

© Jeroen Wiedenhof 2005

Published in 李哲賢 / Lee Jer-shiarn (ed. in chief), 2004 漢學研究國際學術研討會論文集 [Proceedings of the International Conference on Chinese Studies 2004], pp. 387-402. 斗六/ Touliu: 雲林科技大學 / National Yunlin University of Science and Technology, 2005. ISBN 9860040117.

Purpose and effect in the transcription of Mandarin

Jeroen Wiedenhof, Leiden University

Contents

- 0. Sinology and international sinology
 - 1. The transcription of Chinese
 - 1.1. International developments
 - 1.2. Developments in China and Taiwan
 - 2. The Pīnyīn era
 - 2.1. The linguistic impact of Pīnyīn
 - 2.2. The perception of spoken Mandarin
 - 2.3. The transcription of tone sandhi
 - 3. Conclusions
 - Glosses
 - References
-

0. Sinology and international sinology

If sinology is defined as the study of China, international sinology represents the synergy between Chinese and non-Chinese scholarly traditions in this field. The distinction between sinology and international sinology, essential within Chinese academe, is trivial outside China, the international dimension being self-evident. The methodological implications will not be pursued here, as the purpose of the present paper is linguistic. Note, however, that similar issues haunt the use and application of the terms *Chinese linguistics* (a. ‘the study of Chinese languages’, b. ‘linguistics in China’) and *Chinese languages* (a. ‘Chinese languages, Sinitic languages’, b. ‘the languages of China’, including e.g. Mongolian, Uyghur and Zhuàng).

Cooperation between Chinese and non-Chinese traditions in sinology depends on a mutual understanding of differences in purposes and methods. The following sections document such variation in an area central to sinology, viz., the ways in which the Mandarin language is represented in writing. The discrepancies between the Chinese script and the Latin alphabet are consistent with systematic differences in linguistic perceptions and attitudes. Likewise, within alphabetic traditions, the choice of a specific romanization system may facilitate the discussion of some questions in the study of language while obscuring others.

After the rise and fall of a number of romanizations, the Pīnyīn system has gained international acceptance since the final decades of the twentieth century. In this period of relative stability in the transcription of Mandarin, the time has come to take stock.

1. The transcription of Chinese

Systematic documentation of Chinese languages in the Latin alphabet dates back to the sixteenth century, when religious zeal drove missionaries to the shores of southern China. Following the vicissitudes of Sino-foreign relations, the needs of religious, political and economic activities led to linguistic attention to various Mǐn 閩 and Yuè 粵 dialects. And from the earliest missionary contacts, the need for social and political acceptance led to a keen interest in the language of the Chinese official class, or *guānhuà* 官話 ‘Mandarin’.

Figure 1 – Early romanized Mandarin

譯 引 首 普	辛 母 五 德 oe	九 子 齒 阿 oa	大 字 三 音 in	毛 二 三 應 im	夫 自 士 魚 iu	○ 同 鳴 字 父
ce	ce	cin	cin	cin	cin	cin
coe	coa	coa	chin	chin	chin	chin
coe	coa	coa	chin	chin	chin	chin
Koe	Koa	Koa	kin	kin	kin	kin
Koe	Koa	Koa	kin	kin	kin	kin
poe	poa	poa	pin	pin	pin	pin
poe	poa	poa	pin	pin	pin	pin
toe	toa	toa	tin	tin	tin	tin
toe	toa	toa	tin	tin	tin	tin
uo	ua	ua	un	un	un	un
uo	ua	ua	un	un	un	un

A fragment from the facsimile re-print of Niklaas Trigault's 1626 西儒耳目資 *Xīrú ěrmù zī* 'Occidental scholar's auditory and visual guide'.

Trigault (1957, part 1, p. 67)

1.1. International developments

The earliest transcriptions of Mandarin in the Latin alphabet reflected the orthographies of the transcribers' native languages. The Mandarin voiceless retroflex fricative [ʂ], for instance, was recorded as *ch* according to French systems, as *x* in Spanish transcriptions, and as *sh* in spellings based on English, e.g. *choû*, *xū* and *shu*¹ for Pīnyīn *shū* 書 'book'. The popularity of these transcriptions depended not only on the nationality of the user, but also on the availability and success of grammars and guide books such as Niklaas Trigault's 1626 西儒耳目資 *Xīrú ěrmù zī* 'Auditory and visual guide for the occidental scholar' (1957; see Figure 1), Francisco Varo's 1703 *Arte de la lengua mandarina* (2000) and Abel-Rémusat's 1822 *Éléments de la grammaire chinoise* (1857).

By the beginning of the twentieth century, foreign scholars had recreated the sounds of Mandarin in a large number of transcriptions. At the same time, a proliferation of scripts

was beginning to evolve in China (section 1.2). Different spellings were usually associated with national, regional and scholarly traditions; Legeza's (1968-1969) *Guide to transliterated Chinese in the modern Peking dialect* lists details for fifty major systems.¹

Due to the growing global importance of English as a language of science, English-based transcriptions were increasingly accepted in international communication, particularly the Wade-Giles and Yale systems (Table 2). Wade-Giles is the name given to the system used by Herbert Allen Giles (1845-1935) in his Chinese-English dictionary (1892, 1912). This spelling was based on the system created by Thomas Francis Wade (1818-1895), although Giles did not specify his adjustments. Today, the popularity of Wade-Giles is second only to that of the Pīnyīn spelling. The most common version of Wade-Giles is the reasoned adaptation by Yuen Ren Chao (趙元任 Zhào Yuánrèn, 1892-1982) in R.H. Mathews' Chinese-English dictionary (1943: ix-xxi).

