Reconstituting Modern Mandarin
An Empirical Reassessment of Reconstruction by Philological Categories and Sinoxenic Correspondence
Chris Wen-Chao Li
San Francisco State University

Evaluating Reconstruction Methodology

- Saussure’s Proto–Indo–European laryngeals (1879) vs discovery of Hittite (1915)
- R.A. Hall’s (1976): reconstructed proto–Romance vs recorded Latin
- Present study: reconstruction of modern Chinese language using materials similar to those used for historical work (categories and correspondences)

Methods of Traditional Chinese Phonological Reconstruction

- Philological categories (initials, rimes, rime groups, grades)
- Sinoxenic correspondences (c.f. Rosetta Stone)
- Structuralism/descriptive linguistics (symmetry, distribution, naturalness, patterns of change)
- (Comparison with modern dialects)

Sources of Problems in Chinese Historical Phonology

- Language internal: structural complexity
- Language external: language variation and/or artificial conventions (mixtum compositum of earlier traditions)
- By product of reconstruction method employed

Methodology: Syllable Structure

- Root Structure
- (C) (G) V (X)
  - C=consonant
  - G=glide (approximant)
  - V=vowel
  - X=consonant or vowel

Sinoxenic Materials (Chinese Tourist Phrasebooks)
Sinoxenic Materials
(Sample Entries)

- Do you speak English?
  度油司必克 英格力序
- Where are you going?
  惠兒阿油 勾印
- Anything to declare?
  娶泥辛 兔地克淚兒

(From Wu & Brown 2003)

Correspondence Source

- 吳哲夫・Dan F. Brown - 2003 -
  《用中文溜美國話》 - 台北：三思堂 -
- Chapter 1: Basic vocabulary list (pp.6–52)
  - 1,401 characters
- Source language: American English

Philological Categories: Initials List

「太平歌」 (anonymous)

子夜久難明
喜報東方亮
此日笙歌頌太平
眾口齊歡唱

(22 characters)

Philological Categories: Initial Features

<table>
<thead>
<tr>
<th>1 唇 labial</th>
<th>2 舌 lingual</th>
<th>3 齒 dental</th>
<th>4 牙 molar</th>
<th>5 舌根 guttural</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1 牙尖 “tooth tip”</td>
<td>3.2 正齒 “full tooth”</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3.1.1 尖 sharp</td>
<td>3.1.2 圓 blunt</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Philological Categories: Medials

四呼 (4 exhalations): 開、齊、合、撮

凡音皆由內而外——
- 初出於喉，平舌舒唇，謂之「開口」。
- 舉舌對齒，聲在舌，謂之「齊齒」。
- 斂撮而蓄之，聲滿頤輔之間，謂之「合口」。
- 訖響而出聲，謂之「撮口」。

（潘耒《類音》——見竺家寧1991：151）
Philological Categories: Rimes (2)

Exemplars: 开 齐 合 撮

Reconstruction of Initials:

Labial Series

<table>
<thead>
<tr>
<th>Initial</th>
<th>齐</th>
<th>明</th>
<th>方</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reconstruction</td>
<td>[b] (83.9%)</td>
<td>[p] (100%)</td>
<td>[m] (97.8%)</td>
</tr>
</tbody>
</table>

- Initial 方 is used for both [f] (73.1%) or [v] (26.9%)
- [f] is uniquely associated with initial 方
- [v] is also associated with initials 齐 [b] and 明 [p]
- Initial 方 is reconstructed as [f]

Lingual Series

<table>
<thead>
<tr>
<th>Initial</th>
<th>亮</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reconstruction</td>
<td>[l] (43.9%) or [j] (54.6%)</td>
</tr>
</tbody>
</table>

- Initial 亮 is used for both [l] (54.6%) or [j] (43.9%)
- [l] is uniquely associated with initial 亮
- [j] is also associated extensively with initial 日 [j] (87.0%)
- Initial 亮 is reconstructed as tap [ɻ], which shares acoustic properties with both [l] and [j]

Molar Series

<table>
<thead>
<tr>
<th>Initial</th>
<th>歌</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reconstruction</td>
<td>[g] (90.6%)</td>
</tr>
</tbody>
</table>

- Initial 歌 is used for both [g] (90.6%)
- [g] is uniquely associated with initial 歌

Guttural Series

<table>
<thead>
<tr>
<th>Initial</th>
<th>欢</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reconstruction</td>
<td>[h] (100%)</td>
</tr>
</tbody>
</table>

