Contrastiveness and diachronic variation in Chinese nasal codas

Tsz-Him Tsui
The Ohio State University

Abstract: Among the nasal codas across Chinese languages, [-m] underwent sound changes more often than [-n] or [-ŋ] (Zee 1985). The current paper argues that such diachronic variations in Chinese nasal codas could be explained by their relative contrastiveness (Hume et al 2012). The [-m] coda, being the least contrastive, was more variable in perception and production, and thus more likely to undergo sound changes.

1 Introduction

Previous studies found that the bilabial nasal coda [-m] underwent historical sound changes more often than the other codas, [-n] and [-ŋ], across Chinese languages. While such patterns of sound change are described very clearly (Zee 1985), it remains unclear why there is such a tendency for [-m] to change. This paper attempts to fill in this gap, and explain the preference for [-m] to undergo sound change in Chinese languages. The current proposal is that [-m] was the least contrastive nasal coda in Middle Chinese, and was used least often to distinguish minimal pairs. Therefore, changes to [-m] would result in the least amount of confusion among Middle Chinese speakers, and [-m] is more prone to sound change than the other nasal codas.

This short paper is organized as follows: Section 2 reviews the diachronic development of nasal codas in Chinese languages, as well as studies on synchronic perception and production of Chinese varieties. We will suggest that the synchronic data is not conclusive enough for us to infer the perception and production patterns in Middle Chinese. Section 3 describes the quantification of functional load and relative contrastiveness of phonological elements, using tools from information theory (Shannon 1948, Surendran & Niyogi 2003, Hume et al. 2012). Our current study is described in Section 4. We will conclude in Section 5, and suggest that a usage-based account of sound change patterns in Chinese nasal coda is plausible.

2 Nasal codas in Chinese languages

2.1 Diachronic development

Three nasal codas, [-m] [-n] and [-ŋ], were present in Middle Chinese, spoken approximately between 6th to 12th century CE (Pulleyblank 1984). However, contemporary Chinese varieties often lack some of these nasal codas, as they either merged with another nasal or lost in diachronic developments. Zee (1985) tracked the diachronic changes on nasal codas from Middle Chinese to modern Chinese languages, and found that the bilabial nasal coda [-m] underwent both merger and deletion more often than the alveolar [-n] and the velar [-ŋ]. Changes to [-m] only occurred in more Chinese varieties, but also occurred in more widespread geographically. For example in (1), the -m > -n change appeared in 16 Chinese varieties, which belonged to 8 geographically-diverse sub-families. On the other hand, changes to [-n] and [-ŋ] appeared in fewer varieties, and often in more closely-related sub-families.

In proceedings of Northwest Linguistics Conference 29,
University of British Columbia Working Papers in Linguistics 38,
(1) Some major historical sound changes from Middle Chinese to modern Chinese varieties (Zee 1985)

<table>
<thead>
<tr>
<th>Type of change</th>
<th># of varieties involved</th>
<th># of sub-families involved</th>
</tr>
</thead>
<tbody>
<tr>
<td>-m &gt; -n</td>
<td>16</td>
<td>8</td>
</tr>
<tr>
<td>-n &gt; -Ṽ</td>
<td>10</td>
<td>6</td>
</tr>
<tr>
<td>-ŋ &gt; -n</td>
<td>11</td>
<td>8</td>
</tr>
</tbody>
</table>

Zee (1985) also suggested that such tendency to make changes to [-m] may not be explained by previous proposals, which suggested that changes to [-n] would occur most often.

2.2 Can patterns of synchronic variation explain the diachronic patterns?

It may be possible to explain the diachronic tendency for [-m] coda to change through variations in perception or production. For example, among the three nasal codas, [-m] may be the most variably perceived one. In other words, it may be the case that [-m] was often perceived as [-n] or [-ŋ] by Middle Chinese speakers and/or speakers of later Chinese varieties, because of the acoustic similarities between [-m] and [-n / -ŋ] (CHANCE or CHANGE in Blevins 2004). As a result, Middle Chinese speakers may have misheard the [-m] codas, and have reanalyzed them as [-n] or [-ŋ] (Ohala 1981).