Table 2 – Three popular modern romanizations for Peking Mandarin

Note the different treatment of (a) aspiration; (b) retroflex, palatal and alveodental initials; and (c) syllable boundaries. For details on the transcription of tone sandhi see section 2.3.

Wade-Giles (1892)	Yale (1943)	Pīnyīn (1958)	meaning, characters
(a) <i>tá</i> ³	<i>dǎ</i>	<i>dǎ</i>	'hit' 打
<i>t'a</i> ³	<i>tǎ</i>	<i>tǎ</i>	'pagoda' 塔
(b) <i>chuan</i> ⁴	<i>jwàn</i>	<i>zhuàn</i>	'earn' 賺
<i>chüan</i> ⁴	<i>jywàn</i>	<i>juàn</i>	'chapter' 卷
<i>tsuan</i> ⁴	<i>dzwàn</i>	<i>zuàn</i>	'drill' 鑽
(c) <i>Hsiao²-chieh, ní² hao³?</i>	<i>Syáujye, ní² hǎo?</i>	<i>Xiáojie, ní² hǎo?</i>	'How are you, Miss?' 小姐, 你好?

By the middle of the twentieth century, the only international rival to Wade-Giles was the Yale transcription. Devised in 1943 by George A. Kennedy (1901-1960), the system owed much of its popularity to the American war effort, which created a huge demand for applied linguistic skills. When the army field manuals of the 1940s were reworked into college textbooks, the Yale transcription made an international breakthrough. From the 1950s to the 1970s, Yale University's Far Eastern Publications provided Mandarin curricula throughout the Western world with textbooks, dictionaries and language manuals. Generations of sinologists in the United States and in Europe were initiated to the sounds of Mandarin by means of the Yale spelling, which combined the self-evident graphics of tone symbols (high level tone ˉ; high rising tone ˊ; low tone ˋ; high falling tone ˋ; unmarked neutral tone) with distinctly Anglo-Saxon correspondences between graphics and sound values (e.g. *shr, sye, tsz* for Wade-Giles *shih, hsieh, tz'ü*, Pīnyīn *shi, xie, ci*). The influence of

¹ The notion of transcriptional variety even took root in sinological folklore, as the unwritten *rites de passage* of early European sinology were said to include marrying a Chinese partner, conceiving a transcription for the Mandarin tongue, and publishing a translation of the 道德經 *Dào dé jīng*; preferably in that order, to coincide with the scholar's youth, acme and wise old age.

Yale, however, was restricted to the realm of language instruction, and Wade-Giles continued to dominate international research publications.

The creation and development of the *Hànyǔ Pīnyīn* 漢語拼音 transcription, usually abbreviated as Pīnyīn, will be outlined in section 2. In the Cold War era, the use of this system outside China was typically regarded as a political statement, or a deliberate identification with the Chinese communist regime. Like Yale, Pīnyīn entered the academic field by means of educational channels, initially through foreigners learning Mandarin in China and later through textbooks, dictionaries, and linguistic reference works from the People's Republic.

The main boost for the Pīnyīn system came with China's arrival on the international political scene. After the People's Republic of China became a member of the United Nations in 1971, Pīnyīn gradually gained acceptance outside China. In 1982, the international agency for standardization ISO registered Pīnyīn as International Standard #7098 (ISO 1982). Soon Pīnyīn spelling dominated popular journals, newspapers, translated literature and other texts dealing with China. This change is evident, for instance, in international cartography. According to the *Times atlas of the world*, "[e]arly in 1979 Pīnyīn was accepted by most nations of the world as the system to be employed officially for romanized Chinese names" (1983, index, p. 6). However, even today, an exception is usually made for the geography of Taiwan, which follows the Wade-Giles tradition (Figure 3).

Figure 3 – International transcription of Mandarin across the Taiwan Strait: Pīnyīn and Wade-Giles

Within the same plate, place names in China are spelled in Pīnyīn without tone symbols, e.g. “Quanzhou” for 泉州, Pīnyīn *Quánzhōu*, Wade Giles *Ch'üan²-chou¹*; but place names in Taiwan reflect the Wade-Giles system, again dispensing with tones: “Kao-hsiung” and “Kaohsiung” for 高雄, Wade Giles *Kao¹-hsiung²*, Pīnyīn *Gāoxióng*.



Times (1983, plate 23)



National Geographic (1999, plate 103)

In Taiwan itself, the growing international acceptance of Pīnyīn was officially ignored, but the system began to be taken seriously by Taiwanese linguists and dictionary makers in the 1990s. The transition was marked by alternative names such as 國際拼音 *Guójiè Pīnyīn* ‘International Transcription’, 聯合國拼音 *Liánhéguó Pīnyīn* ‘United Nations Transcription’, and, more colloquially, 羅馬拼音 *Luómǎ Pīnyīn* ‘Romanized Transcription’. Despite these developments, the use of Pīnyīn in Taiwan is marginal: it coexists with Gwoyeu Romatzyh, 通用拼音 *Tongyòng Pinyin*, Wade-Giles, Yale, as well as unsystematic English-based spellings. The Taiwanese situation is also heavily politicized, with central and local governments often implementing different transcriptions, or using a variety of systems. In the Taipei subway system, for instance, in-coach digital marquees consistently use Pīnyīn, while signboards in the stations reflect a variety of transcriptions. Hence passengers traveling to 頂溪 *Dǐngxī* on the southbound line may find the name of this station announced as *Dingxi* before arrival, but spelled as *Tinghsi* after alighting on the platform (Figure 4).

