



Research Article

Sha Liu*

Factors in sound change: A quantitative analysis of palatalization in Northern Mandarin

<https://doi.org/10.1515/opli-2022-0226>

received March 18, 2021; accepted November 21, 2022

Abstract: Factors in sound change are still a major subject of debate in the field of linguistics, with the frequency factor perhaps being the most controversial. The present article focuses on palatalization of the velars before high front vowels and glides in Northern Mandarin, because palatalization stretched for more than 100 years and can provide detailed information concerning its contour. Based on a statistical analysis of corpus data of palatalization in Northern Mandarin, the present article argues that the factor of frequency is positively associated with palatalization and the factor of frequency change is negatively associated. Morphosyntactic structure and word class are also claimed to be factors in sound change. However, neither of these two factors has emerged as significant in the current study.

Keywords: frequency, frequency change, morphosyntactic structure, palatalization, sound change, word class

1 Introduction

Factors in sound change are not a fresh topic in the linguistic field. However, it is still a matter of major debate since different theories hold contradictory views. Frequency, morphosyntactic structure, and word class are the three factors often referred to in discussion of sound change (Thomsen 1879, Postal 1968, Phillips 1983, 2001, Bybee 1985, 2012, 2017, Crowley 1997, Pierrehumbert 2001, Donohue 2005, Ogura 2012, Smith 2012). The present study is motivated by the following two points: (1) the long debate concerning the role of the aforementioned three factors in sound change is yet to come to a conclusion; (2) frequency change, a potential factor in sound change, has generally been ignored in previous studies. This article not only intends to find an answer to the question of whether the three factors have a role in sound change but also intends to understand their exact role in sound change if the answer to the first question is affirmative. The present article focuses on palatalization in Middle Mandarin (c.1500–c.1800). Palatalization began at the beginning of the eighteenth century and was completed no later than the beginning of the nineteenth century (Morrison 1815–23, Edkins 1857, 1864, Akiyasu 1967). The whole process was finished in approximately one century. Therefore, it is possible to explore the different phases of this process. In addition, this process involved a relatively large number of morphemes, which makes the statistical analysis possible. Third, previous studies have mainly used data from the Indo-European language family. The sound change mechanism is supposed to be universal, and thus, data from various language families should be investigated. From this respect, Mandarin is a good candidate since it belongs to the Sino-Tibetan language

* **Corresponding author: Sha Liu**, Language Education and Research Center, Fukuoka University, Jonan Ward, Fukuoka City, Fukuoka Prefecture, Japan, e-mail: pengpeng_happy1@yahoo.co.jp

family and has not been widely discussed. This article uses related classic works and corpus to collect data for palatalization, carries out statistical analysis, and reveals factors in this sound change.

This article starts with a sketch of factors in sound change documented in previous studies and then presents the trajectory of palatalization in Middle Mandarin (c.1500–c.1800) and data extraction methods in Section 3. Section 4 focuses on the statistical analysis of corpus data and presents my own proposals concerning factors in sound change. This article arrives at its conclusion in Section 5.

2 Review of previous studies

A growing body of literature has been arguing that an intersection of factors should be considered for sound change discussion (see, e.g., Grimm 1822, Jespersen [1922] 2007, Malkiel 1968, 1969–70, 1976, Postal 1968, Stemberger and MacWhinney 1986, Crowley 1997, Kapatsinski 2010, Wedel et al. 2013). Factors in sound change documented in previous studies include frequency, morphosyntactic structure, and word class. For ease of exposition, this section starts with the word class factor and ends with the most complex factor, frequency.

2.1 Word class

In the second half of the twentieth century, a number of studies demonstrated the possible relation between word class and sound change. To exemplify, King (1969) discusses the final schwa deletion in Yiddish, notices that this rule does not apply when the final schwa is in an adjective inflectional ending, and concludes that this is evidence of morphologically conditioned sound change (reviewed in Robinson and van Coetsem 1973; see also Anttila 1972). More recently, Crowley (1997) focuses on Southern Paamese and Northern Paamese, two languages of Central Vanuatu, and reports that Southern Paamese /l/ corresponds to Northern Paamese /i/, /l/, or zero in all word classes except verbs. According to Crowley (1997), this process shows that at least some sound changes only affect certain word classes. Although both King (1969) and Crowley (1997) claim that word class has a role in sound change, their claim is to some extent vague. Neither of them states clearly the exact relationship between word class and sound change. Phillips (1983, 2006) is also a proponent of the role of word class in sound changes and gives a much more detailed description of its exact role. Phillips (1983, 2006) classifies words into two categories, content words and function words: content words mainly include adjectives, adverbs, nouns, verbs, etc.; function words refer to a wide range of words that normally receive low sentence stress, such as adverbial conjunctions, auxiliary verbs, determiners, prepositions, quantifiers, and so on. Phillips (1983, 2006) points out that a strengthening sound change, the change from /d/ to /t/ in Old High German Isidor, affected function words last, and goes further to state that content words tend to be affected by strengthening sound changes first and function words by weakening sound changes first.

Though intriguing, the role of word class in sound change has been facing challenges (Jasanoff 1971, Blevins and Lynch 2009, Brown 2013). Blevins and Lynch (2009) claim that the sound change discussed in Crowley (1997) applies to all word classes including verbs, but the phonological and morphological components in verb inflections restore the sound change in verbs and make verbs appear to be exceptions. From another perspective, Jasanoff (1971) claims that what appears to be morphologically conditioned is in fact regular sound change partially obscured by subsequent analogy.

At the same time, some scholars adopt the middle way by claiming that no conclusion can be drawn yet and further investigation is necessary (Sihler 2000, Campbell 2013, Manker 2015). For example, although Manker (2015) states that the reason why many examples in Phillips (2006) appear to be limited to certain word classes is that these words more often happen to be utilized within the favorable phonetic environment for the change, Manker (2015) does not give any decisive conclusion, stating that the role of morphological factors in sound changes cannot be absolutely ruled out and needs further extensive investigation (see also Sihler 2000, Campbell 2013).

2.2 Morphosyntactic structure

The second factor to be reviewed is the morphosyntactic structure. Taking the credit for noticing this factor, Postal (1968) focuses on Mohawk, the language spoken by Mohawk people, and notices that [e] is regularly inserted into [kw] sequences except when the [k] is the first person morpheme and the [w] is the first element of the plural morpheme. According to Postal (1968), this language example shows that superficial morphosyntactic structure could also condition sound change (reviewed in Fudge 1972; see also King 1969). Later scholars more plainly state the morphosyntactic structure as the division between bound morphemes and free morphemes. Donohue (2005) discusses the voicing of voiceless stops in Palu'e and sound changes in Modern Indonesian and Bali-Vitu (Austronesian, Oceanic) and points out that the sound change advances in a small set of bound morphemes more completely than in free lexemes. Similarly, Bybee (2002) focuses on word-final /t, d/ deletion in American English and concludes that bound morphemes can affect the deletion process. Guy (1991b) also focuses on word-final /t, d/ deletion in American English and gives a more detailed conclusion: underived or monomorphemic words, such as *mist* and *pact*, undergo deletion at a higher rate than inflected forms such as past tense verbs like *missed* and *packed* (see also Labov et al. 1968, Fasold 1972, Guy 1991a). Baranowski and Turton (2020) report a similar result for word-final /t, d/ deletion in British English. Different scholars hold different views concerning the exact effect of the morphosyntactic structure: although Guy (1991a, 1991b) claims that underived or monomorphemic word-lead /t, d/ deletion in American English, Donohue (2005) reports that bound morphemes were at the forefront of the voicing of voiceless stops in Palu'e.