To infer perception patterns in Middle Chinese and subsequent Chinese varieties, we may look into synchronic data in modern Chinese varieties. For example, if erroneous perceptions of [-m] occur more often than those of [-n] or [-ŋ] among speakers of modern Chinese varieties, we would have reasons to believe that [-m] may also have been the most variably perceived nasal coda for Middle Chinese or subsequent Chinese speakers. Because of such biases in mis-perceiving [-m] as [-n] or [-ŋ], [-m] was reanalyzed as [-n] or [-ŋ] more often than the other way around. Therefore, mergers of [-m] into another nasal occurred more often than other potential nasal mergers.

Chen & Guion-Anderson (2011) studied the confusability of nasal codas in Southern Min Taiwanese and Beijing Mandarin. Southern Min Taiwanese has all three nasal coda [-m], [-n] and [-ŋ], while Beijing Mandarin now has [-n] and [-ŋ] only. The study found that Taiwanese speakers indeed misheard [-m] the most, which is also the coda that underwent sound change most often. Thus, perceptual confusability appears to be a potential explanation for the diachronic pattern Chinese varieties — the [-m] coda may have more variably perceived in Middle Chinese, and thus listeners were more likely to reanalyze the [-m] coda as another phoneme than reanalyzing [-n] or [-ŋ]. As a result, [-m] underwent sound change more often than [-n] and [-ŋ].

However, this perceptual explanation is not as conclusive as we hoped it to be. Chen & Guion-Anderson (2011) also found that while [-m] is the most confusable nasal coda in Taiwanese, [-ŋ] is somewhat less confusable than [-m], while [-n] is the least confusable coda. On the other hand, Beijing Mandarin speakers misidentified [-n] more often than [-ŋ]. Therefore, there does not seem to be any pan-Chinese perceptual hierarchy or confusion matrix; [-n] is less confusable than [-ŋ] in Taiwanese, but is more confusable than [-ŋ] in Mandarin. Since different Chinese varieties may have different confusion matrices for their nasal codas, we do not have any conclusive evidence, based on perceptual experiments alone, that [-m] was more often misidentified across most Chinese varieties.

If perception evidence is not conclusive, how about production studies? It may be the case that the [-m] coda, while not the most variably perceived one, was the most variably produced coda in Middle Chinese. The [-m] coda may have often been produced as [-n] and/or [-ŋ], and speakers of...
Middle Chinese gradually chose [-n] or [-ŋ], rather than [-m], as the underlying form (CHOICE in Blevins 2004). Thus, in subsequent Chinese varieties, [-m] is often merged into other nasal codas.

Again, such claims about productions of [-m] in Middle Chinese may be inferred from synchronic data. Some modern Chinese varieties still retain all three nasal codas, [-m] [-n] and [-ŋ]. If speakers of such Chinese varieties produce the [-m] codas more variably, often as [-n] or [-ŋ], than the alveolar [-n] or velar [-ŋ], we have reasons to believe that such production pattern may have occurred in Middle Chinese as well.

Zee (1999) studied the production of nasal codas Hong Kong Cantonese, which has all three nasal codas [-m] [-n] and [-ŋ]. He found that, while variations in production of nasal codas existed, the codas are not equally variable. Specifically, Cantonese speakers' production of [-ŋ] was often misproduced as [-n], while [-n] could also be produced as [-ŋ], albeit to a lesser extent. Most importantly, however, the bilabial [-m] appeared to be the most stably articulated coda, and was only rarely produced as [-n]. Thus, we do not have any evidence that [-m] was the most variably produced coda in Middle Chinese, based on our current synchronous data.

In sum, the synchronous patterns are inconclusive. We have no evidence that a pan-Chinese hierarchy of variability exist for perception and/or production. Thus, we shall look into patterns in Middle Chinese itself in order to explain the tendency for [-m] to undergo sound change.

3 Contrastiveness

Speech sounds are often categorically perceived (Liberman et al. 1957). Speakers are often able to identify contrastive sounds in the ends of a phonetic continuum, and are inconsistent in identifying sounds in the center of the continuum. As a result, phonological contrasts were often conceptualized as binary, in terms of phonological features (e.g. Chomsky & Halle 1968). For example, a speech sound may be [+ labial] or [-labial], while without any intermediate stages between [+ labial] and [-labial].

