

## The Classification of Sinitic Languages: What Is “Chinese”?

Victor H. Mair (梅維恒)  
University of Pennsylvania

*Chinese as the name of a language is a misnomer. It has been applied to numerous dialects, styles, and languages from the middle of the 2<sup>nd</sup> millennium BC. Sinitic is a more satisfactory designation for covering all these entities.... The present-day spoken languages are not mutually intelligible (some are further apart than Portuguese and Italian), and neither are the major subdivisions within each group. Søren Egerod (1988)*

*I think we should refer to the complex of linguistic forms now called Chinese “dialects” as the Sinitic language family. Jerry Norman (2009a)*

*If one’s terms are imprecise, one’s language will be illogical. 名不正，言不順. Analects, 13.3*

Terminological imprecision, particularly with regard to the Chinese word *fangyan* and its translation by the English word “dialect,” has resulted in a situation whereby Sinitic language taxonomy may variously be described as chaotic, impenetrable, or functionally absent. For such a large, diverse agglomeration of languages as Sinitic, this is an unacceptable state of affairs. Through rigorous definition and careful analysis, it is possible to arrive at a clearer understanding of the nature of the relationships among the constituent languages of the Sinitic Language Group/Family (SLG/F).

Key words: Sinitic languages, languages versus dialects, topolects, *Putonghua*

---

\* For assistance in researching various aspects of this study, I am deeply grateful to Jerry Norman, W. South Coblin, Tsu-Lin Mei, Wolfgang Behr, William H. Baxter, Guillaume Jacques, Bai Weiguo, Huang Heqing, Xu Wenkan, Yao Dehuai, Sanping Chen, Jidong Yang, David Prager Branner, Thomas S. Mullaney, Mark Elliott, Robert Lindsay, Linda Chance, Alban Kojima, S. Robert Ramsey, William C. Hannas, Neil Kuster, Minkyung Ji, and Daniel Sou. The author alone, however, is accountable for all facts and interpretations presented herein.

## 1. Preliminaries

The present study is an extension of the author's "What is a Chinese 'dialect/topolect'?" Reflections on some key Sino-English linguistic terms" (Mair 1991). This is an updating of that paper and should be read in conjunction with it. We may summarize the major findings of that article as follows:

i. Loose usage of terms such as *fangyan* 方言 (universally, but mistakenly, rendered as 'dialect' in English), Chinese, and *Hanyu* 漢語 'Han language' has drastic implications for scientific, linguistic classification.

*Corollary:* Only with precise terminology can we have accurate, exacting classification.

ii. A more precise translation of *fangyan*, one that effectively neutralizes its deleterious consequences for the scientific classification of the languages of East Asia, is 'topolect.' This is a word which means exactly the same thing as *fangyan*, no more and no less, and which has been accepted by the *American Heritage Dictionary of the English Language* and other authorities.

iii. The number of *fangyan* 'topolects' and their taxonomic status is uncertain.

iv. The number of so-called *da fangyan* 大方言 ("major topolects") and their taxonomic status is uncertain.

v. The demographically huge aggregation of speech forms loosely and collectively known as *Hanyu*<sup>1</sup> 'Han [People's] Languages[s]' is fully deserving of rigorous classification comparable to that which has been applied to Indo-European, Tibeto-Burman, Afro-Asiatic (Semitic-Hamitic), Amerindian, Austronesian, Austro-Asiatic, and virtually all the other major language stocks on earth. It is inconceivable that *Hanyu* has been and is taxonomically undifferentiable throughout time and space. Yet the most authoritative reference works in China, such as the *Yuyan wenzi* 語言文字 [Language and Script] volume of the *Zhongguo da baike quanshu* 中國大百科全書 [Great Encyclopedia of China], fail to divide *Hanyu* into constituent subdivisions, treating it as a single entity. This is in stark contrast to their treatment of much smaller groups such as Tibeto-Burman, Miao-Yao, and Zhuang-Dong, which they divide into numerous branches, sub-branches, and languages.