Figure 4 – Mandarin transcriptions in Taiwan

In the course of one trip, travellers on the Taipei subway may find the same station being announced in different transcriptions. This example shows Pīnyīn and Wade-Giles, without tones for both systems and without apostrophes and hyphens for Wade-Giles. The photos were taken in October 2004.



In-coach marquee

Platform signboard

The international success of Pīnyīn has led to the disappearance of various national transcriptions of Mandarin. While Wade-Giles continues to play a significant role in international sinology, the academic use of Spanish, French, German and other transcriptions for Mandarin has been on the wane. In the popular press, likewise, Pīnyīn now often replaces national systems of transcription. The Pīnyīn system has even led to cases of genuine linguistic change outside China, as long-established terms for Chinese proper names are sometimes exchanged for Pīnyīn reading pronunciations. Thus the *Bloomsbury English*

dictionary defines *Peking* [p^hi'k^hi:ŋ] as the “former name for *Beijing*” [bei'dzɿŋ], which is still listed in compounds such as *Peking duck* and *Peking man* (2004: 1386). Languages may display various degrees of influences of this type. Developments similar to those in English can be found, for instance, in modern Dutch; but in German, which likewise uses Pīnyīn in the press, spoken forms tend to remain faithful to traditional names such as *Mao Tse-tung* [mao ts^he:t^huŋ], *Peking* [p^he:k^hiŋ] and *Taipeh* [t^hai'p^he:].

The overall effect of the Pīnyīn system's international success has been one of standardization. Mandarin is the only living Chinese language for which a single transcription system has been almost universally accepted. For other major Chinese languages, such as Cantonese Yuè and Taiwanese Mǐn, competing transcriptions and competing scripts continue to generate confusion among students and disputes among scholars. But this standardization in the Pīnyīn era has its downsides, too. Like any written representation of a spoken language, Pīnyīn spelling tends to highlight some aspects of Mandarin while failing to draw attention to others. Examples will be presented in sections 2.1 to 2.3.

1.2. Developments in China and Taiwan

The modern Chinese linguistic tradition started at the close of the nineteenth century; the usual point of departure is the 1898 publication of 馬氏文通 *Mǎ Shì Wén Tōng* ‘Mr. Mǎ's interpretation of texts’, a Latin-inspired grammar of Classical Chinese written by 馬建忠 *Mǎ Jiànzhōng* (1845-1900). In the social and political upheaval marking the last years of the 清 *Qīng* empire and the formative years of the Chinese Republic, a number of daunting linguistic problems arose, including the position of Mandarin as a national language, the future of the character script, and the replacement of Classical Chinese by 白話 *báihuà* ‘modern written Chinese’ as the educational and literary standard. John DeFrancis (1950) gives an overview of this pivotal stage in the history of Chinese.

Many spelling systems were created in late imperial and early Republican China, but only two had wide-ranging success (Table 5). The first of these is 注音符號 *Zhùyīn Fúhào* (literally, ‘Phonetic Spelling Symbols’), which goes by a surprising number of alternative names, such as 國音字母 *Guóyīn Zìmǔ* ‘Alphabet for the National Pronunciation’, 國語注音符號 *Guóyǔ Zhùyīn Fúhào* ‘Phonetic Spelling Symbols for the National Language’, 注音字母 *Zhùyīn Zìmǔ* ‘Phonetic Spelling Alphabet’ and simply 注音 *Zhùyīn* ‘Phonetic Spelling’. Colloquially, the system is also called ㄅㄆㄇㄏ *Bōpomōfo*, after the names of the first four symbols in this scheme (Table 6). English names include *Mandarin Phonetic Symbols*, *National Phonetic Alphabet* and *National Phonetic Letters*. Among transcriptions still prominent today, *Zhùyīn Fúhào* is the only one which favors native symbols instead of the Latin alphabet. This feature was part of its birthright, for the native Japanese 假名 *kana* system served as a role model at a time when Chinese nationalist feelings prevailed. The shape of the symbols reflects Chinese calligraphic conventions (Table 6), and their arrangement in strings (Table 5) represents the traditional analysis of the Mandarin syllable into an initial, a medial, a final and a tone, each element spelled by a separate symbol, e.g. ㄓㄨㄢˋ, Pīnyīn *zh + u + an + ˋ > zhuàn* ‘earn’.

Table 5 – Three influential Chinese transcriptions for Peking Mandarin

Compare Table 2; for the treatment of tone sandhi see section 2.3.

	Zhùyīn Fúhào (1913)	Gwoyeu Romatzyh (1926)	Pīnyīn (1958)	meaning, characters
(a)	ㄉㄚˇ ㄊㄚˇ	daa taa	dǎ tǎ	‘hit’ 打 ‘pagoda’ 塔
(b)	ㄉㄨㄢˋ ㄐㄨㄢˋ ㄉㄨㄢˋ	juann jiuann tzuann	zhuàn juàn zuàn	‘earn’ 賺 ‘chapter’ 卷 ‘drill’ 鑽
(c)	ㄊㄩㄢˊ ㄉㄨㄢˊ ㄉㄨㄢˊ ㄉㄨㄢˊ ㄉㄨㄢˊ ㄉㄨㄢˊ	Shyau.jie, ni hao?	Xiáojié, ní hǎo?	‘How are you, Miss?’ 小姐, 你好?