In contrast, Renwick et al. (2014) focus on word-final /t, d/ deletion in British English and report that their results show little support for the role of morphosyntactic structure. What is most surprising may be that completely opposite conclusions concerning the role of morphosyntactic structure have been drawn from the same phenomenon, word-final /t, d/ deletion in English.

To sum up, controversies concerning the morphosyntactic structure factor are first whether it has a role in sound change, and second, whether bound morphemes or free morphemes change faster in sound change.

2.3 Frequency

Word class and morphosyntactic structure are not the only two factors put forward in previous studies. The frequency factor is possibly the one that has generated the most discussion. There has been a long-running argument between academics over whether the frequency is relevant. Even scholars who argue that the frequency plays a role in sound change cannot agree with each other on the exact frequency effect. In the following, the frequency effect is approached from the perspectives of high frequency and low frequency and relative frequency.

2.3.1 High frequency vs low frequency

Since the nineteenth century, the frequency effect has been studied in detail. Grimm (1822) discusses the connection between irregularity and high-frequency auxiliary verbs. Thomsen (1879) gives a few frequent Romanic verbs, pointing out that they are exceptions to normal phonetic development (see also Thomsen 1920). Jespersen ([1922] 2007) more plainly states that frequent words, due to their common and recurrent usage, tend to be exceptions to sound change.

What is interesting is that more than a century has passed, and yet the frequency effect is still not agreed on or well understood. Bybee (1985, 2000, 2002), Pierrehumbert (2001), and Smith (2012), among others, claim that frequent words lead sound change and less frequent words follow, a view different from that of Grimm (1822), Thomsen (1879, 1920), and Jespersen ([1922] 2007). Bybee (2002) studies the deletion of final /t/ and /d/ in American English and finds that the deletion rates in high-frequency words are

statistically higher than in low-frequency words. According to Bybee (2002), this is evidence that frequent words lead sound change. Hay and Foulkes (2016) report that frequent words lead a change in the pronunciation of word-medial intervocalic /t/ in New Zealand English. On the contrary, other scholars claim that low-frequency words are at the forefront of changes and high-frequency words follow (Labov 1989, Bermúdez-Otero 2007, Hay et al. 2015). Labov (1989) focuses on Philadelphia a-tensing and reports that the most frequent words lag behind in the shifting to the tense class. Hay et al. (2015) observe pronunciation changes in New Zealand English over a 130-year period and claim that their data show that low-frequency words led these changes and high-frequency words lagged behind.

Unlike in the aforementioned two paragraphs, scholars referred to in this paragraph doubt the role of frequency in sound change. Fruehwald et al. (2013) assert that the phonological change progresses in every context at the same rate over time irrelevant of frequency and name this as the constant rate effect (see also Kroch 1989, 1994, Pintzuk 1991, Santorini 1992, 1993, Dinkin 2008). In addition, Zellou and Tamminga (2014) study the co-articulatory vowel nasality in Philadelphia English and claim that their study shows that the changes in nasality are independent of any frequency effect. Similarly, Labov (2010) examines the role of frequency in several different phonetically gradual changes and concludes that the role of frequency in these changes is minimal, if not zero, a view agreed on by Kiparsky (2014).

2.3.2 Relative frequency

Wedel et al. (2013) approach the frequency effect from a different perspective, proposing that the relative frequency, instead of absolute frequency, is the factor relevant to sound changes. Wedel et al. (2013) use 19 systems of phoneme contrasts from nine different languages to explore factors in phoneme merger, such as the *cot* ~ *caught* merger in North American English, and report that no measures based on absolute word frequency could satisfactorily predict merger probability for their data. Instead, Wedel et al. (2013) report that the frequency ratio or the relative frequency of minimal pair members could predict merger well, with minimal pairs whose members are more similar in frequency having less tendency to merge.

2.4 Proposal of this article

As reviewed in Section 2.1, the role of word class in sound change is yet to be determined and is thus a focus of this study. Similarly, the role of morphosyntactic structure in sound change is not fully understood either. As noted in Section 2.2, first, this article needs to answer whether the morphosyntactic structure has a role in sound change; second, this article needs to find out whether bound morphemes or free morphemes lead sound change, if the answer to the first question is affirmative.

Concerning the most debated factor, frequency, this article explores whether it has a role in sound change. This article extracts numerical data of frequency and investigates its correlation with sound change by statistical analysis. In addition, the present study refers to the concept of relative frequency. Unlike Wedel et al. (2013), the present study does not refer to minimal pair members and thus cannot give relative frequency from the perspective of minimal pair. Instead, the study compares frequencies of all related morphemes and assigns them to one of the three levels, low, medium, and high.

The present study also has its own proposal: the factor of frequency change. Previous studies have mainly focused on the frequency effect and have not referred so much to frequency change. If the frequency is correlated with sound change, however, frequency change may also be associated with sound change. To exemplify, Pierrehumbert (2001), among others, claims that sound change usually affects the most frequent lexical items first. Assuming this to be correct, lexical items with increased frequency during the time concerned seem more likely to undergo a certain sound change earlier than lexical items with decreased frequency. The reason is that lexical items with increased frequency seem to be more active and more accessible to a related sound change than lexical items with decreased frequency during the time concerned.

The present study also refers to the concept of relative frequency change by comparing the frequency of a related morpheme in the time of palatalization with its frequency in an earlier period to calculate the frequency change of this morpheme and assigning it to one of two levels, decreased frequency and increased frequency.

This section has given factors referred to in previous studies and my own proposals. The controversies concerning these factors can be solved only by a close look at empirical data. Section 3 presents a brief description of the language data for the present study.

3 Data

This article takes palatalization in Northern Mandarin (AD 1324–Present) as its language sample. Palatalization in Northern Mandarin stretched from the beginning of the eighteenth century until the beginning of the nineteenth century and was completed without exception (Morrison 1815–23, Edkins 1857, 1864, Wang 1957, Akiyasu 1967, Chen 1976, Pulleyblank 1984). It can provide detailed information concerning morphemes that were at the forefront of this sound change, its contour, and so on. Section 3.1 briefly describes this process, and Section 3.2 expounds data extraction from *CNCORPUS* (Xiao 2010) and *Piaotongshi xinshi yanjie* (New Edition of Vernacular Exposition of Pak the Interpreter; Kim [1765] 2005).

3.1 Trajectory of palatalization in Northern Mandarin

Palatalization concerns the velars ([k], [k^h], and [x]): the velars were palatalized before high front vowels and glides to [tɕ], [tɕ^h], and [ç] (Pulleyblank 1984; see also Chen 1976, Li 1999, Handel 2017). The velar palatalization rule is given in (1) according to Chen (1976), Pulleyblank (1984), Li (1999), and Handel (2017).¹

(1) Velar palatalization rule

$$\begin{aligned} k &\rightarrow t\text{ɕ} / ______ i, j \\ k^h &\rightarrow t\text{ɕ}^h / ______ i, j \\ x &\rightarrow \text{ç} / ______ i, j \end{aligned}$$

Palatalization illustrated in (1) is a regular sound change since it does not have exceptions: all morphemes that met the environment have been palatalized (Wang 1957, Pulleyblank 1984). After the completion of palatalization, the alveolopalatal triplexes ([tɕ], [tɕ^h], [ç]) are in complementary distribution with the velar triplexes ([k], [k^h], and [x]) in Modern Standard Chinese (c.1900–Present), with the alveolopalatal triplexes before high front vowels and the velar triplexes in all other phonetic environments. Three syllables are exemplified in Table 1 to illustrate the velar palatalization rule in (1).