---

<sup>1</sup> To avoid an excess of what might be interpreted as scare quotes, all occurrences of '*Hanyu*' in this paper should be understood as prefaced by 'so-called.' The reason for this is that, in its current undifferentiated, unanalyzed state, '*Hanyu*' is nugatory in discussions of the scientific classification of Sinitic languages. Due to similar developmental and historical reasons, the same holds for '*Han*,' '*Hanzu*,' and '*Han minzu*.' See §3.

vi. Ideally, linguists should strive for a logical division of *Hanyu* into branches, sub-branches, languages, dialects, and sub-dialects, just as they do for all other major language groupings.

vii. Traditional understandings of *fangyan*, which are still operative in contemporary discussions of the nature of *Hanyu*, are incompatible with modern, linguistic classification.

viii. ‘Mandarin’ is an accurate and appropriate English translation of *guanhua* 官話, the forerunner of *guoyu* 國語 ‘National Language’ as used in Taiwan,<sup>2</sup> *huayu* ‘Hua [People’s] Language’ as used in Singapore, and *Putonghua*, ‘Common Speech’ as used in People’s Republic of China.

ix. ‘Mandarin’ is not equivalent to ‘Chinese,’ whatever the latter may signify.

x. The fact that there are approximately two dozen different ways to refer to Mandarin in Chinese, including several that end in *wen* 文 ‘writing’<sup>3</sup> is a clear indication of the nomenclatural imbroglio surrounding this language (or congeries of languages and dialects).

xi. The various speech forms of *Hanyu* significantly differ among themselves in respect to phonology, morphology, lexicon, idiomatic usage, grammar, syntax, and all other aspects of linguistic composition and construction.

xii. *Hanyu* simultaneously means two very different things: Modern Standard Mandarin (MSM) and the totality of all languages spoken by the so-called Han peoples. Because this multivalence of *Hanyu* has not hitherto been sufficiently emphasized, much less recognized, it leads to subconscious confusion concerning the interrelationships among the components of the Sinitic Language Group/Family (SLG/F).<sup>4</sup> Cf. §3.1.

xiii. Cantonese and Mandarin are separate languages. Cantonese is not a ‘dialect’ of Mandarin or of *Hanyu*, and it is grossly erroneous to refer to it as such. Since Cantonese and Mandarin are separate languages (or, perhaps more accurately, separate branches), it is wrong to refer to them as ‘dialects.’ The same holds for Hokkien, Shanghainese, and so forth.<sup>5</sup>

---

<sup>2</sup> For a discussion of the history of the term *guoyu*, see Mair (1994).

<sup>3</sup> For a brilliant, but little known, disquisition on the history of the relationship between writing and language in the EAH, see Lu Xun (1934, reprint 2005).

<sup>4</sup> If efforts to link Sinitic with other major language groups continue to be as unconvincing as they have been to date, it may well be that Sinitic will end up being classified as a family unto itself. Because it remains to be determined whether Sinitic is a group or a family, I provisionally style it the Sinitic Language Group/Family (SLG/F).

<sup>5</sup> N.B.: Mair (1991) and other works by the author cited in the Bibliography serve to document diverse unreferenced statements herein. See also Norman (1988), Ramsey (1987), DeFrancis (1984), and Zhou (2003).

## 2. Presuppositions

i. Language groups are classifiable, both internally and externally (Lass 1997, Campbell 1997).

ii. A lect<sup>6</sup> is a discrete, distinct form of speech that is identified with a particular population, whether such a population be determined by place, ethnicity, occupation, or other criterion.

iii. There are countless lects in China, many of which are completely or partially unintelligible to each other.

iv. Mutual (un)intelligibility<sup>7</sup> is determinable and quantifiable (Casad 1974, Cheng 1982, 1991, 1998, and Tang & van Heuven 2009).

v. Mutual (un)intelligibility is significant and should be taken into consideration in the classification of the SLG/F.

vi. Phonological, morphological, lexical, and even syntactical criteria may also be used for genetic classification, with phonological criteria being particularly relevant in the case of Sinitic (Coblin 2009).