The other prominent Chinese transcription at this time was Gwoyeu Romatzyh 國語羅馬字 (*Guóyǔ Luómǎ zì* ‘Romanization of the National Language’), abbreviated as *Guólúo* in Mandarin and as *GR* in English. The system was conceived by Y.R. Chao and endorsed by the Ministry of Education in 1926. GR is the only well-known system which spells tones without using diacritical symbols, thus steering clear of the technicalities of producing, storing and printing tone marks. Regrettably, however, it lacks unique solutions for each tone; only the neutral tone is consistently spelled with a dot preceding the spelling of the corresponding first tone syllable. The representation of the other tones may differ for different vowels. For instance, GR spells *.da, da, dar, daa, dah* for Pīnyīn *da, dā, dá, dǎ, dà*; but *.bau, bau, baur, bao, baw* for Pīnyīn *bao, bāo, báo, bǎo, bào*. In retrospect, GR might have stood a better chance at survival in the digital age, for in contrast to other transcriptions, it offers no complications for plain text systems and no distinctive computer font requirements.

Table 6 – Graphical characteristics of 注音符號 Zhùyīn Fúhào

The 注音符號 *Zhùyīn Fúhào* transcription makes use of traditional brushstrokes to create simple characteroid symbols consisting of maximally four strokes. It thus maintains Chinese calligraphical features while avoiding the impression of graphical complexity usually associated with character writing.

The first four symbols in the system, ㄅ *bō*, ㄆ *pō*, ㄇ *mō* and ㄉ *fō*, representing the labial consonants *b, p, m* and *f*.

Regular brush characters (楷書 *kǎishū* ‘Model Script’) with few strokes: *dāo* ‘knife’, *xī* ‘sunset’, *kǒu* ‘mouth’ and *wáng* ‘perish’.

Figure 7 – A field manual for dialect research

56 臻開三

臻開三: 眞(臻)質(櫛)				
	平	上	去	入
	眞(臻)	軫	震	質(櫛)
幫滂並明	彬·賓檳檳 貧·頻頻繁 閩閩越·民	憫敏抿(*搯)	·殞鬢	筆·畢必 ·匹一匹布,一匹馬 弼密·蜜
端透定				
泥(娘)來	鄰麟麟		吝吝吝	栗

The 方言調查字表 *Fāngyán diàochá zì biǎo* 'Character tables for dialect research' were compiled in the 1950s by the Academia Sinica and have been distributed to generations of linguistic field workers.

Fāngyán (1964: 56)

In the academic world, meanwhile, philological studies in Classical Chinese now co-existed with an interest in the diversity of colloquial forms. Genetic approaches to linguistic change were adopted (Branner 2000: 8) and modern dialects started to be described in comprehensive reference grammars. By the middle of the 20th century these developments had brought about a revolution in Chinese linguistic scholarship in areas such as syntax, morphology, and linguistic theory. In traditionally well-developed fields such as phonology and lexicography, a sense of continuity was maintained, past achievements paving the way for a new commitment to the documentation of the sounds and words of living speech traditions. In this way modern Chinese linguistics bridged the gap between traditional rhyme studies and modern dialectology. The legacy of Y.R. Chao has been mentioned above; it is no coincidence that he was both a pioneer of the international phonetic movement and a ground-breaking Chinese dialectologist.

One result of the blend between Chinese traditions and Western methods is the field guide shown in Figure 7. Modern Chinese dialect investigation manuals not only include tables to elicit character readings but also assume an active knowledge of the International Phonetic Alphabet. Nonetheless, the elicitation of character readings cannot provide a reliable starting point for dialect classification. As Branner (2000: 25) puts it, lexical inventories based on character readings are "designed to show the relationship of a dialect to the reading tradition of medieval China".

Thus, for the genetic linguist, the Chinese script interferes with the discovery of systematic correspondences between dialects. In linguistics at large, likewise, Chinese character bias impedes the inventory of expressions lacking written representations. Paradoxically, while Chinese historical linguistics owes much of its success to the early invention of writing, the same script often prevents an adequate description of contemporary language use. As a result, everyday modes of expression tend to escape linguistic notice whenever they differ from written usage. Examples from Peking Mandarin are (1) the high frequency of the nominal predicate, (2) the prevalence of *bu* 'not' over *bù ~ bú* 'not'; (3) the existence of the

negative verb *bú* ‘not to be’; and (4) the forward shift of alveodental fricatives and affricates. An example from Taiwan Mandarin is (5) the wide-spread use of the object marker *bǎ* without an overt object.² These phenomena are illustrated below. The character representations in parentheses are fair though inadequate approximations of these spoken expressions; the references describe further details.