- Initial 欢 is used for both [h] (100%)
- [h] is uniquely associated with initial 欢

Category-internal symmetry dictates initial 众 should be [sɻ] instead of [s]. Possible explanations for discrepancy:
- Rarity of [sɻ] cluster in English, as opposed to [dɻ] and [tɻ]
- Possible merger of [sɻ] and [s]
- CCVX syllable structure dictates that initial must be a single consonant, rather than a cluster such as [dɻ], [tɻ] or [sɻ]. Thus the initials in this category are most likely not clusters, but rather consonants with rhotic properties, i.e., retroflex consonants.
- Series reconstructed as retroflex:

<table>
<thead>
<tr>
<th>Initial</th>
<th>唱</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reconstruction</td>
<td>[dɻ] (75%)</td>
</tr>
</tbody>
</table>

- Initial 唱 is used for both [dɻ] (75%)
- [dɻ] is also associated with initial 唱

“Full Tooth” Series

<table>
<thead>
<tr>
<th>Initial</th>
<th>笙</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reconstruction</td>
<td>[sɻ] (78.6%)</td>
</tr>
</tbody>
</table>

- Initial 笙 is used for both [sɻ] (78.6%)
Reconstruction of Initials: “Tooth Tip (sharp)” Series

<table>
<thead>
<tr>
<th>Initial</th>
<th>子</th>
<th>此</th>
<th>頌</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reconstruction</td>
<td>[z]</td>
<td>(76%)</td>
<td>[t]</td>
</tr>
</tbody>
</table>

English correspondents of initials 子 and 頌 suggest series is alveolar sibilant.
Value of initial [z] overlaps with that of initial [t] in 正齒 “full tooth” series.
Since [z] is not an alveolar sibilant, and there is only one instance of this initial, we assume that it is in error, and reconstruct initial 頌 as voiceless alveolar affricate [ts], in line with the alveolar sibilant properties of other initials in the series.

Reconstructed Initials:

<table>
<thead>
<tr>
<th>Initial</th>
<th>久</th>
<th>齊</th>
<th>喜</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reconstruction</td>
<td>[dz]</td>
<td>(53.6%) or [ts]</td>
<td>(100%)</td>
</tr>
</tbody>
</table>

Series reconstructed as palato-alveolar 久 [dʒ]・借 [tʃ]・喜 [ʃ]・喜 [ʃ]
Data shows that alveolars 子 [z] and 頌 [s] may be palatalized before [i] and merge with 久 [dʒ] and 喜 [ʃ].
Merger of [zi] (fricative) and [dʒ] (affricate) suggests that phonetic value of 子 may actually be affricate [dz] rather than fricative [z] ȉȉ with makes more symmetrical in the 齒頭(尖) “tooth tip (apical)” series.

Reconstruction of Medials (1): Zero initial Onset

<table>
<thead>
<tr>
<th>齊口</th>
<th>齊口</th>
<th>齊口</th>
<th>齊口</th>
</tr>
</thead>
<tbody>
<tr>
<td>(j)</td>
<td>85.3% begin with [j], [i], [Unix]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(w)</td>
<td>96.9% begin with [w], [u], [Unix]</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(y)</td>
<td>98.1% begin with other than the above vowels and approximants</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Distribution:
- Used with mid to high vowels [i], [I], [u], [Unix] or empty filler vowel
- Only follows palato-alveolars [dʒ], [ʃ], [ʃ], [s], [s]|

Phonetic detail
- Adjacent consonants are [+palatal], [+round]
- Adjacent vowels are [+high]

Reconstruction
- Reconstructed as [+round, +high, +front] vowel [y]

Reconstructed Medials

<table>
<thead>
<tr>
<th>medial</th>
<th>齊</th>
<th>齊</th>
<th>齊</th>
<th>齊</th>
</tr>
</thead>
<tbody>
<tr>
<td>(j)</td>
<td>(i)</td>
<td>(w)</td>
<td>(v)</td>
<td></td>
</tr>
</tbody>
</table>
Reconstruction of Finals:

**Monophthongal (1)**

- 衣 reconstructed as [i]
  - 95% used for English [ɪ], [i] and [ɪ]
  - Consistent with 齊齒 label

- 鳥 reconstructed as [u]
  - 48% used for English [û] and [ʊ] (NOTE: another 49% used as filler)
  - Consistent with 合口 label

