

# Making sense of Yuan-era transcriptions of non-Chinese words

Historiography of the Mongol Yuan Empire

# Outline

- ▶ The Target Language: Chinese historical phonology
- ▶ The Source Languages
- ▶ The Devices

# The Target Language: Chinese historical phonology

- ▶ How do we know how Chinese characters were once pronounced? (Also known as What are our sources for the phonological history of Chinese?)
  - ▶ Direct or semi-direct attestation
    - ▶ Transcription into alphabetic scripts: Takes you back to Tang
    - ▶ Chinese transcriptions of known non-Chinese terms: Take you back to Han
    - ▶ Foreign readings of Chinese characters: Takes you back to Tang
  - ▶ Reconstructed attestation
    - ▶ Rhyming dictionaries: Takes you back to Han
    - ▶ Internal reconstruction from existing Chinese dialects: Takes you back to Tang
    - ▶ Character forms: Takes you back to Zhou
    - ▶ Tibeto-Burman comparison: Takes you back to prehistory

# Studying Chinese historical phonology: transcription into alphabetic scripts

- ▶ Transcription may be popular (limited to target language phonetic resources) or academic (using special conventions to match source language phonetics)
- ▶ Transcriptions may be descriptive (based on one spoken dialect) or synthetic (aiming to synthesize a variety of dialects)
  - ▶ As a rule, only academic systems can be synthetic
- ▶ Some major systems
  - ▶ Pinyin: academic, Putonghua (Beijing-based), 1955-1958
  - ▶ Wade-Giles: academic, synthetic (approximates Beijing), developed 1859-1892
  - ▶ École française d'Extrême-Orient (EFEO) system: academic, synthetic (approximates Nanjing), 1700s-1902
  - ▶ Manchu system: academic, ?, Beijing Mandarin? 1632
  - ▶ Pagba Chinese: academic, ?, ?, 1271
  - ▶ Dunhuang Tibetan: popular, Shazhou (Dunhuang), 800-900

# Studying Chinese historical phonology: Chinese transcriptions of known non-Chinese terms

- ▶ Major methodological hurdle: you already have to know something about ancient Chinese pronunciation to match it with attested Chinese transcriptions
- ▶ Again as with Chinese transcriptions, academic vs. popular, descriptive vs. synthetic
- ▶ By far the most important single body of this data is Chinese transcriptions of Sanskrit dharanis and other Buddhist terms
  - ▶ Studied most productively by W. South Coblin
  - ▶ Older ones N-S Dynasties
  - ▶ More recent ones fixed in Tang era
- ▶ Some geographical terminology goes back to Han dynasty:
  - ▶ Example Alexandria as *Wu-yi-shan-li* 烏弋山離 >> A-ye-shan-liai

# Studying Chinese historical phonology: Foreign readings of Chinese characters

- ▶ Three main types
  - ▶ Sino-Japanese *kanji* 漢字: mostly Chang'an Tang, some Southern dynasties Wu
  - ▶ Sino-Vietnamese *chữ Hán* 字漢: mostly Tang (?)
  - ▶ Sino-Korean *hanja* 漢字: mostly Five Dynasties-Northern Song (?)
- ▶ In all three cases, pronunciation shaped by phonetics of receiving language at time of reception
  - ▶ For example, ancient Japanese had no *ts*, rendered all Chinese *ts* as *s*
- ▶ And by subsequent phonetic evolution in receiving language
  - ▶ For example, Japanese *p* > *h*, *f*, or *tu* > *tsu*
  - ▶ For example, Korean *ly*, *ny* > *y*

# Studying Chinese historical phonology: The Reconstruction methods

- ▶ Rhyming dictionaries: Takes you back to Han
- ▶ Internal reconstruction from existing Chinese dialects: Takes you back to Tang
- ▶ Character forms: Takes you back to Zhou
- ▶ Tibeto-Burman comparison: Takes you back to prehistory
- ▶ “Use a star when you say that, pilgrim”