- (1) Tā Táiwān rén. (她台灣人。)
 3 PNTáiwān person Wiedenhof (1995: 65, 2004: 78-79)
 ‘She’s from Taiwan.’
- (2) Wǒ bu qù. (我不去。)
 1.SG not go Wiedenhof (1995: 63, 2004: 26)
 ‘I’m not going.’
- (3) Tā bú Táiwān rén. (她還台灣人。)
 3 not.be PNTáiwān person Wiedenhof (1995: 62-73)
 ‘She’s not from Taiwan.’
- (4) [t̪hã˥˥ ˨˩ ʊe˨˩ zəɿ˥˥]
 zán sā rér Wiedenhof (2004: 63); compare Pulleyblank (1984: 29)
 1.INC three.pieces body
 ‘us three’
- (5) Nǐ kéyi bá dǎ kāi. (妳可以把打開。)
 2 may OBJ hit open Chao (1968: 330), Wiedenhof (2004: 139)
 ‘You can open it.’

2. The Pīnyīn era

After the establishment of the People’s Republic in 1949, efforts to reform the Chinese script continued. The committees which developed *Hànyǔ Pīnyīn* (漢語拼音, literally, ‘Chinese spelling’), or Pīnyīn for short, drew upon the work of Soviet linguists. The transcription known as Latinxua (拉丁化 *Lādīnghuà* ‘Latinization’, also called *Sin Wenz*, 新文字 *Xīn wénzì* ‘New Script’) had been designed for Chinese minorities in the Soviet Union. The choice of the Roman rather than the Cyrillic alphabet for Soviet minority languages was a deliberate break with the assimilation policies of the tsarist era (DeFrancis 1950: 107). When Chinese linguists and communist activists embraced the system, a sizeable Latinxua movement arose in the middle of the 1930s, with representations in several Chinese cities and overseas Chinese communities. Latinxua, as the spelling of its name suggests, ignored tones as a matter of principle. It also targeted various Mandarin as well as non-Mandarin dialects. During the Second World War, interest in the system collapsed; by the end of the 1940s, the Latinxua movement had been abandoned in communist and Kuomintang areas alike.

² Similar constructions without overt objects are also found for the coverbal use of verbs such as *gěi* ‘give’ and *zài* ‘be in’ (Wiedenhof 2004: 127). As pointed out by one anonymous reviewer, the use of *bǎ* without overt object in Taiwan Mandarin is reinforced by a similar construction of the Taiwanese Mǐn pretransitive marker *ka*⁷, e.g. *Gua*²⁻¹ *u*⁷⁻³ *ka*⁷⁻⁷ *khuann*³⁻³. ‘I looked at it.’ (*Taiwanese* 1990: 7; compare Klötter 2005: 23).

After almost a decade of preparation, the People's Republic officially adopted Pīnyīn as a national standard in 1958. Two years earlier, the first batch of simplified Chinese characters had been issued, and Pīnyīn was to play a supportive role in the promulgation of the new character script. Early revolutionary ideas about the abolition of Chinese characters in favor of the Latin alphabet were set aside, and the publication of the *Jiǎnhuà zì zǒngbiǎo* 'Comprehensive list of abbreviated characters' in 1964 guaranteed the future of the character script.

During the 1960s, Pīnyīn spread quickly throughout the People's Republic, due to its official endorsement and active propagation. The international acceptance of Pīnyīn since the 1970s has been described above; today, Pīnyīn is by far the most influential alphabetic spelling for Mandarin.

2.1. The linguistic impact of Pīnyīn

Due to the dominance of Pīnyīn, contemporary students of Mandarin are less familiar with other systems. As a result, access to earlier linguistic sources may be limited. For example, it is little known today that just half a century ago, the Mandarin object marker was *bǎi*. Linguistics students studying example sentences from Chao's *Grammar of spoken Chinese* (1968) tend to favor the character versions and to ignore his now unfamiliar Gwoyueu Romanzylh transcriptions. Chao's consistent GR spelling of the object marker as *bae* (Figure 8) tends to be misconstrued as Pinyin *bǎ* (GR *baa*), while the character 把 does not bear out the difference. An unambiguous and recent instance of linguistic change (Wiedenhof 2004: 136-137) is thus easily overlooked.

The controversy about the marking of tones in Chinese romanization had been a bone of contention ever since native developments took over from missionary spelling systems. This issue seemed settled with the victory of Pīnyīn over Latinxua, which had systematically avoided tone. In Pīnyīn, tones are represented by means of diacritical marks over the vowel: first tone *bāo*, second tone *báo*, third tone *bǎo*, fourth tone *bào*; the neutral tone is unmarked, *bao*. This graphical convention has had two regrettable consequences.

Figure 8 – The object marker *bǎi*

Excerpt from the description of the pretransitive Mandarin object marker *bǎi* in Yuen Ren Chao's monumental *Grammar of spoken Chinese*, which uses the Gwoyeu Romatzyh (GR) transcription. Chao's GR spelling *bae* (Pīnyīn *bǎi*) is easily confused with the modern spoken form, which is transcribed as *bǎ* in Pīnyīn and as *baa* in GR. Chao does quote *bǎ* as a citation form (1968: 343, note 48); his rare examples with *bǎ* instead of *bǎi* belong to formal styles (e.g. on pp. 345-346).

(5) Second Verbs with Objects. Instead of—or in addition to—a complement, a second verb may take an object other than the object of the pretransitive, as illustrated below.