The transcriptions in Table 1 are given according to Baxter and Sagart (2014), with reference to *Zhongyuan Yinyun* (Rhymes of the Central Plain; Zhou [1324] 1996) and reconstructions of *Zhongyuan Yinyun* in Pulleyblank (1984, 1991) and Chou (1993). In Table 1, the phoneme [k] of the morpheme ‘urgent’ (急) was palatalized to [tɕ] during palatalization shown as the velar palatalization rule (1). In fact, the morpheme ‘urgent’ (急) not only underwent palatalization but also lost its final consonant [p]. In this way, [kip] became [tɕi] in Middle Mandarin (c.1500–c.1800). The deletion of consonants in the final position seems especially common in languages (Mortensen 2012). Mandarin is not an exception. The system of codas,

¹ Different scholars use slightly different symbols for the environment of the velar palatalization rule. Some scholars use [i] and [j], while others use [i] and [y]. This article follows Pulleyblank (1984) and indicates the environment uniformly as [i] and [j].

Table 1: Palatalization in Mandarin

Early Mandarin ² (c.1300–c.1500)	Middle Mandarin (c.1500–c.1800)	Gloss
[kip]	[tɕi]	‘urgent’ 急
[k ^h jt]	[tɕ ^h i]	‘beg, ask’ 乞
[xjeh]	[ɕi]	‘opera’ 戏

namely, final consonants, was dramatically reduced when Middle Chinese (AD 601–AD 1336) evolved into Mandarin (AD 1324–Present) (Wang 1957, 1985, Pulleyblank 1984, 1991). All final voiceless stop consonants, namely, [k], [p], and [t], disappeared altogether (Chen 1976, Pulleyblank 1984). As a result, syllables like [kip] and [k^hjt] in Table 1 finally became [tɕi] and [tɕ^hi], respectively. A detailed discussion of the loss of final consonants will be omitted here since it is not the focus of the present article. The morpheme ‘beg, ask’ (乞) was [k^hjt] in Early Mandarin (c.1300–c.1500). It not only underwent the palatalization but also lost its final consonant [t]. In this way, the morpheme ‘beg, ask’ (乞) finally became [tɕ^hi] in Middle Mandarin (c.1500–c.1800). The same process applies to the morpheme ‘opera’ (戏).

Palatalization was first recorded in the anonymous Chinese rhythm dictionary *Yuanyin Zhengkao* (On the alveolar and alveolo-palatal consonants; [1743] 2015). This dictionary contains transcriptions in the Manchu alphabet, a phonographic system. The dictionary gives the pronunciations of all related morphemes as [ki-], [k^hi-], and [xi-] since it takes the position that the palatalized pronunciations of [tɕi], [tɕ^hi], and [ɕi] are not correct and argues that the pronunciations of [ki-], [k^hi-], and [xi-] should be followed. The argument in this dictionary provides evidence that palatalization had begun before the middle of the eighteenth century, but it does not provide information concerning which morphemes had undergone palatalization. It was not until the rhyme dictionary *Lishi Yinjian* (The rhyme book of Liruzhen; Li [1805] 1992) that palatalization was shown to have been completed for the first time. In sum, the discussion in Section 3.1 presents two important points in time: the period before the middle of the eighteenth century as the beginning of palatalization and the beginning of the nineteenth century as the time for the completion of palatalization.

3.2 Data extraction

The most reliable source of data seems to be corpora with data in the related period and dictionaries and historical books compiled by native speakers of Mandarin during the time concerned.

3.2.1 Dictionaries and historical books

The choice of dictionaries and books during the time concerned is not as straightforward as it appears. As noted in Section 3.1, *Yuanyin Zhengkao* ([1743] 2015), the first Chinese rhyme dictionary that recorded palatalization, does not indicate which morphemes had been palatalized. It only states that palatalized pronunciations are not correct. In addition, dictionaries and books compiled by Chinese scholars before the twentieth century use *fanqie* (反切), a traditional method of indicating the pronunciation of a Chinese character by using two other Chinese characters. For example, the pronunciation of the character 唐 might be represented as follows: 徒郎. It roughly means that the initial of 唐 is the same as that of 徒, and the final

² The period of Mandarin (AD 1324–Present) is usually divided into three phases: Early Mandarin (c.1300–c.1500), Middle Mandarin (c.1500–c.1800), and Modern Standard Chinese (c.1900–Present) (Dong 1954, Stimson 1966, Shi 2002, Shen 2015, 2020). The classic book *Zhongyuan Yinyun* (Rhymes of the Central Plain; Zhou [1324] 1996) marks the beginning of Early Mandarin (c.1300–c.1500; Norman 1988, Li 1999, Shen 2015, 2020). The focus of this article, palatalization, falls in the period of Middle Mandarin (c.1500–c.1800).

of 唐 is the same as that of 郎. This representation makes it circular and thus difficult to understand the pronunciations of characters and morphemes they represent, since the Chinese writing system is a representative logographic system, not a phonographic system like English. In contrast, Chinese language learning textbooks used in Korea use Hangul (the Korean alphabet). They can provide a relatively clear picture about the pronunciation of Mandarin during the time concerned.

Piaotongshi Xinshi Yanjie (New Edition of Vernacular Exposition of Pak the Interpreter; Kim [1765] 2005) is the first Chinese language learning textbook used in Korea to show palatalized pronunciations in Mandarin (Kim 1989, Jin 2001, Lin 2010). Chinese characters in this textbook are transcribed in Hangul (the Korean alphabet). The book was compiled in Beijing, the capital of China, and focuses on oral Mandarin at that time (Kim 1989, Lin 2010). As stated in Section 3.1, palatalization was first recorded in the anonymous Chinese rhythm dictionary *Yuanyin Zhengkao* ([1743] 2015). *Piaotongshi xinshi yanjie* (Kim [1765] 2005) was compiled almost in the same period as the rhyme dictionary. Thus, it can provide information about morphemes that were at the forefront of palatalization. This study first relies on *Piaotongshi xinshi yanjie* (Kim [1765] 2005) to locate morphemes that had been palatalized in the middle of the eighteenth century and those that had not. Take the two morphemes ‘urgent’ (急) and ‘opera’ (戏) in Table 1 as examples. The morpheme ‘urgent’ is transcribed as ㄹ in *Piaotongshi xinshi yanjie*, or [tɕi] in the International Phonetic Alphabet. Therefore, this morpheme had been palatalized in the middle of the eighteenth century and was at the forefront of palatalization according to *Piaotongshi xinshi yanjie*. The morpheme ‘opera’ is transcribed as ㅁ ([xi]), which means that the morpheme had not been palatalized in the time of the book. Thus, the morpheme ‘opera’ lagged behind in the palatalization. In this way, morphemes leading the palatalization and morphemes that lagged behind can be identified. The next step is to compare morphemes leading the palatalization with morphemes lagging behind in this process to locate factors involved in the palatalization.

3.2.2 Corpus

As noted, the three factors, word class, morphosyntactic structure, and frequency, are of interest here. This study uses *CNCORPUS* to extract frequency data. *CNCORPUS* is composed of two databases: an Old Chinese and Middle Chinese database and a Mandarin database. It also permits searching for data according to Chinese dynasties. Another benefit of the corpus is its capacity, 120 million Chinese characters from a wide range of sources. However, *CNCORPUS* does not give phonetic transcriptions of these Chinese characters. Therefore, the exact phonetic transcription of the related morphemes for this article relies on the book *Piaotongshi xinshi yanjie*.