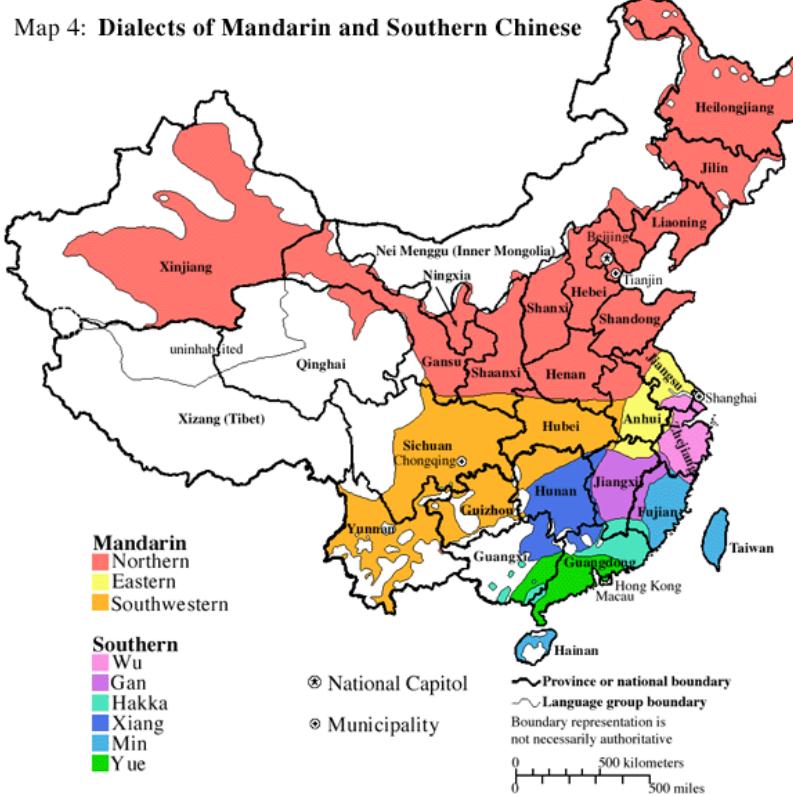
For example: *jian* 盱 < \*klam

# Studying Chinese historical phonology: Rhyming Dictionaries

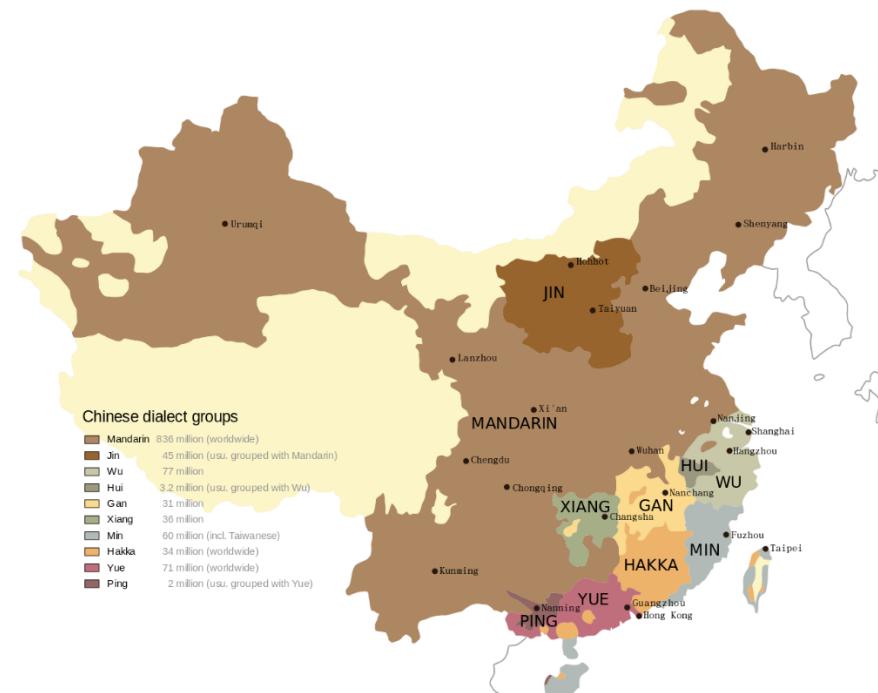
- ▶ This the dominant body of data in Chinese linguistic reconstruction
- ▶ Rhyming dictionaries establish *classes*
- ▶ The linguist then reconstructs what those classes might have been, phonetically
- ▶ Kalgren introduced this to Sinology
- ▶ Pulleyblank is the latest summarizer of this data
- ▶ Rhyming dictionary data is:
  - ▶ Pervasively synthetic
    - ▶ Both synchronically (by space)—tries to create a standard that would make sense of all educated Chinese persons' speech
      - ▶ “Three literati from Hangzhou, Nanjing, and Beijing walk into a moon-viewing party . . .”
    - ▶ And diachronically (by time)—tries to fit contemporary speech into historic speech patterns
  - ▶ As a result, it's always more complicated than seems realistic