(a) With Cognate Object. A cognate object, as we have seen, may consist of the verb itself, or the number of times, extent, or duration of the action. Examples are: 把衣裳燙一燙! *Bae i.shang tanq .i.tanq!* 'Give the clothes an ironing!', whence the tentative aspect in the form of reduplication: *Bae i.shang tanq .tanq!* 'Just iron the clothes!' Sometimes, instead of V.V for V-i-V, the first verb is omitted, resulting in i-V, as in 把眼睛一翻 *bae yean.jing i-fan* instead of *bae yean.jing fan.i.fan* 'roll the eyes once,—just roll the eyes'. With number of times: 他把全文讀了一遍. *Ta bae chyuan-wen dwule i-biann.* 'He read the whole text once through.' With extent: 得把這摞紙軋着点儿 *deei bae jey-luoh jyy yahj .deal* 'must cover (with a paper-weight, etc.) this pile of paper a little', 把袖子放三寸 *bae shiowtz fanq san-tsuenn* 'let out the sleeves by three inches'. With duration: 把水煮了好半天 *bae shoei juule hao bann-tian* 'have boiled the water a good long time', 把他騙了十年 *bae ta piannle shyr-nian* 'deceived him for ten years'.

Chao (1968: 349)

- (1) Technically, tone symbols used to be a printer's nightmare; they continue to cause distress for users of word processing software. In the early days of the digital age, the ASCII standard offered two Pīnyīn tone marks for most vowels, if only because the second and fourth tones happened to coincide with the French *aigu* and *grave* accents. The transcription of the first and third tones became available in separate symbol fonts, although these often failed to handle tone marks over the Pīnyīn vowel *ü*, as in *lǜ* 'green'. The ISO-endorsed Unicode standard fully incorporates the Pīnyīn system. It has gained acceptance in recent years, but for the average user it remains difficult to determine if a given computer font complies with Unicode; and if so, to devise word processing shortcuts for the insertion of Pīnyīn tone symbols.
- (2) From a linguistic perspective, tone symbols may contribute to misconceptions about the nature of tone in tone languages, because these diacritics stand graphically separated from Latin letters. By contrast, the principle of graphical tone incorporation has the advantage that it visually represents the phonemic nature of tone. To illustrate, readers will readily appreciate that Gwoyeu Romatzyh *bau*, *baur*, *bao* and *baw* are distinct words meaning 'wrap', 'thin', 'plump' and 'newspaper', just as the phonemic distinctions between English *bait*, *bat*, *beat* and *boat* are reflected in the orthography. The superimposed tone marks of Pīnyīn *bāo*, *báo*, *bǎo* and *bào*, on the other hand, are easily regarded as superfluous, especially by native speakers of non-tone languages,

who may lump these expressions together. Most Western texts will eliminate the tone marks of Pīnyīn (i.e., “Pinyin”), as illustrated in Figure 3. Curiously, maps and newspapers ignore tone symbols in *Chéngdé* and *Mǎ Róng*, but do print the same diacritics in names such as *Chénéville* and *Milošević*.

To sum up, the success of Pīnyīn over Latinxua created the possibility of recording tones, but due to the relatively complex graphics of the tone symbols, they are little used outside the areas of language education and linguistics. Even Chinese passports, despite their obvious identificational function and the high frequency of identical personal names, do not specify Pīnyīn tones.

2.2. The perception of spoken Mandarin

As shown above, the character script influences the way spoken Mandarin is perceived and described. The Pīnyīn transcription likewise affects the perception and analysis of speech sounds. Consider, for instance, the phonological status in Mandarin of the schwa vowel [ə], which occurs in expressions such as the subordinative particle *de* and the perfective suffix *le*. Such grammatical function words typically have neutral tones, but they may acquire a fourth tone citation reading, for instance in a linguistic discussion: *Zhèibīān bú yòng de yě xíng*. ‘In this case it’s also fine to leave the *de* out.’; *Yòng yí ge lè hái shì liǎng ge lè?* ‘Did he use one or two *le*’s?’. The pronunciation of *lè* in the last sentence displays a clear phonemic contrast with that of the verb *lè* ‘be amused’. The citation form of the particle *le* is pronounced as *lè* [lə̌], while the verb *lè* ‘be amused’ has a diphthong: [ɿʌ̌]. This fact is little known because the Pīnyīn transcription does not recognize the distinction. In the Wade Giles system, these two vowels are distinguished by the circumflex accent. Thus *lê*⁴ ‘be amused’ contrasts with *le*⁴ ‘the particle *le*’; *shê*² ‘snake’ with *she*² in *she*²·*me* ‘what’, etcetera.

In the case of *zhè* ‘this’, both [ʈʂə̌] and [ʈʂʌ̌] occur, the latter usually in formal situations, e.g. news broadcasts in the national media. The stylistic difference suggests that the form [ʈʂʌ̌] arose as a hypercorrection, but this cannot be confirmed because of the uncertain source of the vowel. Neater evidence is provided by *shéme* ‘what’, derived from *zjəp mjuət* 什物 ‘appliances, utilities’ > ‘thing’ > ‘which thing’ > ‘what’ (Zhāng 1982: 302-303, Norman 1988: 119-120). Many dictionaries and textbooks transcribe the expression as *shénmo* or *shénme*. The -n- in these spellings is due to the Chinese script, for the character 甚 in 甚麼 *shéme* ‘what’ is also used to write Classical Chinese *shèn* ‘tremendous’. The same scriptural influence also leads to hypercorrect pronunciations such as [ʂən̩ ɿmɔ̌] and [ʂən̩ ɿmɔ̌] by speakers from non-Mandarin backgrounds.