CNCORPUS permits searching for data according to Chinese dynasties. The book *Piaotongshi xinshi yanjie* (Kim [1765] 2005) was compiled in the middle of the eighteenth century. Palatalization ended at the beginning of the nineteenth century (Li [1805] 1992, Morrison 1815–23, Edkins 1857, Edkins 1864, Akiyasu 1967). The two important time points – the middle of the eighteenth century and the beginning of the nineteenth century – fall within the Qing Dynasty (1644–1912). Accordingly, the focus was given to the frequency data in the Qing Dynasty (1644–1912). In addition, because frequency change was assumed to be a factor in sound change, frequency before the time concerned was required to calculate frequency change. The Qing Dynasty (1644–1912) was preceded by the Ming Dynasty (1368–1644). However, data for the Yuan Dynasty (1271–1368) and data for the Ming Dynasty (1368–1644) are combined together in the *CNCORPUS*. Thus, data from both the Yuan and Ming Dynasties (1271–1644; hereafter the YM Dynasties) were extracted. In other words, the focus was mainly on the frequency data in the Qing Dynasty (1644–1912), and the frequency change from the YM Dynasties (1271–1644) to the Qing Dynasty (1644–1912). Take the morpheme ‘urgent’ in Table 1 as an example again: its frequency is 3,592 (1644–1912) and 4,684 (1271–1644). Its frequency change is thus –1,092, which means that its frequency has decreased from 1271–1644 to 1644–1912. Concerning the morpheme ‘opera’, its frequency is 1,072 (1644–1912). Since its frequency is 1,045 between 1271 and 1644, it has an increased frequency in 1644–1912 compared to its frequency in 1271–1644.

3.3 Methodology

This study first relies on *Piaotongshi xinshi yanjie* (Kim [1765] 2005) to locate morphemes that had been palatalized in the middle of the eighteenth century and those that had not. Then the article uses the CNCORPUS to look for related information concerning both palatalized morphemes in the middle of the eighteenth century and unpalatalized morphemes in the same period. Finally, the article uses GraphPad Prism version 8.0.0 for Windows (GraphPad 2019; henceforth GraphPad software) to carry out statistical analysis to locate factors in palatalization.

According to my manual research result shown in Table 2, 70.21% of morphemes of [ki-] in Early Mandarin (c.1300–c.1500) have palatalized pronunciations in *Piaotongshi xinshi yanjie* (Kim [1765] 2005), 33 morphemes of 47; 66.67% of morphemes of [k^{hi}-] in Early Mandarin have palatalized pronunciations, 20 morphemes of 30. A total of 51.61% of [xi-] in Early Mandarin have palatalized pronunciations, 16 morphemes of 31. The whole palatalization rate in *Piaotongshi xinshi yanjie* (Kim [1765] 2005) is 63.89%. Palatalization had been completed in *Lishi Yinjian* (Li [1805] 1992). The trajectory of palatalization is shown in Figure 1.

It is necessary at this point to emphasize that morphemes of [ki-], [k^{hi}-], and [xi-] with the combination of one of the final voiceless stops ([k], [p], and [t]) and morphemes of [tɕi], [tɕ^{hi}], and [ɕi] are included in this study. More specifically, morphemes of [kik], [kip], [kit], [k^{hi}ik], [k^{hi}ip], [k^{hi}it], [xik], [xip], [xit], [tɕi], [tɕ^{hi}], and [ɕi] are considered for the present study. Morphemes with a complex nucleus or a final other than [k], [p], or [t] are not included in this study since they may also be involved in other sound changes. The present study tries to exclude any other sound change process as much as possible to solely focus on palatalization. For example, morphemes of [kiak] are not included since they have a complex nucleus [ia]. Morphemes of [ɕin] are not included either because they have the final [n]. From here on, [ki-] will be used to stand for [kik], [kip], and [kit] for convenience of exposition. The same applies to [k^{hi}-] and [xi-].

In order to locate factors in palatalization observed in Table 2 and Figure 1, statistical analysis is conducted in Section 4, taking all factors discussed in Section 2 into consideration.

Table 2: Percentages of unpalatalized and palatalized morphemes in *Piaotongshi xinshi yanjie* (Kim [1765] 2005)

Unpalatalized			Palatalized		
[ki-]	[k ^{hi} -]	[xi-]	[tɕi]	[tɕ ^{hi}]	[ɕi]
29.79% (14/47)	33.33% (10/30)	48.39% (15/31)	70.21% (33/47)	66.67% (20/30)	51.61% (16/31)
Total 36.11% (39/108)			Total 63.89% (69/108)		

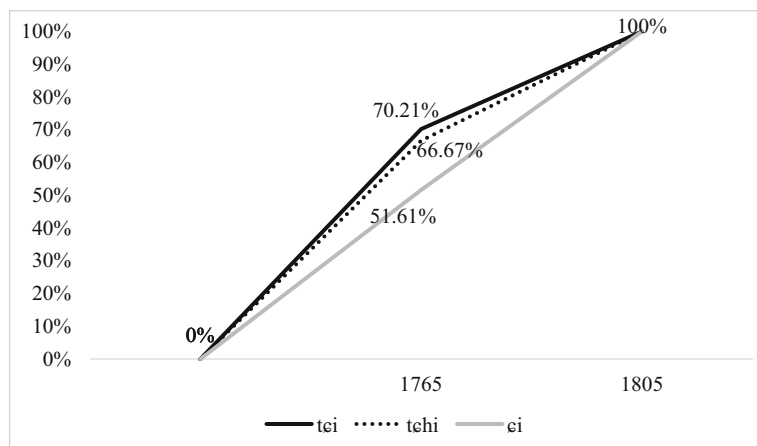


Figure 1: The trajectory of palatalization.

4 Statistical analysis for factors

For ease of exposition, this section will first explain the factors for statistical analysis and then the statistical analysis results.

4.1 Factors for the binary logistic regression analysis

Multiple logistic regression analysis is performed using the GraphPad software to examine the relationship between factors and palatalization. The dependent factor is palatalization, which has binary outcomes: *no* and *yes*, where *no* means that a morpheme under discussion is transcribed as [ki-], [k^hi-], or [xi-] in *Piaotongshi xinshi yanjie* (Kim [1765] 2005), and *yes* means that a relevant morpheme is transcribed as [tɕi], [tɕ^hi], or [ɕi] in *Piaotongshi xinshi yanjie* (Kim [1765] 2005). Table 3 presents a summary of the factors and factor levels for the analysis.

The factor of *word class (1644–1912)* with two levels *content word* and *function word* was configured to examine the contradictory claims concerning the role of word class in sound change discussed in Section 2.1. The time period *1644–1912* means that the data were extracted in the time frame of the Qing Dynasty (1644–1912). Adjectives, nouns, verbs, and so on are classified as content words; adverbial conjunctions, auxiliary verbs, determiners, prepositions, quantifiers, and so on are classified as function words. This is in line with the dichotomy of words in Phillips (1983, 2001, 2006). It may appear ideal to classify words into adjectives, adverbial conjunctions, determiners, nouns, prepositions, verbs, and so on. However, too many factor levels are not preferable for statistical analysis. Mashi Wentong (Master Ma’s linguistic overview; Ma 1898), a classic work on Chinese grammar, has been referred to. To exemplify, the morpheme ‘urgent’ (急) is classified as a content word in Ma (1898). However, as not every related morpheme can be located in Ma (1898), Baxter and Sagart (2014) have also been referred to to decide the word class of each morpheme. The Chinese character 乞 will be used as an example: this character is annotated as ‘beg’ in Baxter and Sagart (2014) and is thus a verb. According to Phillips (1983, 2001, 2006), a verb is a content word. Therefore, this study takes the morpheme ‘beg’ (乞) as a content word.