# Studying Chinese historical phonology: Internal reconstruction from existing Chinese dialects (oops—Sinitic languages!)

Dartmouth's dialect map

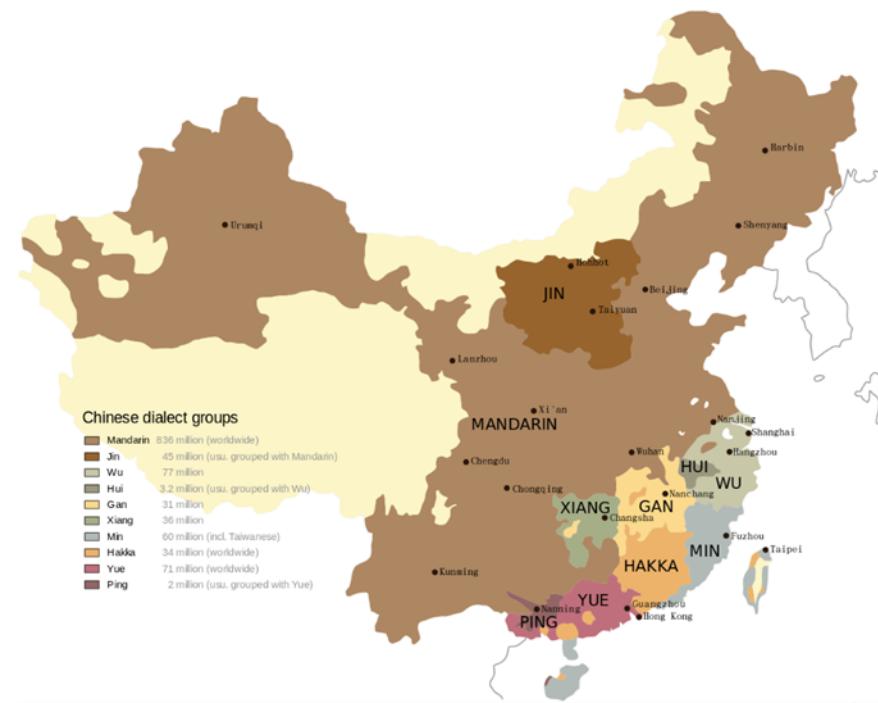


PRC's dialect map



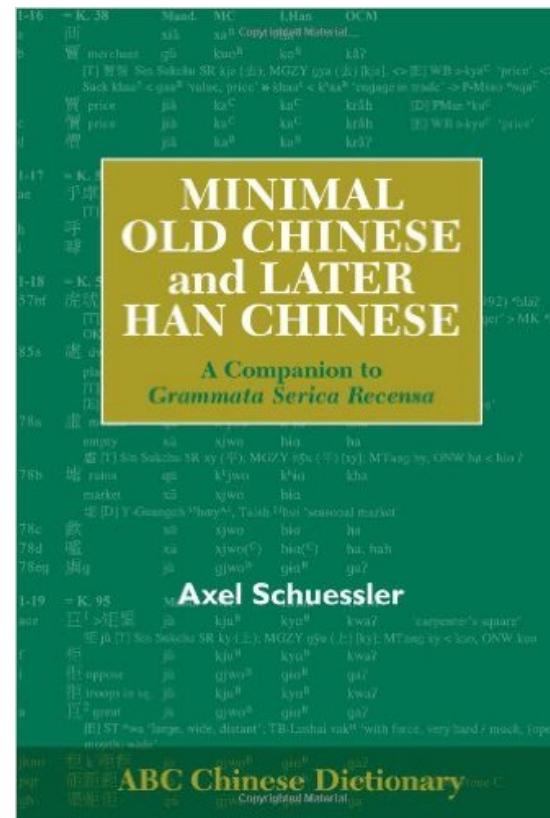
# Studying Chinese historical phonology: Internal reconstruction from Sinitic languages