## 3. Problems

i. There are more than two dozen synonyms for the national language of China (Mair 1991:11). The two most frequently encountered nowadays are *Hanyu* and *Putonghua*. What is the relationship between *Hanyu* and *Putonghua*? Unmistakably, *Putonghua* constitutes a subset of *Hanyu*, since the former term was adopted by the government of the People's Republic of China government after 1950 as a designation for the contemporary lingua franca of the nation. That is to say, *Putonghua* is a particular type of *Hanyu*; it is the 'common speech' of the nation in contrast to all of the other forms of *Hanyu* which are considered to interfere with or inhibit communication among the Han

---

<sup>6</sup> Although 'lect' usually refers to social or regional varieties within a speech community, I accept this convenient usage of SIL because it is the most neutral designation I can think of for a separately identifiable pattern or form of speech, even more so than topolect. SIL (formerly known as the Summer Institute of Languages), which runs the influential, authoritative online list of languages entitled *Ethnologue*, has been charged with the task of assigning ISO numbers to lects. The assignment of an ISO number to a lect means that it has been officially recognized by the world linguistic community as a separate language.

<sup>7</sup> The reason for the formulation 'mutual (un)intelligibility' is that 'intelligibility' and 'unintelligibility' correspond to each other in an obverse fashion. For example, if lect A is considered to be 40% intelligible to speakers of lect B, it may also said to be 60% unintelligible to speakers of lect B.

peoples. In contrast to *Putonghua*, *Hanyu* is recognized diachronically not only as having a much lengthier reach backward in time (supposedly to the very beginning of the SLG/F), but also synchronically beyond the limits of the lingua franca to embrace all the local and regional varieties of related speech forms that are more or less mutually unintelligible. Indeed, the very *raison d'être* for *Putonghua* is to serve as a means of communication for speakers of various lects of *Hanyu* who would otherwise not be able to talk to each other unconstrainedly (or at all). But *Hanyu*, in the sense of a comprehensive entity that includes all forms of language presumably spoken by individuals of Han ethnicity, throughout the time of existence of the Han people, is a modern construct.

ii. *Hanyu*, as employed in current cultural and political discourse, is a thoroughly modern, nationalist concept. Fundamentally, *Hanyu* signifies the language(s) spoken by the *Hanzu* 漢族 ‘people[s] of Han ethnicity’, while *Hanzu* is short for *Han minzu* 漢民族 ‘the Han nation/people/ethnos’. Both *Hanzu* and *minzu*—with their current signification of nationhood—are expressions that cannot be found before around the turn of the 20<sup>th</sup> century and are the result of nationalistic discourse that came to East Asia from the West. More specifically, they are likely to have been calqued in Japan or under Japanese influence to cope with new concepts concerning nationhood and ethnicity (Wilkinson 2000:708, Fogel 2005, Doak 2007, Huang 2009a, Xu 2009, and Yao 2009).

The hallowed appellation ‘Han’ was originally the self-designation for two celebrated dynasties: the Former/Western Han (202 BC-AD 23) and the Later/Eastern Han (25-220), separated by Wang Mang’s Xin (‘New’) dynasty (9-23). Naturally, the subjects of the two Han dynasties were contemporaneously referred to as *hanren* ‘Han person[s]’, but, with the collapse of the Eastern Han, such terms fell out of use. In the whole of the massive *Zhu Zi yulei* 朱子語類 [Classified Conversations of Master Zhu], the leading Neo-Confucian scholar of his age, Zhu Xi (朱熹; 1130-1200) employs the term ‘Han person(s)’ twenty-seven times, but in each case he means ‘people of the Han Dynasty’, not an ethnically distinct group (Zhao 2009). This is particularly revealing in light of the fact that *Zhu Zi Yulei* has a considerable proportion of vernacular passages. Although the name Han was sporadically resurrected by several ephemeral political entities in the centuries that followed, the last dynasty to call itself Han fell in 1364, after less than four years of existence (Beck 1986:371-373).