Recent studies, however, began to examine the gradient nature of phonological contrasts. While these study acknowledged the binary nature of phonological contrasts, they suggested that certain contrasts are stronger/more useful than others, and used tools from information theory (Shannon 1948) to quantify the usefulness, or functional load, of phonological contrasts. Surendran & Niyogi (2003) studied the probability (p) and uncertainty (i.e. entropy, Η) in predicting word forms in corpora, with and without particular phonological contrasts, using the following definition; the more uniformly distributed the word forms are, and/or the more word forms we have, the more uncertainty (and thus, higher entropy value) we would have.

\[ Η = \Sigma p \cdot \log_2 p \]

If we were to merge any two sounds (say a and b) and eliminate this phonological contrasts in a corpus, homophones would be created, and the number of word form types would be decreased. Since there are fewer word forms to choose from, the entropy in predicting word forms would be decrease. The more often a phonological contrast is used to distinguish minimal pairs, the more important it is for communication purposes — it would create a lot of confusion in communication if speakers and listeners could not distinguish minimal pairs that appear frequently. The importance of a particular phonological contrast could be measured in terms of the amount of decrease in entropy in word-form prediction when that contrast is lost. Thus, Surendran and Niyogi (2003) quantified the functional load of a particular phonological contrast as followed:
(3) Usefulness / Functional load between two sounds, \(a\) and \(b\) (Surendran & Niyogi 2003):

\[
\text{Functional load} = \frac{H_{a\neq b} - H_{a=b}}{H_{a=b}}
\]

where
- \(H_{a\neq b}\) is the entropy when \(a\) and \(b\) are contrastive
- \(H_{a=b}\) is the entropy when \(a\) and \(b\) are merged

This way, functional load is defined as the proportional decrease in entropy when a phonological contrast is lost. In other words, functional load of a phonological contrast is the amount of uncertainty, which that contrast is responsible for resolving.

While the definition in (3) quantifies the importance of a particular phonological contrast, Hume et al. (2012) extended this methodology to quantify the importance of a particular sound. Their study focused on the vowel deletion and insertion patterns in French, which target the vowels [ø] and [œ]. This is unexpected, since these front rounded vowels, [ø] and [œ], are often deemed as marked vowels (e.g. Chomsky & Halle 1968), while previous studies suggested that marked vowels are less likely to be inserted (e.g. Rice 1999). To account for such vowel deletion/insertion patterns in French, Hume et al. (2012) first calculated the functional loads in every single phonological contrast between the vowels in French. They then averaged the functional loads associated with a particular vowel to achieve the relative contrastiveness of a particular vowel. This is summarized in (4) below:

(4) Relative contrastiveness of a particular sound \(a\) (Hume et al. 2012):

\[
\text{Relative contrastiveness} = \frac{\sum_{m \in M} H_{a\neq x} - H_{a=x}}{|M|}
\]

where
- \(M\) is the set of all possible mergers, \(m\)
- \(x\) are the other sounds that are contrastive to \(a\), the sound in focus
- \(H_{a\neq x}\) is the entropy when \(a\) and \(x\) are contrastive
- \(H_{a=x}\) is the entropy when \(a\) and \(x\) are merged

Hume et al. (2012) found that the vowels [ø] and [œ] are the least contrastive, and thus insertion or deletion of them would not hinder communication as much as inserting/deleting other vowels. Thus, they concluded that contrastiveness is a contributing factor to the French vowel insertion and deletion patterns.

4 Current study

4.1 Materials and methods

The current study capitalizes on the proposed quantification of relative contrastiveness, as described in the previous session, and investigates whether the diachronic variation in nasal codas in Chinese is related to the contrastiveness of nasal codas in Middle Chinese. Given that the bilabial nasal coda [-m] appears to be more volatile, and participated in sound changes more often than [-n] and [-ŋ], our hypothesis is that the [-m] coda is used less often to distinguish minimal pairs, and is less contrastive, than [-n] and [-ŋ] in Middle Chinese. Changes to [-m] would create less confusion in communication, and Middle Chinese speakers were more prone to variably produce and/or perceive the [-m] coda than [-n] or [-ŋ]. Thus, [-m] was more likely to be reanalyzed by Middle Chinese speakers, and undergo historical sound change (Blevins 2004).