- ▶ Two ways of look at this:
  - ▶ “conservative” regions vs. “progressive” regions
    - ▶ Implies conservative regions are speaking a language “just like” some past dynasty’s
    - ▶ At some point in the past (usually identified as Tang) Sinitic language was homogenous
  - ▶ Evolving family, with different isoglosses and common trends
- ▶ Hardly any examples of non-prestige dialects being used as target language for transcriptions



# Studying Chinese historical phonology: Character forms & Tibeto-Burman

- ▶ Neither relevant for our work
- ▶ But both are reconstructive and *highly* controversial
- ▶ Best guide (for my money): Axel Schuessler's *Minimal Old Chinese and Later Han Chinese*



# Application: Yuan-era Pronunciation of Chinese Characters

- ▶ Two main sources
  - ▶ *Menggu ziyun* 蒙古字韻
    - ▶ The first academic transcription of Chinese—ever
  - ▶ *Zhongyuan yinyun* 中原音韻
    - ▶ A rhyming dictionary—but one that appears to be shockingly descriptive
      - ▶ Note: it actually has a dialect in the title
- ▶ Debate: to what degree is *Menggu ziyun* synthetic?
- ▶ In my experience, the target dialect for transcriptions is accurately represented by *Zhongyuan yinyun*

# Application: Yuan-era Divergences from modern Putonghua

- ▶ /j/ < /dz/ or /g/
  - ▶ /q/ < /ts/ or /k/
  - ▶ /x/ < /s/ or /h/
  - ▶ /-n/ < /-n/ or /-m/
  - ▶ /-e/ < /o/ or /-au/
  - ▶ /en/ < /in/, /eng/ < /ing/
  - ▶ /uan/ < /on/, /uang/ < /ong/
  - ▶ /ong/ < /ung/
  - ▶ /ie/ < /iai/
  - ▶ /ue/ < /io/
  - ▶ /au/ < /au/ or /eu/
  - ▶ /ou/ < /əu/
- 
- ▶ NB: These are quite simplified reconstructions
  - ▶ A number of them still found in later transcription systems
    - ▶ E.g. Yuezhou 岳州 > Yojeu
    - ▶ E.g. Jiangxi 江西 > Giyangsi
  - ▶ But also note: later transcriptions have archaic elements not seen in (Daidu) Yuan Mandarin
    - ▶ E.g. Ngan-king 安慶

# Application: Yuan-era Divergences from modern Putonghua

- ▶ Note also:
- ▶ There are a number of irregular developments
- ▶ i.e. words jump from one class to another
  - ▶ E.g. Shuò 撥; cf. SHWAW 瓢; cf. CHAU<sup>3a</sup> 瓢 (p. 138); cho~chol (+<sup>丁</sup>) (§§424, 505); cho~chö (p. 36)
  - ▶ Shuò 碩; cf. alternative Mandarin shí; cf. 碩德八刺 < Tib. Siddhi-pa-la [Šidiibala]

# The Source Languages

- ▶ What are the potential languages which can be source language of transcriptions?
- ▶ Conventional wisdom:
  - ▶ Mongolian was officially used, but really Persian and/or Turkic was the non-Chinese language the immigrants were all speaking
- ▶ I'm skeptical: Why?
- ▶ Both Persian and Turkic have the phoneme /z/; Mongolian does not
  - ▶ Every once in a while you see a transcription where /z/ is being represented
  - ▶ But it's quite rare
  - ▶ So, Mongolian appears to be the dominant language (but note symbiosis with eastern Turkic dialects, Uyghur & Öng'üt)

# The Source Languages

- ▶ Mongolian in the Uyghur script dominant
- ▶ Also Mongolian in the Pagba Script
- ▶ Uyghur & Öng'üt Turkic
- ▶ Persian
- ▶ Jurchen
- ▶ Note: all of these also found in the *Hua-Yi yiyu* 華夷譯語 vocabularies from the following Ming dynasty, and other vocabularies