The factor of *morphosyntactic structure (1644–1912)* with two levels *free* and *bound* was constructed to test whether morphosyntactic structure has a role in sound change. *Free* and *bound*, respectively, mean that a related morpheme is mainly used as a free morpheme and as a bound morpheme. According to Packard (2015), if a morpheme can stand alone as an independent word, it is a free morpheme; if it must be used together with other language materials, then it is a bound morpheme (see also Chao 1968, Hsieh 2016). However, the distinction between free and bound morphemes is not always clear-cut in Chinese: bound morphemes may sometimes be used as free morphemes (Packard 2015, 2020). The present study decides the

Table 3: Factors for the binary logistic regression model

Factor	Factor level	Count (%)
Word class (1644–1912)	Content word	99 (91.67)
	Function word	9 (8.33)
Morphosyntactic structure (1644–1912)	Free	67 (62.04)
	Bound	41 (37.96)
Normalized frequency (1644–1912) Frequency dummy (1644–1912)	Numerical data	
	Low	36 (33.33)
	Medium	36 (33.33)
Frequency change from between 1271 and 1644 to between 1644 and 1912 Frequency change from between 1271 and 1644 to between 1644 and 1912 dummy	High	36 (33.33)
	Numerical data	
	Decrease	67 (62.04)
	Increase	41 (37.96)

morphosyntactic status of a morpheme depending on whether it is mainly used as a free morpheme or a bound morpheme in example sentences extracted from the CNCORPUS: if a morpheme is used as a free morpheme in most sentences, it is marked as *free*; otherwise it is marked as bound. To exemplify, the morpheme ‘urgent’ (急 [tʃi]) in Table 1 is marked as a free morpheme because it is used as a free morpheme in more than 90% of all sentences extracted from the corpus, while the morpheme ‘daughter-in-law, wife’ (媳 [ci]) is marked as a bound morpheme since it is almost always used with another morpheme found in sentences extracted from the corpus.

As noted in Section 2.3, the frequency factor is claimed to have an association with sound change by some scholars, although disagreement exists concerning whether frequent words or less frequent words lead sound change (Grimm 1822, Bybee 1985, 2000, 2002, Pierrehumbert 2001, Bermúdez-Otero 2007, Ogura 2012, Smith 2012, Hay and Foulkes 2016). The raw numerical data of frequency between 1644 and 1912 were examined in the first place. The cross-tabulation analysis carried on the GraphPad software revealed that it was rejected as a statistically significant factor for palatalization ($p = 0.46$). Thus, the raw data of frequency between 1644 and 1912 were normalized on the GraphPad software and reported as *Normalized frequency (1644–1912)* in Table 3.

The factor of *frequency dummy (1644–1912)* was configured due to the debate in Section 2.3.1 concerning whether low-frequency words or high-frequency words lead sound change. It was also partly configured in line with claims in the study by Wedel et al. (2013), reviewed in Section 2.3.2. Wedel et al. claim (2013) that the relative frequency of minimal pair members, instead of the absolute frequencies, is a significant predictor of phoneme merger. The present study does not refer to their relative frequency. Instead, the study refers to all morphemes involved in palatalization. Thus, it took into account the relative frequencies of all related morphemes. The factor *frequency dummy (1644–1912)* has three factor levels: *low*, *medium*, and *high*. Each level takes one third of data: the one third of data with the lowest frequencies in this column is marked as *low*; the one third of data with the highest frequencies is marked as *high*; and the leftover one third of data between *low* and *high* is *medium*. Therefore, *low* here does not refer to frequency lower than a specific count. Instead, it means that the frequency of a certain morpheme is among the lowest frequencies of morphemes involved in palatalization. For example, the morpheme ‘urgent’ [tʃi] in Table 1 is marked as *medium* since its frequency of 3,592 is among the second tertile in terms of frequency.

The next factor *frequency change from between 1271 and 1644 to between 1644 and 1912* (henceforth *frequency change*) was introduced since it seems possible that morphemes with increased frequencies and morphemes with decreased frequencies may have undergone palatalization at different rates. For example, the frequency of the morpheme ‘urgent’ is 4,684 between 1271 and 1644 and 3,592 between 1644 and 1912. Thus, the frequency change from between 1271 and 1644 to between 1644 and 1912 for the morpheme ‘urgent’ is $-1,092$, where the negative number means that its frequency has decreased from between 1271 and 1644 to between 1644 and 1912. In contrast, the frequency of the morpheme ‘opera’ is 1,045 between 1271 and 1644 and 1,072 between 1644 and 1912. As a result, the frequency change for the morpheme ‘opera’ is 27, where the positive number means that its frequency has increased from between 1271 and 1644 to between 1644 and 1912. It is possible to normalize the raw data of frequency change by adding all numbers with the absolute of the most negative. In this way, the most negative number will become zero, and all the other numbers become positive. However, the focus of frequency change is whether related frequencies have increased or decreased. It therefore serves the present study better to use the raw data, instead of normalized data.

The raw numerical data for the aforementioned factor were converted to categorical data with two levels: *decrease* and *increase*, with *decrease* as the reference level. The reason is that it seems that whether frequency has decreased or increased is also a factor in sound change. Therefore, the factor of *frequency change from between 1271 and 1644 to between 1644 and 1912 dummy* (hereafter *frequency change dummy*) was introduced. The morphemes ‘urgent’ and ‘opera’ will be exemplified again: the frequency of ‘urgent’ decreased 1,092 counts from between 1271 and 1644 to between 1644 and 1912, so it is marked as *decrease* for the factor of *frequency change dummy*; the frequency of ‘opera’ increased 27 counts, so it is marked as *increase* for the same factor.

4.2 The binary logistic regression results

The statistical relationships between the six factors in Table 3 and the dependent variable were assessed using the multiple logistic regression of the GraphPad software. Model selection was guided by AIC (Akaike Information Criterion; Akaike 1974, Burnham and Anderson 2004), calculated probability (p value), and VIF (variance inflation factor; Rawlings et al. 1998, James et al. 2017). In Table 4, the results are reported.

The dependent variable has two categories: *unpalatalized* and *palatalized*, with *unpalatalized* as the reference level. A p -value smaller than 0.05 was considered statistically significant.

From here on, time phases will be omitted for easy exposition. For example, *word class (1644–1912)* will be shortened to *word class*. The factor of *frequency change from between 1271 and 1644 to between 1644 and 1912* will be shortened to *frequency change*.

The binary logistic regression analysis for Table 4 indicates that the factors of *word class* and *morpho-syntactic structure* did not emerge as statistically significant ($p = 0.34, 0.82$).

The next two factors present an interesting result: while *frequency dummy* is statistically significant ($p = 0.04$), *normalized frequency* is not shown as a statistically significant factor ($p = 0.18$). The factor *frequency dummy* shows a significant effect in the direction of increasing the possibility for palatalization from low frequency, medium frequency, to high frequency, as indicated by the positive coefficient value (estimate = 0.10). The correlation between this factor and palatalization is visualized in Figure 2.

Two tendencies can be observed in Figure 2: (1) the number of unpalatalized morphemes, shown as striped bars, declines from the low and medium levels to the high level of the frequency dummy factor; (2) the number of palatalized morphemes, shown as gray bars, increases from the low and medium levels to

Table 4: The binary logistic regression results

Variable	Estimate	Std. Error	z	p
(Intercept)	0.66	0.30	2.23	0.03 *
Word class (1644–1912)	1.46	0.86	1.70	0.34
Morphosyntactic structure (1644–1912)	–0.02	0.07	0.22	0.82
Normalized frequency (1644–1912)	0.02	0.01	1.33	0.18
Frequency dummy (1644–1912)	0.10	0.05	1.93	0.04 *
Frequency change from between 1271 and 1644 to between 1644 and 1912	–0.01	0.00	1.21	0.23
Frequency change from between 1271 and 1644 to between 1644 and 1912 dummy	–0.96	0.43	0.38	0.02 *

Notes: * = $p < 0.05$.

|z| stands for the absolute value of z as given by the GraphPad software.