In order to quantify the contrastiveness of the nasal codas in Middle Chinese, an XML copy of Guangyun (Murakoshi 2004) was used. Guangyun is a rhyme dictionary, compiled between 1007
and 1008 during the Song Dynasty. The original listed out syllables in Middle Chinese, ordered by their onsets, rhymes and tones, and is still the main source for reconstruction of Middle Chinese phonology (Norman 1988). While the original Guangyun included only the traditional Chinese names for the onsets, rhymes and tones, the current XML Guangyun contained the proposed reconstructions in IPA symbols as well. For the current study, these reconstructions in IPA are used.

(5) Sample entries from the XML Guangyun (Murakoshi 2004)

```xml
<voice_part ipa="dʰuŋ˩ onyomi="トウ">
  <word_head id="w108a0402">
    同<note>齊也...<fanqie>徒紅切</fanqie></note></word_head>
</voice_part>

<voice_part ipa="ţĭuŋ˥˩ onyomi="チュウ">
  <word_head id="w108b0407">
    中<note>平也成也...<fanqie>陟弓切</fanqie>又陟仲切四</note></word_head>
  <word_head id="w108b0701">
    衷<note>善也正也通也中也又衷衣褻衣也</note></word_head>
  <word_head id="w108b0702">
    忠<note>無私也敬也直也厚也...</note></word_head>
</voice_part>
```

The XML Guangyun included 25,334 entries. The reconstructed pronunciations were extracted from each entry. After tallying the entries, we found 3745 syllable types in total, and the number of tokens for each syllable shape types from 1 to 87.

(6) Histogram of syllable tokens in Guangyun
We calculated the entropy in predicting syllable types in the Guanyun corpus, before and after simulating the mergers of the nasal codas pairs [-m]~[-n], [-m]~[-ŋ], and [-n]~[-ŋ], as described previously in (4). Using these entropy values, we then followed (2–4) and quantified the functional load of each of the coda pairs, as well as the relative contrastiveness of each of the nasal codas.

4.2 Results

The findings are summarized in (7) and (8):

Functional loads of the phonological contrasts between nasal codas in Guanyun

\[
\begin{align*}
[-m] & : 0.006486 \\
[-n] & : 0.004061 \\
[-ŋ] & : -0.007576
\end{align*}
\]

(7) Relative contrastiveness of nasal codas in Guanyun

\[
\begin{align*}
[-m] &= (0.006486 + 0.004061) / 2 = 0.005273 \\
[-n] &= (0.006486 + 0.007576) / 2 = 0.007031 \\
[-ŋ] &= (0.004061 + 0.007576) / 2 = 0.005818
\end{align*}
\]

The bilabial nasal coda [-m] appeared to be less contrastive (0.005273) than [-n] (0.007031) and [-ŋ] (0.005818), confirming our hypothesis. Thus, [-m] contributed less to distinguish minimal pairs than [-n] and [-ŋ] in Middle Chinese. Variation in perception or production of [-m] would create less confusion than those of [-n] or [-ŋ]. Therefore, it was less important for Middle Chinese speakers to correctly produce or perceive [-m] than other nasal codas. [-m] would be more prone to variations in production or perception, and reanalysis by listeners. As a result, [-m] underwent historical sound changes more often than [-n] and [-ŋ].

5 Conclusion

From our results, it appears that the relative contrastiveness may explain a pattern of sound change in varieties of Chinese. The [-m] coda in Middle Chinese underwent sound changes more often than the other two codas, [-n] and [-ŋ]. Since [-m] was used the least in distinguishing minimal pairs, its lack of contrastiveness may explain its tendency to undergo historical sound change. Such results reaffirm that patterns of language use may influence variability in production and perception of sounds, and may explain certain sound change patterns (e.g. Bybee 2006).

References


Smith, Bridget. 2010. *Eth and theta: a tale of two phonemes*. Columbus: The Ohio State University, MS.