# The Source Languages: Defining Source Language Vocable sets

- ▶ That is, what dictionaries should you use?
- ▶ But since transcriptions are mostly names, name dictionaries particularly useful, as are atlases with alphabetical indexes
- ▶ Biggest desideratum is a Mongolian name dictionary that is not keyed to just modern Mongolian usage
- ▶ Like *Onomasticon Turicum*

# The Source Languages: Peculiarities of Uyghur (and) Mongolian

- ▶ Vowel Harmony
- ▶ Suffixation
- ▶ Phonotactics (syllable structure, vowels, consonants)
- ▶ Strong & Weak distinctions and lenition

# Vowel Harmony

- ▶ What is it? Division of vowels into classes, such that any given vocable has vowels only from one class
- ▶ This extends to both derivational and agglutinative suffixes
- ▶ Lexical assignment to the vowel-harmonic classes is *extremely stable*
- ▶ Yuan-era Chinese academic transcribers follow vowel harmony, but not in rounded vowels
- ▶ In Mongolian, /i/ is neutral
- ▶ Uyghur has /ɨ/ (masculine) and /i/ (feminine)

Masculine/“Back” /Velar	Feminine/“Front” /Palatal
/a/	/e/
Rounded vowels	
/o/	/ö/
/u/	/ü/
Consonantal allophones	
/q/	/k/
/g/ (γ, Γ or ǵ, Ǵ)	/g/

# Vowel Harmony in Chinese transcriptions

Question: what vowel harmony does 干 represent?

Masculine/“Back” /Velar	Feminine/“Front” /Palatal	Masculine/“Back” /Velar	Feminine/“Front” /Palatal
/a/	/e/	納	捏
Rounded vowels		Rounded vowels	
/o/	/ö/	那, 諾	那
/u/	/ü/	奴, 詮	奴
Consonantal allophones		Consonantal allophones	
/q/	/k/	哈, 合	怯, 克, 可
/g/ ( $\gamma$ , Γ or $\dot{g}$ , $\dot{G}$ )	/g/	哈, 合	哥

# Vowel Harmony in Chinese transcriptions: Suffixation

- ▶ Vowel Harmony determined by first syllable
  - ▶ E.g. 朱兒徹台
  - ▶ E.g. 月哥察兒
  - ▶ E.g. 禿滿 for tümen
- ▶ Derivational and agglutinative suffixes all have two forms: masculine and feminine
- ▶ Chinese transcriptions *frequently* use just the masculine form for all such suffixes.
- ▶ Sometimes even for just second syllables
  - ▶ May be connected to features of Mongolian phonetic realization

# Middle Mongolian Phonotactics

- ▶ syllable structure
  - ▶ CV or CVC
  - ▶ Initially C can be an unrepresented (virtually silent) glottal stop
- ▶ Vowels
  - ▶ /o/, /ö/ in second syllable only allowed following another /o/, /ö/
- ▶ Consonants
  - ▶ In syllable-final position, no strong/weak distinction, no affricates allowed
    - ▶ In other words no t-d, K-G (what's this? q-ğ, k-g), p-b minimal pairs, no final č, no final ĥ
    - ▶ Early convention was to transcribe them as strong, now as weak

# Strong & Weak distinctions and lenition

- ▶ Most languages we work with have a two-way strong-weak distinction
- ▶ But three-way, even four-way distinctions found
- ▶ *BUT*

Strong	Weak
/k/	/g/
/q/	/g̚/
/t/	/d/
/č/	/j/
/ts/	/dz/
/p/	/b/

# Strong & Weak distinctions and lenition

- ▶ Most languages we work with have a two-way strong-weak distinction
- ▶ But three-way, even four-way distinctions found
- ▶ BUT how they are realized differs broadly
- ▶ Two simple version:
  - ▶ Strong: unvoiced, weak: voiced
    - ▶ Continental European, Middle Eastern languages, Japanese
  - ▶ Strong: aspirate, weak: unaspirated
    - ▶ Mandarin, modern Tibetan
- ▶ Mixed versions
  - ▶ Modern English, Middle Mongolian
- ▶ Mandarin speakers tend to hear all intervocalic Mongolian plosives as weak
  - ▶ Weakening is called “lenition”