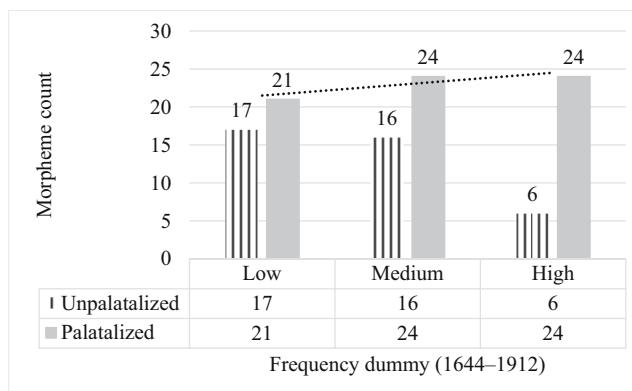


Figure 2: Frequency dummy. The words *low*, *medium*, and *high* in the first row stand for the *low*, *medium*, and *high* levels of the frequency dummy factor, respectively. The words *unpalatalized* and *palatalized* in the leftmost column stand for the two categories of the dependent variable, *unpalatalized morphemes* and *palatalized morphemes*, respectively.

the high level of the same factor. In short, Figure 2 clearly indicates that morphemes in the high-frequency level tend to undergo palatalization earlier.

Table 4 also shows that the factor of *frequency change* is independent of palatalization ($p = 0.23$), while *frequency change dummy* is statistically significantly associated with palatalization ($p = 0.02$). The factor *frequency change dummy* has *decrease* as the reference level. Morphemes with increased frequencies have a lower probability of undergoing palatalization first, as indicated by the negative coefficient value (estimate = -0.96). In other words, morphemes with decreased frequencies were at the forefront of palatalization. The correlation between this factor and palatalization is visualized in Figure 3.

The graphic illustration of quantitative data in Figure 3 provides evidence for the following two tendencies: (1) the number of unpalatalized morphemes, represented as striped bars, increases from the *decrease* level to the *increase* level of the frequency change dummy factor; (2) the number of palatalized morphemes, represented as gray bars, shows a reversed pattern.

To sum, data for this study show different mechanisms for the factors of frequency dummy and frequency change dummy.

5 Discussion and conclusion

This corpus-based study used statistical analysis to locate factors in palatalization. The present study shows that both morphosyntactic structure and word class have been rejected as statistically significant factors in palatalization, while the factors of frequency dummy and frequency change dummy emerged as statistically significant. Results from this study argue against the claim that frequency has minimal or even no role in sound change.

A few extra words will be added to explain the categorical data of frequency and frequency change in this article. The factor *frequency dummy* has three levels, *low*, *medium*, and *high*, with each level taking one third of data. As a result, high frequency is a comparative concept: *high* is defined by comparison with frequencies of all related morphemes. The factor *frequency change dummy* is configured with two levels: *decrease* and *increase*. *Decrease* and *increase* are comparative concepts: a morpheme under discussion should compare with its frequency in an earlier period for its frequency change. The fact that categorical data of frequency, with three levels of low, medium, and high, are statistically significantly correlated with palatalization, rather than numerical data and normalized numerical data, presents an intriguing idea: frequency should be considered from a relative perspective. Frequencies of all morphemes involved in a certain sound change should be compared with each other. Since different morphemes are involved in different sound changes, there is no unified standard of high frequency for all sound changes. Concerning

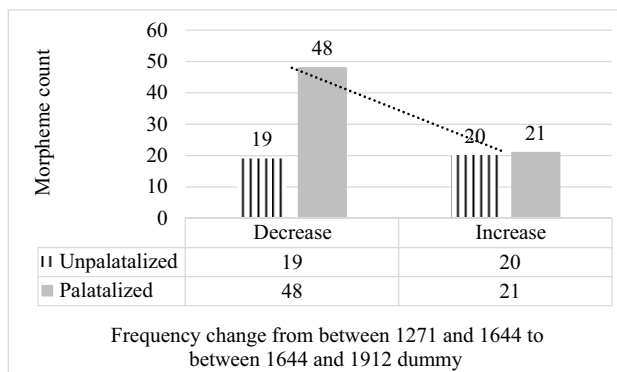


Figure 3: Frequency change dummy. The words *decrease* and *increase* in the first row, respectively, stand for the *decrease* and *increase* levels of the frequency change dummy factor. The words *unpalatalized* and *palatalized* in the leftmost column, respectively, stand for *unpalatalized morphemes* and *palatalized morphemes*, the two categories of the dependent variable.

the frequency change factor, its categorical data with *decrease* and *increase* levels, instead of numerical data, are shown to be statistically significantly correlated with palatalization. This indicates that the exact count of frequency change is not strongly correlated with sound change, but whether frequency has decreased or increased is. The most surprising part is that frequency change has a negative correlation with palatalization: morphemes with decreased frequencies tend to lead palatalization. In short, morphemes with high frequency and decreased frequencies were at the forefront of palatalization. This may be explained by the following. On the one hand, low-frequency morphemes were too inactive to undergo palatalization first, and this is why high-frequency morphemes tended to lead palatalization. On the other hand, increased frequency indicates that a related morpheme was more frequently used. If this morpheme had undergone palatalization first, it may lead to difficulty in communication. Thus, morphemes with increased frequencies tended to lag behind in palatalization.

Acknowledgments: For help in getting this article to its final form, I am grateful to Prof. Eiji Yamada, Prof. Hajime Takeyasu, Prof. Kuixin Zhao, and Prof. Moon Changyun for advice and discussion, to Prof. Jefferson M. Peters and Prof. Stephen Howe for editing my article, and to the anonymous reviewers and the editors of the present journal for detailed and helpful feedback. All remaining errors are my own responsibility.

Funding information: This work was funded by JSPS Grant-in-Aid for Early-Career Scientists (KAKENHI-PROJECT-20K13072).

Author contributions: The author has accepted responsibility for the entire content of this manuscript.

Conflict of interest: The author states no conflict of interest.