-迂 reconstructed as [y]
  - 舌面 label demands presence of vowel [y]
  - Used only for English syllables with rounded palatal consonants and high to mid-high vowels

**Monophthongal (2)**

- 啊 reconstructed as [e]
  - Compromise between [a] (25.6%), [æ] (24.8%) and [a] (21.7%)

- 嘿 reconstructed as [o]
  - Compromise between [a] and [au]
  - Neutral vowel --- used often as filler vowel (71.3%)

- 秆 reconstructed as [i]*
  - Highest percentage: 開口 [e] (39.2%); 合口 [we] (25%)

**Monophthongal (3)**

- 喔 reconstructed as [o]
  - Used often for neutral vowel [a] in the environment of labial/rounded consonants
  - (47.4%), e.g., in syllable “ber” in September, October, November, December — lends a rounded quality to the vowel
  - Also used for [a] (15.8%), [ɔ] (10.5%) and [ou] (10.5%)

-日 reconstructed as [ʊ]
  - Follows consonants [s] (74.6%), [z] (16.1%), and [ŋ] (7.7%) in absence of vowel

**Diphthongal**

- 哀 reconstructed as [ai]
  - 開口 variant is 100% used for English [ai]; 合口 variant is 50% [aw]

- 嘈 reconstructed as [au]
  - 開口 variant (except for fillers) is always used for English [au]

- 諂 reconstructed as [ei]*
  - 60% used for English [ei]; 32% used for English [ı]

- 牠 reconstructed as [ou]
  - 32.2% used for English [ou]; 25.8% used for English [ı]

- 叀 reconstructed as [i̯]
  - 51% used for English [ı]; 19.2% used for English [l]
  - Reconstructed as compromise between [ı] and [l]

**Nasal Endings (1)**

- 英 reconstructed as [in]
  - Highest percentage: 85.7% used for English [ıŋ]
  - Consistent with 齊齒 label

- 因 reconstructed as [in]
  - Highest percentage: 56.3% used for English [ın] or [ın]
  - Remaining 43.7% used for English [ın], which is already represented by initial 英
  - Consistent with 齊齒 label

- 毋 reconstructed as [yn]
  - 舌面 label demands presence of vowel [y]
  - Used only for English syllables with rounded palatal consonants, followed by [ın], [un] or [än]

**Nasal Endings (2)**

- 安 reconstructed as [en]
  - Compromise between [ın] (41% 50%), [än] (41%) and [ın] (9.1%)

- 應 reconstructed as [en]
  - Highest percentage: 36.4% used for English [ın] (齊齒)
  - Compromise of [ın] (36.4%), [än] (18.2%) and [ın] (9.1%)

- 懸 reconstructed as [än]
  - Compromise of [ın] (61.6%) and [än] (19.2%)
Reconstruction of Finals: Nasal Endings (3)

- 韏 
  - reconstructed as [an]*
  - Compromise of [an] (33.3%), [un] (33.3%), [am] (16.7%) and [sun] (16.7%)
- 韋 
  - reconstructed as [on]*
  - Compromise of [an] (25%), [enan] (25%), [oun] (25%) and [sun] (25%)
- 韄 
  - reconstructed as [an]*
  - 100% used for English [an]
  - Reconstructed as [an], with a slightly lower and more retracted vowel, to avoid overlap with initial [un]

Vowel Chart (13 vowels)

Reconstructed Finals (1)

<table>
<thead>
<tr>
<th>Exemplar</th>
<th>開</th>
<th>齊</th>
<th>合</th>
<th>撮</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>e</td>
<td>je</td>
<td>we</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>o</td>
<td>wo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>j</td>
<td>j</td>
<td>w</td>
<td>j</td>
</tr>
<tr>
<td>5</td>
<td>u</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>aj</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>au</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>a</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>ai</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>ou</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>ou</td>
<td>j</td>
<td>w</td>
<td></td>
</tr>
</tbody>
</table>

Reconstructed Finals (2)

<table>
<thead>
<tr>
<th>Exemplar</th>
<th>開</th>
<th>齊</th>
<th>合</th>
<th>撮</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>e</td>
<td>je</td>
<td>we</td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>en</td>
<td>wen</td>
<td></td>
<td></td>
</tr>
<tr>
<td>14</td>
<td>jn</td>
<td>yen</td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td>an</td>
<td>wan</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>in</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>jn</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>an</td>
<td>jan</td>
<td>waj</td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>jon</td>
<td>won</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>an</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>i</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>ar</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Comparison with Present-Day Mandarin (Initials)