2.3. The transcription of tone sandhi

In the alphabetical world, the popular notion of a ‘word’ is by no means linguistically inspired. It is a notion based on the script — a word is a series of letters separated from other words by spaces. In Mandarin, too, the popular notion of the basic units of language is motivated by written norms, but due to the nature of the character script, the Chinese notion comes closer to the morpheme (Chao 1968: 137-139).

Consequently, for Mandarin systems of transcription, the syllable is the maximally relevant unit. While the identification of Pīnyīn 詞 cǐ ‘words’ as syllables joined together in writing has been the subject of extensive discussion and regulation (*Zhèngcè fǎguī* 1996; for some background and discussion see also *Pīnxié fǎ* 1953 and *Pīnyīn zhèngcǐ* 1985), no living tradition has materialized so far, for the simple reason that the character script dominates all ordinary functions of writing.

One clear victim of this state of affairs is the transcription of tone sandhi, which typically transgresses syllable boundaries. Take, for example, the first tone sandhi rule taught in Mandarin courses: a third tone before another third tone is pronounced as a second tone, e.g. nǐ ‘you’ + hǎo ‘good’ > ní hǎo? ‘How are you?’. The rule is valid almost universally, with the exception of major syntactic boundaries between two third tones, as in *Nǐ? Hǎo.* ‘You? Okay.’, which consists of two separate sentences.

Now consider the first line from a dialogue in a modern textbook which, like most textbooks on the market today, does not indicate tone sandhi in its Pīnyīn transcriptions (Figure 8). The expression ‘how are you?’ is spelled as “nǐ hǎo” in the first line, but the same line contains another sequence of two third tone marks in the sentence “Rènshi nǐ hěn gāoxìng”, meaning ‘Glad to meet you’. Interestingly, the tone sandhi rule does not apply here, because the sentence consists of two clauses, i.e. literally *Rènshi nǐ...* ‘Knowing you...’ and *...hěn gāoxìng.* ‘...[I am] very glad’. If the tone sandhi rule were to be applied here, the personal pronoun would become part of the second clause, yielding an unintended *Rènshi, ní hěn gāoxìng.* ‘That you know them makes you happy.’. In other words, students of the language have to figure out where tone sandhi rules apply even though the text is transcribed in a system which is fully equipped to distinguish these different tonal readings.

Figure 8 – Transcribed Mandarin in a modern textbook

Lǐ Yǒu: Xiǎoyīn, nǐ hǎo. Rènshi nǐ hěn gāoxìng ⁽¹⁾.
 Gāo Xiǎoyīn: Rènshi nǐmen wǒ yě hěn gāoxìng.
 Lǐ Yǒu: Nǐmen jiā hěn dà ^(G2), yě hěn piàoliang.
 Xiǎo Gāo: Shì ma? ⁽²⁾. Qǐng zuò, qǐng zuò.
 Wáng Péng: Xiǎoyīn, nǐ zài ^(G3) nǎr gōngzuò?
 Gāo Xiǎoyīn: Wǒ zài xuéxiào gōngzuò. Nǐmen xiǎng hē diǎnr ^(G1) shénme?

Fragment of a dialogue for beginning students in the *Integrated Chinese* series (Tao e.a. 1997: 93).

In sum, in the only domain which retains Pīnyīn tone marks, viz. the study of language, students are often made to decode transcribed texts. This situation is understandable in view of Chinese character writing traditions, but defies the purpose of the Pīnyīn transcription system, i.e., to record the sounds of spoken Mandarin.

3. Conclusions

Any assessment of purpose and effect in the transcription of Mandarin has to take into account that it fulfills different needs inside and outside China. In China, transcription is

used as an educational tool and as an orthography serving foreign communities. The broad educational significance of Pīnyīn is limited to elementary schools, where the system serves as a precursor to character orthography. Early plans to abolish Chinese characters altogether have long disappeared from the political agenda. Outside China, a distinction can be made between general and specialist applications. International journalism, commerce and tourism dispense with tone marks. The use of tone marks is thus restricted to linguistic environments such as language learning and dialect fieldwork.

The impact of graphical representations of language on the way we perceive and analyse language comes as no surprise, for the dominant role of the visual cortex in the observation processes of the human brain is well known. “While other senses such as hearing and touch are essential, visual information dominates our perceptions and frames the way we think” (Gazzaniga e.a. 1998: 123). For the student of language, therefore, the only hope is to keep both ears wide open.

Glosses

1	first person
2	second person
3	third person
INC	inclusive, i.e. the notion ‘we’ including the hearer
OBJ	object marker
PN	proper name
SG	singular

References

For reprinted titles, the year given in the left column represents the edition consulted.

- Abel-Rémusat (1857) [Jean Pierre] Abel-Rémusat, *Éléments de la grammaire chinoise: ou Principes généraux du kou-wen ou style antique, et du kouan-hoa, c'est-à-dire, de la langue commune généralement usitée dans l'empire chinois*. Reprint of the 1822 *Imprimerie royale* edition. L. Léon de Rosny, ed., Paris: Maisonneuve, 1857. A second reprint with an Introduction by Alain Peyraube has been published in Paris: Ala Productions, 1987.
- Bloomsbury (2004) Kathy Rooney, ed., *Bloomsbury English dictionary*. London: Bloomsbury, 2004. Second edition of the *Encarta world English dictionary*, 1999.
- Branner (2000) David Prager Branner, *Problems in comparative Chinese dialectology: The classification of Miin and Hakka*. Berlin: Mouton de Gruyter, 2000.
- Chao (1968) Yuen Ren Chao, *A grammar of spoken Chinese*. Berkeley: University of California Press, 1968.
- DeFrancis (1950) John DeFrancis, *Nationalism and language reform in China*. Princeton: Princeton University Press, 1950.
- Fāngyán (1964) 中國科學院語言研究所 Zhōngguó Kēxué Yuàn Yǔyán Yánjiū Suǒ [Department of Linguistic Research, Academia Sinica], ed., 方言調查字表