Aspirate	Unaspirate	Voiced
[t'] or [t <sup>h</sup> ]	[t]	[d]
	Italian t Persian t	Italian d Persian d
Pinyin t	Pinyin d	
English t (initial)		English d (initial)
Mongolian t (initial)	Mongolian d (initial)	
	Mongolian t (intervocalic)	Mongolian d (intervocalic)

# The Devices

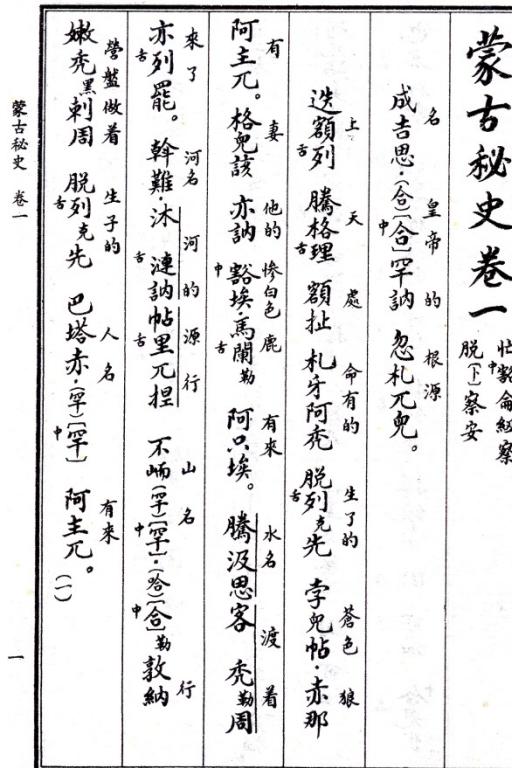
- ▶ The standard character set
- ▶ Diacriticals
- ▶ How were systems transmitted?
- ▶ The concision tendency
- ▶ The logographic tendency

# The standard character set

- ▶ Each dynasty uses a somewhat standardized character set for transcribing sounds
  - ▶ Partly dependent on sound evolution
  - ▶ Partly dependent on conventions
- ▶ Each dynasty, based on the dominant source language will select certain phonetic features in target language to stress and ignore others
  - ▶ For example, let's transcribe *küsen*
    - ▶ Early/mid-Yuan character set: 曲先
    - ▶ Qing character set: 庫森
  - ▶ For example, let's transcribe Mongolian *sula* “vacant, unemployed”
    - ▶ Yuan character set: 速刺
    - ▶ Qing character set: 蘇拉
  - ▶ Contrast
    - ▶ Jurchen Jin: 薩合釐 Sahalian
    - ▶ Yuan:曲憐居 or 起釐谷 Kürelgü

# Diacriticals

- ▶ For Ming era transcriptions in *Hua-Yi yiyu* 華夷譯語 and *Menggu mishis* 蒙古秘史, systems with diacriticals preserved
    - ▶ Final non-nasal consonants (b, t, k)
    - ▶ Distinguish r and l
    - ▶ Distinguish q~g from h
    - ▶ Linked to academic transcriptions of full texts



# How were the practices transmitted?

- ▶ Were there standard reference works?
- ▶ Survive from the Ming, what about in the Yuan?
- ▶ Or was it based on an apprenticeship situation?

# The Concision Tendency

- ▶ The aim is reduce the number of syllables
- ▶ Ways to do this
  - ▶ Omit final non-nasal consonants (esp. k, t, b, but sometimes r or s)
  - ▶ Represent final liquids (r or l) by -n
  - ▶ Use final nasal, but include reduced-size diacritical character to represent the exact consonant

# The logographic tendency

- ▶ Common roots understood by transcribing scribes
- ▶ When making derivations, they would try to preserve the root
- ▶ E.g. tümen 禿滿
  - ▶ Tümeder 禿滿迭兒
  - ▶ Why is *man* 滿 transcribing /-me-/?