References

- Akaike, Hirotugu. 1974. "A new look at the statistical model identification." *IEEE Transactions on Automatic Control* 19(6), 716–23.
- Akiyasu, Tōdō. 1967. "Postwar studies of the Chinese language." In: *Current trends in linguistics*, Vol. II, edited by Yuen Ren Chao, Richard B. Noss, and Joseph K. Yamagiwa, p. 633–44. The Hague: Mouton.
- Anttila, Raimo. 1972. *An introduction to historical and comparative linguistics*. New York: Macmillan.
- Baranowski, Maciej, and Danielle Turton. 2020. "TD-deletion in British English: New evidence for the long-lost morphological effect." *Language Variation and Change* 32(1), 1–23.
- Baxter, William H. and Laurent Sagart. 2014. *Old Chinese: A new reconstruction*. New York: Oxford University Press.
- Bermúdez-Otero, Ricardo. 2007. "Diachronic phonology." In: *The Cambridge handbook of phonology*, edited by Paul de Lacy, p. 497–517. Cambridge: Cambridge University Press.
- Blevins, Juliette and John Lynch. 2009. "Morphological conditions on regular sound change?: A reanalysis of *l-loss in Paamese and Southeast Ambrym." *Oceanic Linguistics* 48(1), 111–29.
- Brown, Esther L. 2013. "Word classes in phonological variation: Conditioning factors or epiphenomena?" In *Selected proceedings of the 15th Hispanic linguistics symposium*, edited by Chad Howe, Sarah Blackwell, and Margaret Lubbers Quesada, p. 179–86. Somerville, MA: Cascadia Proceedings Project.
- Burnham, Kenneth P. and David R. Anderson. 2004. "Multimodel inference: Understanding AIC and BIC in model selection." *Sociological Methods & Research* 33, 261–304.
- Bybee, Joan L. 1985. *Morphology: A study of the relation between meaning and form*. Amsterdam: John Benjamins.
- Bybee, Joan L. 2000. "The phonology of the lexicon: Evidence from lexical diffusion." In *Usage-based models of language*, edited by Michael Barlow and Suzanne Kemmer, p. 65–85. Stanford: CSLI.
- Bybee, Joan L. 2002. "Lexical diffusion in regular sound change." In *Sounds and systems: Studies in structure and change*, edited by David Restle and Dietmar Zaefferer, p. 59–74. Berlin: Mouton de Gruyter.
- Bybee, Joan L. 2012. "Patterns of lexical diffusion and articulatory motivation for sound change." In *The initiation of sound change: Perception, production, and social factors*, edited by Maria-Josep Solé and Daniel Recasens, p. 211–34. Philadelphia, PA: John Benjamins.
- Bybee, Joan L. 2017. *Grammatical and lexical factors in sound change: A usage-based approach*. Cambridge: Cambridge University Press. <https://www.cambridge.org/core/journals/language-variation-and-change/article/grammatical-and->

- lexical-factors-in-sound-change-a-usagebased-approach/628862C61B7CFA9FE4569CC0E2687E2A (accessed February 9, 2021).
- Campbell, Lyle. 2013. *Historical linguistics* (3rd edn). Cambridge, MA: MIT Press.
- Chao, Yuen-Ren. 1968. *A grammar of spoken Chinese*. Berkeley: University of California Press.
- Chen, Matthew Y. 1976. "From Middle Chinese to modern Peking." *Journal of Chinese linguistics* 4(2/3), 113–277.
- Chou, Fa-ka. 1993. *A pronouncing dictionary of Chinese characters in Archaic & Ancient Chinese, Mandarin & Cantonese*. Beijing: Zhonghua Book.
- Crowley, Terry. 1997. *An introduction to historical linguistics* (3rd edn). New York: Oxford University Press.
- Dinkin, Aaron. 2008. "The real effect of word frequency on phonological variation." *University of Pennsylvania Working Papers in Linguistics* 14, 97–106. <http://itre.cis.upenn.edu/myl/papers/Dinkin2007.pdf> (accessed December 14, 2020).
- Dong, Tonghe. 1954. *Zhongguo Yuyinshi [The History of Chinese Phonology]*. Taipei: China Cultural Publication Foundation.
- Donohue, Mark. 2005. "Syntactic and lexical factors conditioning the diffusion of sound change." *Oceanic Linguistics* 44(2), 427–42.
- Edkins, Joseph. 1857. *A grammar of the Chinese colloquial language, commonly called the Mandarin dialect*. Shanghai: London Mission.
- Edkins, Joseph. 1864. *A grammar of the Chinese colloquial language, commonly called the Mandarin dialect* (2nd edn). Shanghai: Presbyterian Mission.
- Fasold, Ralph W. 1972. *Tense marking in Black English: A linguistic and social analysis*. Arlington, VA: Center for Applied Linguistics.
- Fruehwald, Josef, Jonathan Gress-Wright, and Joel C. Wallenberg. 2013. "Phonological rule change: The constant rate effect." In *NELS 40: Proceedings of the 40th annual meeting of the north east linguistic society*, Vol. I edited by S. Kan, C. Moore-Cantwell, and R. Staubs, p. 219–30. California: CreateSpace Independent Publishing Platform. https://www.pure.ed.ac.uk/ws/portalfiles/portal/14416788/Fruehwald_Gress_Wright_Wallenberg_Phonological_Rule_Change.pdf (accessed December 12, 2020).
- Fudge, E.C. 1972. "Review of Postal 1968." *Journal of Linguistics* 8, 136–56.
- GraphPad Prism. 2019. *Version 8.0.0 for Windows*. San Diego, California: GraphPad Software, Computer software.
- Grimm, Jacob. 1822. *Deutsche Grammatik*, Vol. I. Göttingen: Dietrich.
- Guy, Gregory R. 1991a. "Contextual conditioning in variable lexical phonology." *Language Variation and Change* 3(2), 223–39.
- Guy, Gregory R. 1991b. "Explanation in variable phonology: An exponential model of morphological constraints." *Language Variation and Change* 3(1), 1–22.
- Handel, Zev. 2017. "The Sinitic languages: Phonology." In *The Sino-Tibetan languages*, (2nd edn), edited by Graham Thurgood and Randy J. LaPolla, p. 85–113. London: Routledge.
- Hay, Jennifer B., Janet B. Pierrehumbert, Abby J. Walker, and Patrick LaShell. 2015. "Tracking word frequency effects through 130 years of sound change." *Cognition* 139, 83–91.
- Hay, Jennifer B. and Paul Foulkes. 2016. "The evolution of medial/t/over real and remembered time." *Language* 92(2), 298–330.
- Hsieh, Shu-kai. 2016. "Chinese linguistics: Semantics." In *The Routledge encyclopedia of the Chinese language*, edited by Chan Sin-Wai, p. 203–14. New York: Routledge.
- James, Gareth, Daniela Witten, Trevor Hastie, and Robert Tibshirani. 2017. *An Introduction to statistical learning* (8th edn). New York: Springer.
- Jasanoff, Jay. 1971. "A generative approach to historical linguistics." *Romance Philology* 25, 74–85.
- Jespersen, Otto. (1922) 2007. *Language: Its nature and development*. London: Routledge.
- Jin, Jishi. 2001. "Palatalized velars and alveolar sibilants and their transcriptions in Korean works." *Studies of the Chinese language* 281(2), 159–65.
- Kapatsinski, Vsevolod. 2010. "What is it I am writing? Lexical frequency effects in spelling Russian prefixes: Uncertainty and competition in an apparently regular system." *Corpus Linguistics and Linguistic Theory* 6(2), 157–215.
- Kim, Changjo. (1765) 2005. *Piaotongshi Xinshi Yanjie [New Edition of Vernacular Exposition of Pak the Interpreter]*. Beijing: Zhonghua Book.
- Kim, Youngman. 1989. "Middle Mandarin phonology: A study based on Korean data." *PhD thesis*. Columbus, Ohio: The Ohio State University.
- King, Robert D. 1969. *Historical linguistics and generative grammar*. Englewood Cliffs, NJ: Prentice-Hall.
- Kiparsky, Paul. 2014. "New perspectives in historical linguistics." In *The Routledge handbook of historical linguistics*, edited by Claire Bownen and Bethwyn Evans, p. 64–102. New York: Routledge.
- Kroch, Anthony S. 1989. "Reflexes of grammar in patterns of language change." *Language Variation and Change* 1, 199–244.
- Kroch, Anthony S. 1994. "Morphosyntactic variation." In *Proceedings of the thirtieth annual meeting of the Chicago Linguistics Society*, Vol. II edited by K. Beals, p. 180–201.
- Labov, William. 1989. "Exact description of the speech community: Short a in Philadelphia." In *Language change and variation* edited by Ralph W. Fasold and Deborah Schiffrin, p. 1–57. Amsterdam: John Benjamins.
- Labov, William. 2010. *Principles of linguistic change, Vol. III: Cognitive and cultural factors*. Oxford: Blackwell.