- Fāngyán diàochá zì biǎo* [Character tables for dialect research]. Revised edition, 北京 Peking, 科學出版社 Kēxué Chūbǎnshè, 1964. First edition 1955.
- Gazzaniga (1998) Michael S. Gazzaniga, Richard B. Ivry and George R. Mangun, *Cognitive neuroscience: The biology of the mind*. New York: W.W. Norton, 1998.
- Giles (1912) Herbert A. Giles, *A Chinese-English dictionary*. First edition 1892, second edition Shanghai: Kelly & Walsh and London: Bernard Quaritch, 1912.
- ISO (1982) *International Standard 7098: Documentation – Romanization of Chinese, ISO 7098-1982 (E)*. [Genève:] International Organization for Standardization, 1982.
- Klötter (2005) Henning Klötter, *Written Taiwanese*. Wiesbaden: Harrassowitz, 2005.
- Legeza (1968-1969) Ireneus László Legeza, *Guide to transliterated Chinese in the modern Peking dialect*. In two volumes. Leiden: E.J. Brill, 1968-1969.
- Mathews (1943) R.H. Mathews, *Chinese-English Dictionary: Revised American Edition*. Y.R. Chao and M.Y. Wang, eds. Cambridge, Massachusetts: Harvard-Yenching Institute, 1943.
- National Geographic* (1999) *National Geographic atlas of the world: Seventh edition*. Washington: National Geographic Society, 1999.
- Norman (1988) Jerry Norman, *Chinese*. Cambridge: Cambridge University Press, 1988.
- Pīnxié fǎ* (1953) 林漢達 Lín Hàndá e.a., eds., 漢語的詞兒和拼寫法 *Hànyǔ de cí ér hé pīnxié fǎ* [Chinese words and their transcription], 第一集 *dì yì jí* [Part 1]. 北京 Peking: 中華書局 Zhōnghuá Shūjú, 1953.
- Pīnyīn zhèngcí* (1985) 《汉语拼音证词法论文选》编辑组 *Hànyǔ Pīnyīn zhèngcí fǎ lùn wén xuǎn* Biānjí Zǔ [Editors of *Pīnyīn orthography: Selected articles*], ed., 汉语拼音证词法论文选 *Hànyǔ Pīnyīn zhèngcí fǎ lùn wén xuǎn* [Pīnyīn orthography: Selected articles] 北京 Peking: 文字改革出版社 Wénzì Gǎigé Chūbǎnshè, 1985.
- Pulleyblank (1984) E.G. Pulleyblank, *Middle Chinese: A study in historical phonology*. Vancouver: University of British Columbia Press: 1984.
- Taiwanese* (1990) *Taiwanese: Book 2*. 台中 Taichung: Maryknoll Language Service Center, 1985. Reprinted with slight corrections, 1990.
- Tao e.a. (1997) Tao-chung Yao, Yuehua Liu, Liangyan Ge, Yea-fen Chen, Nyan-ping Bi and Xiaojun Wang, *Integrated Chinese: Traditional character edition, Textbook, Level 1, Part 1*. Boston: Cheng & Tsui, 1997.
- Trigault (1957) 金尼閣 Jīn Nígé [Niklaas Trigault], 西儒耳目資 *Xīrú ěr mù zī* [Occidental scholar's auditory and visual guide], [杭州 Hángzhōu,] 1626. Facsimile reprint, 北京 Peking: 文字改革出版社 Wénzì Gǎigé Chūbǎnshè, 1957.
- Varo (2000) Francisco Varo, *Arte de la lengua mandarina*. Canton: 1703 edition of a manuscript completed in 1682. Reprinted in W. South Coblin & Joseph A. Levi, eds., *Francisco Varo's Grammar of the Mandarin language (1703): An English translation of 'Arte de la lengua mandarina'*. Amsterdam: John Benjamins, 2000. With an Introduction by Sandra Breitenbach.

- Wiedenhof (1995) Jeroen Wiedenhof, *Meaning and syntax in spoken Mandarin*. Leiden: CNWS Research School, 1995.
- ___ (2004) Jeroen Wiedenhof, *Grammatica van het Mandarijn* [A grammar of Mandarin]. Amsterdam: Bulaaq, 2004.
- Zhāng (1982) 张惠英 Zhāng Huìyīng, 释《什麼》“Shì shéme” [An explanation for shéme ‘what’]. 中国语文 Zhōngguó Yǔwén, 1982, no. 4, pp. 302-305.
- Zhèngcè fǎguī (1996) 国家语言文字工作委员会政策法规室 Guójiā Yǔyán Wénzì Gōngzuò Wěiyuánhùi, Zhèngcè Fǎguī Shì [Policies and Regulations Office, National Committee for Language and Writing], ed., 国家语言文字政策法规汇编 Guójiā yǔyán wénzì zhèngcè fǎguī huìbiān [National policies and regulations on language and writing: Collected documents]. 北京 Peking: 语文出版社 Yǔwén Chūbǎnshè, 1996.