<table>
<thead>
<tr>
<th>微</th>
<th>韋</th>
<th>韁</th>
<th>韇</th>
<th>韴</th>
<th>韘</th>
<th>韙</th>
<th>韚</th>
<th>韛</th>
<th>韜</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>b</td>
<td>p</td>
<td>m</td>
<td>f</td>
<td>p</td>
<td>m</td>
<td>f</td>
<td>p</td>
<td>pȹ</td>
</tr>
<tr>
<td>13</td>
<td>d</td>
<td>t</td>
<td>n</td>
<td>Ȏ</td>
<td>t</td>
<td>n</td>
<td>l</td>
<td>t</td>
<td>tȹ</td>
</tr>
<tr>
<td>14</td>
<td>dz</td>
<td>ts</td>
<td>s</td>
<td>ts</td>
<td>tsȹ</td>
<td>s</td>
<td>s</td>
<td>s</td>
<td>sȹ</td>
</tr>
<tr>
<td>15</td>
<td>dȢ</td>
<td>tȓ</td>
<td>ȓ</td>
<td>tǥ</td>
<td>tǥȹ</td>
<td>Ț</td>
<td>Ț</td>
<td>tȒ</td>
<td>tȒȹ</td>
</tr>
<tr>
<td>16</td>
<td>g</td>
<td>k</td>
<td>k</td>
<td>kȹ</td>
<td>g</td>
<td>k</td>
<td>k</td>
<td>k</td>
<td>kȹ</td>
</tr>
<tr>
<td>17</td>
<td>h</td>
<td>Ȉ</td>
<td>x</td>
<td>Ȉ</td>
<td>h</td>
<td>Ȉ</td>
<td>x</td>
<td>Ȉ</td>
<td>Ȉ</td>
</tr>
</tbody>
</table>

Comparison with Present-Day Mandarin (Initials)

- Accurately reconstructed: 31.8%
- Phonation
  - Loss of aspiration contrast (reconstructed as voicing contrast)
- Place of articulation
  - Alveopalatals reconstructed as palato-alveolars
  - Velar fricative reconstructed as pharyngeal
  - Lateral [l] reconstructed as tap [Ȏ]
  - (Successful reconstruction of retroflex series due to distributional symmetry and syllable structure constraints)
Comparison with Present-Day Mandarin (Medials)

<table>
<thead>
<tr>
<th>Reconstructed</th>
<th>Actual</th>
</tr>
</thead>
<tbody>
<tr>
<td>[f]</td>
<td>[f]</td>
</tr>
<tr>
<td>[i]</td>
<td>[i]</td>
</tr>
<tr>
<td>[w]</td>
<td>[w]</td>
</tr>
<tr>
<td>[y]</td>
<td>[y]</td>
</tr>
</tbody>
</table>

- reconstruction successful due to small inventory, but requires insight into:
  - Phonetic detail (redundant features):
    - e.g., reconstruction of [y]
  - Distributional symmetry

Comparison with Present-Day Mandarin (Rimes 1)

<table>
<thead>
<tr>
<th>Reconstruction</th>
<th>Modern Standard Chinese</th>
</tr>
</thead>
<tbody>
<tr>
<td>開 齊 合 撮</td>
<td>啊 a ja wa</td>
</tr>
<tr>
<td>開 齊 合 撮</td>
<td>喔 o wo</td>
</tr>
<tr>
<td>開 齐 合 撮</td>
<td>喔 o wo</td>
</tr>
<tr>
<td>開 齐 合 撮</td>
<td>啊 a ja wa</td>
</tr>
</tbody>
</table>

Comparison with Present-Day Mandarin (Rimes 2)

<table>
<thead>
<tr>
<th>Reconstruction</th>
<th>Modern Standard Chinese</th>
</tr>
</thead>
<tbody>
<tr>
<td>開 齐 合 撮</td>
<td>問 an wan</td>
</tr>
<tr>
<td>開 齐 合 撮</td>
<td>煙 jɑn ɡɔn</td>
</tr>
<tr>
<td>開 齐 合 撮</td>
<td>恩 Ǩn wǨn</td>
</tr>
<tr>
<td>開 齐 合 撮</td>
<td>因 in</td>
</tr>
<tr>
<td>開 齐 合 撮</td>
<td>晕 yn</td>
</tr>
<tr>
<td>開 齐 合 撮</td>
<td>閤 Ǩŋ jǨŋ wǨŋ</td>
</tr>
<tr>
<td>開 齐 合 撮</td>
<td>嗯 jʊn w珺</td>
</tr>
</tbody>
</table>