- Labov, William, Paul Cohen, Clarence Robins, and John Lewis. 1968. *A study of the nonstandard English of Black and Puerto Rican speakers in New York City (Cooperative Research Report No. 3288)*. Washington, DC: U.S. Office of Education.
- Li, Ruzhen. (1805) 1992. *Lishiyinjian [The rhythm book of Liruzhen]*. Xi'an: Shaanxi People's Education.
- Li, Wen-Chao. 1999. *A diachronically-motivated segmental phonology of Mandarin Chinese*. New York: Peter Lang.
- Lin, Binhui. 2010. *A selected collection of Chinese language learning textbooks used outside of China: Research into ancient Chinese novels and operas (the 14th century ~ 20th century)*. Changsha: Hunan People's.
- Ma, Jianzhong. 1898. *Mashi Wentong [Master Ma's linguistic overview]*. Shanghai: The Commercial Press, Shanghai. <https://ctext.org/wiki.pl?if=gb&chapter=223793>.
- Malkiel, Yakov. 1968. *Essays on linguistic themes*. Oxford: Blackwell.
- Malkiel, Yakov. 1969–70. "Sound changes rooted in morphological conditions: The case of Old Spanish/sk/changing to/dk/." *Romance Philology* 23, 188–200.
- Malkiel, Yakov. 1976. "Multi-conditioned sound change and the impact of morphology on phonology." *Language* 52(4), 757–78.
- Manker, Jonathan. 2015. "Phonetic sources of morphological patterns in sound change: Fricative voicing in Athabaskan." *UC Berkeley Phonology Lab Annual Report*, p. 243–94.
- Morrison, Robert. 1815–23. *A dictionary of the Chinese language, in three parts*. Macao: East India.
- Mortensen, David R. 2012. "The emergence of obstruents after high vowels." *Diachronica* 29(4), 434–70.
- Norman, Jerry. 1988. *Chinese*. New York: Cambridge University Press.
- Ogura, Mieko. 2012. "The timing of language change." In *The handbook of historical sociolinguistics* edited by Juan Manuel Hernández-Campoy and Juan Camilo Conde-Silvestre, p. 427–50. West Sussex: Blackwell.
- Packard, Jerome L. 2015. "Morphology: Morphemes in Chinese." In *The Oxford handbook of Chinese linguistics* edited by William S-Y. Wang and Chaofen Sun, p. 263–305. New York: Oxford University Press.
- Packard, Jerome L. 2020. Chinese morphology. In *The Routledge encyclopedia of the Chinese language* edited by Chan Sin-wai, p. 215–26. New York: Routledge.
- Phillips, Betty S. 1983. "Lexical diffusion and function words." *Linguistics* 21, 487–99.
- Phillips, Betty S. 2001. "Lexical diffusion, lexical frequency, and lexical analysis." In *Frequency and the emergence of linguistic structure* edited by Joan Bybee and Paul Hopper, p. 123–36. Amsterdam: John Benjamins.
- Phillips, Betty S. 2006. *Word frequency and lexical diffusion*. New York: Palgrave Macmillan.
- Pierrehumbert, Janet B. 2001. "Exemplar dynamics: Word frequency, lenition and contrast." In *Frequency and the emergence of linguistic structure* edited by Joan Bybee and Paul Hopper, p. 137–57. Amsterdam: John Benjamins.
- Pintzuk, Susan. 1991. "Phrase structures in competition: Variation and change in Old English word order." *PhD thesis*. Philadelphia, Pennsylvania: University of Pennsylvania.
- Postal, Paul. 1968. *Aspects of phonological theory*. New York: Harper and Row.
- Pulleyblank, Edwin G. 1984. *Middle Chinese: A study in historical phonology*. Vancouver: University of British Columbia Press.
- Pulleyblank, Edwin G. 1991. *Lexicon of reconstructed pronunciation in Early Middle Chinese, Late Middle Chinese, and Early Mandarin*. Vancouver: University of British Columbia Press.
- Rawlings, John O., Sastry G. Pantula, and David A. Dickey. 1998. *Applied regression analysis: A research tool* (2nd edn). New York: Springer.
- Renwick, Margaret, Ladan Baghai-Ravary, Rosalind Temple and John Coleman. 2014. "Deletions in big data? The phonetics of word-final (t,d) in the Audio BNC." *Oral presentation at BAAP 2014 (British Association of Academic Phoneticians)*.
- Robinson, Orrin W. III and Frans van Coetsem. 1973. "Review article of King 1969." *Lingua* 31, 331–69.
- Santorini, Beatrice. 1992. "Variation and change in Yiddish subordinate clause word order." *Natural Language & Linguistic Theory* 10, 595–640.
- Santorini, Beatrice. 1993. "The rate of phrase structure change in the history of Yiddish." *Language Variation and Change* 5, 257–83.
- Shen, Zhongwei. 2015. "Early Mandarin seen from Ancient Altaic scripts: The rise of a new phonological standard." In *The Oxford handbook of Chinese linguistics* edited by William S-Y. Wang and Chaofen Sun, p. 91–104. New York: Oxford University Press.
- Shen, Zhongwei. 2020. *A phonological history of Chinese*. New York: Cambridge University Press.
- Shi, Yuzhi. 2002. *The establishment of Modern Chinese grammar: The formation of the resultative construction and its effects*. Amsterdam: John Benjamins.
- Sihler, Andrew. 2000. *Language history: An introduction*. Amsterdam: Benjamins.
- Smith, K. Aaron. 2012. "Frequency and language change." In: *English historical linguistics: An international handbook*, Vol. II edited by Alexander Bergs and Laurel J. Brinton, p. 1531–46. Berlin: Mouton de Gruyter.
- Stemberger, Joseph Paul and Brian MacWhinney. 1986. "Frequency and the lexical storage of regularly inflected forms." *Memory & Cognition* 14, 17–26.
- Stimson, Hugh M. 1966. *The Jongyuan In Yunn: A guide to Old Mandarin pronunciation*. New Haven: Yale University Press.
- Thomsen, Vilhelm. 1879. Andare-andar-anar-aller: En kritisk-etymologisk undersøgelse. *Det philolhist*, p. 197–214. Kjøbenhavn: Samfunds mindeskrift.

- Thomsen, Vilhelm. 1920. "Den gotiske sprogklasses indflydelse på den finske. En sproghistorisk undersøgelse." In *Samlede afhandlinger*, Vol. II, p. 49–264. Copenhagen: Kristiana.
- Wang, Li. 1957. *Hanyushigao [A sketch of the history of Chinese]*. Beijing: Science.
- Wang, Li. 1985. *Hanyu Yuyinshi [The history of Chinese Phonology]*. Beijing: China Social Sciences.
- Wedel, Andrew, Jackson, Scott, and Kaplan Abby. 2013. "Functional load the lexicon: Evidence that syntactic category and frequency relationships in minimal lemma pairs predict the loss of phoneme contrasts in language change." *Language and Speech* 56(3), 395–417.
- Xiao, Hang. *CNCORPUS*. 2010. <http://www.aihanyu.org/cncorpus/index.aspx> (accessed June 17, 2021).
- Yuanyin Zhengkao [On the alveolar and alveolo-palatal consonants]. (1743) 2015. Beijing: Capital Normal University Press. <https://ctext.org/library.pl?if=gb&res=81809&remap=gb> (accessed August 28, 2020).
- Zellou, Georgia, and Meredith Tamminga. 2014. "Nasal coarticulation changes over time in Philadelphia English." *Phonetics* 47, 18–35.
- Zhou, Deqing. (1324) 1996. "Zhongyuan Yinyun." In *Chuanshi cangshu jingku*, Vol. II: Language and word. Hainan: Hainan International.