used in Hong Kong 港式中文差比句的类型与特点. *Journal*

of Yunnan Normal University (philosophy and social sciences) 云南师范大学学报 *(哲学社会科学版)* 44(6): 8–16.

Endnotes

[1]Traditionally called dialects, and they are also commonly called Sinitic languages by typologists. The term Sinitic languages is adopted here in the context of the discussion of language contact. Please see chapter 13 of this volume for a more detailed discussion on Sinitic languages/Chinese dialects.
[2]Some consequences of such contacts are discussed in the chapter on neologism (Chapter 34 of this volume).
[3]Data from National Statistics (Taiwan), URL: https://www.stat.gov.tw/mp.asp?mp=4
[4]Data from Lau (1993) and *Statistics Singapore* (URL: http://www.singstat.gov.sg/)
[5]Data from *Statistics Singapore* (URL: http://www.singstat.gov.sg/).