Comparison with Present-Day Mandarin (Rimes)

- 52.9% accurately reconstructed; 26.5% with crucial inaccuracies
- Rimes with ending [ŋ] reconstructed as rimes with ending [n], but vowel quality preserved.
  - Scarcity of [ŋ] ending plus back vowel combinations in English
- Rimes [ɛ] and [ei] flip-flopped
- Vowel quality differences
  - [a] reconstructed as [Ǡ] in rimes [a] and 安 an
  - [y] reconstructed as [Ȟ] in rime [y]
  - [i] reconstructed as [ui] in rime [i]

Methodological Implications (1)

- Phoneme to phoneme mappings -- difficult to reconstruct phonemes and distinction absent or rare in language of comparison
  - Alveopalatals [tɕ], [tʃᵢ], [ɕᵢ]
  - Velar fricative [x]
  - Rimes [an], [an] and [Nŋ]

Methodological Implications (2)

- Phonetic detail -- reconstruction of key distinctions require access to phonetic detail (redundant features):
  - Lip rounding in English [dʒ], [tʃ], [ʃ], [s], [z] led to reconstruction of medial [y] and finals [y], [ym]
- Distributional symmetry leads to more accurate reconstruction
  - Initial 齐 reconstructed as retroflex [j] due to consideration of within-category symmetry and syllable structure
Methodological Implications (3)

- Imperfect learning (L2 transfer) on the part of the glossator may lead to inaccurate reconstructions (Lass 1997: 83)
- English word-initial voicing contrast re-interpreted as Mandarin aspiration contrast → Mandarin aspiration contrast reconstructed as voicing contrast
- Inaccurate rendition of initial [l] because [l] used to transcribe English [ɨ]
- How much more perfect are the foreign language abilities of historical authors compared with present-day glossators (who have access to the foreign language on a daily basis thanks to globalization and mass media)?

Methodological Implications (4)

- Rime as single unit
- Reference to vowel quality and endings scarce
- Cause of reconstruction of finals with [ŋ] as [n]

Application to Chinese Historical Phonology

Proliferation of rhyme dictionaries

- Yan Zhitui (531–590) on rhyme dictionaries of his day:
  自茲厥後，音韻鋒出，各有土風，遞向非笑，“指馬”之喻，未知孰是。
  (inconsistent descriptions attributed to regional variation)

Contemporary Rhyme Classifications

- 13 rhymes (十三輻)
- 18 rhymes (中華新韻)
- 20 rhymes (中華韻典)
- 22 rhymes (張允和)
- 22 rhymes (張鼎盛)

Areas of contention (inclusion of lip-rounding, common historical source)

- Rimes 唱 [ŋ] vs 啊 [a] (inclusion of lip-rounding)
- Rimes 糾 [z] vs 之 [ŋ] (common historical source)
- Rimes 安 [an] vs 煙 [en] (common historical source)
- Rimes 因 [in] vs 暈 [yn] (inclusion of lip-rounding)
- Rimes 極 [ŋ] vs 王 [in] (common historical source)
Interpreting Philological Records

1. Linguistic complexity:
   - "the Middle Chinese sound system that Karlgren reconstructed is too complex to be true" (Ao 1991: 338)

2. Language–external factors giving rise to complexity:
   - Sources represent not a real, living homogeneous language, but "a maximally differentiated compromise between the reading traditions of [cultural centers of east central China], further complicated by the inclusion of distinctions attested in various earlier rhyme dictionaries". (Norman & Coblin 1995: 578)
   - Philological categories may not be phonetically unique
   - Complexity as byproduct of reconstruction method?
   - Variant rhyme classifications for MSC
   - Excessive number of [n]-ending rimes (8): [in], [yn], [sn], [in], [en], [an], [en], [en], [en], [en]

Future Research

- Reconstructions based on classificatory schemes in additional rhyme books
- Reconstructions based on English–Chinese correspondences in different tourist guides (look for regional variation)
- Reconstructions based on correspondences with languages other than English (e.g., Japanese, Korean, German, French)

Chris Wen-Chao Li, San Francisco State University
- Tel.: 001 415 338 1034
- E-mail: wenchao@sfsu.edu