As a matter of fact, in the centuries from the demise of the Eastern Han up to modern times, terms such as *han* 漢, *hanzi* 漢字, *hanren* 漢人, and *han'er* 漢兒 were used (often derogatorily<sup>8</sup>) by the northern peoples (e.g., Särbi, Khitans, Jurchens,

<sup>8</sup> The derogatory connotations of ‘Han’ have continued into modern times. Amazingly, when not used in its neologistical ethnic sense, *han* in contemporary usage frequently conveys a pejorative sense, as in the following common expressions: *chunhan* 蠢漢 ‘dull person’, *benhan* 笨漢 ‘stupid man’, *shahan* 傻漢 ‘foolish man’, *lanhan* 懶漢 ‘lazy man’, *wulaihan* 無賴漢 ‘rascal’,

Mongols) who established control over all or large parts of what is now called ‘China’ to refer to their indigenous subjects. Hence, in *Bei shi* 北史 [History of the Northern Dynasties], we find such expressions as *gouhan* 狗漢 ‘dog Han’ (92.3053; also in *Bei Qi shu* 北齊書 [History of the Northern Qi] 50.692-693) and *kongtouhan* 空頭漢 ‘empty-headed Han’ (54.1967). An especially lackluster whitewashing of such a deprecatory term by a later editor occurs in the basic annals of emperor Wenxuan 文宣 (r. 550-559) in *History of the Northern Dynasties* and in *History of the Northern Qi* where the emperor refers to an official that he despises as a *chihan* 痴漢 ‘silly Han’. Sima Guang 司馬光 (1019-1086), in *Zizhi tongjian* 資治通鑑 (Comprehensive Mirror for Aid in Government), 166.5149-5150 changes this to *chiren* 痴人 ‘silly person’. The renowned synthesizer’s systematic, stubborn bowdlerization of any references to *han* in the official histories is almost laughably documented at 190.5959 of the *Comprehensive Mirror*, which has the Tang emperor Gao Zu declare, “This son has long been occupied with military affairs outside and was taught by a bookish student. He’s no longer the lad of old.” 此兒久典兵在外，爲書生所教，非復昔日子也。 In *Jiu Tang shu* 舊唐書 [Old Tang History] 64.2415-2416, this was: “This son has already been exclusively occupied with military affairs outside for a long time and was taught by a book-reading Han. He’s no longer my lad of old.” 此兒典兵既久，在外專制，爲讀書漢所教，非復我昔日子也。 In *Xin Tang shu* 新唐書 [New Tang History] 79.3540, the equivalent passage reads thus: “[This] son has long been occupied with military affairs and was corrupted by a Confucian student. He’s no longer my son of old.” 兒久典兵，爲儒生所誤，非復我昔日子。 The lengths to which the Confucian historian went to avoid mention of any sort of Han becomes truly comical at 194.6096 which relates an expostulation of the Tang emperor Tai Zong: “When the emperor had departed from the court, he angrily said, ‘It may be necessary to kill this old country bumpkin.’” 上嘗罷朝，怒曰：會須殺此田舍翁。 Compare this to an event as recorded in *Da Tang xinyu* 大唐新語 [New Accounts of the Great Tang] 1.13: “When the emperor had departed from the court, he said to himself, ‘I’ll have to kill off this yokel-Han.’” 太宗嘗罷朝，自言：殺卻此田舍漢。 The parallel passage in *Sui Tang jiahua* 隋唐佳話 [Fine Tales from the Sui and Tang] A7 has:

---

*manghan* 莽漢 ‘boorish fellow’, *menwaihan* 門外漢 ‘outsider’, *liulanghan* 流浪漢 ‘wastrel’, and *zuihan* 醉漢 ‘drunkard’. Positive expressions incorporating *han*, such as *haohan* 好漢 ‘brave man’, *tiehan* 鐵漢 ‘strong man’, *yinghan* 硬漢 ‘man of fortitude’, and *nanzihan* 男子漢 ‘real man’, are encountered much less frequently than those with negative connotations, have historically often been ironically employed by non-Sinitic overlords with reference to their autochthonous subjects, and usually are related to macho or outlaw characteristics. In fact, so accustomed are Chinese to referring to disreputable persons as ‘such-and-such a *han*’ that *-han* has become a standard suffix in the media for a person with negative associations, e.g., *bulianghan* 不良漢 ‘good for nothing guy’, *jiejihan* 劫機漢 ‘airplane hijacker’, etc. (Chen 2008a).

“Having departed from the court, Taizong angrily said, ‘It may be necessary to kill this yokel-Han.’”<sup>9</sup> 太宗曾罷朝，怒曰：會須殺此田舍漢。Sima Guang’s allergy to any expression involving *han* in the mouth of someone with northern associations even caused him to nullify a famous, colorful statement of Empress Wu Zetian 武則天 (625-705). Admittedly with a degree of arch ambiguity, the empress had said, “I need a good Han. Are there any?” 朕要一好漢，有乎？ (*New Accounts of the Great Tang* 6.94) Sima Guang’s wordy rendering of the Empress’s inquiry is innocuous and awkward: “I would like to obtain an excellent scholar-official that I can use. Who can do it?” 朕欲得一佳士用之，誰可者？ (*Comprehensive Mirror* 207.6551, first year of the Jiushi [Lasting Vision] reign period 700). Also in *Fine Tales from the Sui and Tang* B33, we find the Tang emperor Gao Zong (r. 650-683) referring to a harsh official who has caned a favored but miscreant eunuch as being “this savage Han” 此漢獐。In Feng Yan’s 封演 (*jinshi* 756) *Feng shi wenjian ji* 封氏聞見記 [Things Seen and Heard by Mr. Feng], 10.88, 10.93, a certain Wang Yanguang 王嚴光 declares, “Take this [ignorant] Han to use as turtle bait!” 取此等[蒙]漢以充鰲餌。 (Wang Yanguang is also identified as a turtle catcher in *Lei shuo* 類說 [Classified Conversations], 1.6.26a.) And in Zhuang Chuo’s 莊綽 12<sup>th</sup> century *Ji lei bian* 雞肋編 [Chicken Rib Chapters], B70, a Song emperor derides a fish merchant with the words, “This Han is malicious [literally, ‘toxic’]!” 這漢毒也。 In all of these instances, ‘Han’ may be interpreted as ‘fellow’, or worse. At best, toward the end of the Six Dynasties and continuing through the early Tang (while northern influences at court were still very strong), *Han* implied an educated, effete person, including Sinitically acculturated northerners (Chen 2008b). In Zen/Chan texts, *Han* occurs regularly as a colloquial term meaning ‘fellow’ (Broughton 1999:146).

In the “Benji” 本紀 [Basic Annals] of Shi Zu 世祖 of *Yuan shi* 原史 [History of the Yuan], 13.268, Sinitic-speaking Khitan and Jurchen individuals living in certain areas were legally classified as *hanren*. This shows clearly that *han* in the Mongol period was not thought of as an ethnic category in the same way it is today. The ethnolinguistic indeterminacy of *han* in pre-modern times is also borne out by a passage in the *Taiping huanyu ji* 太平寰宇記 [Universal Geography of the Taiping Era [976-983]], 35.11a (Wenhai chubanshe reprint (this passage is inexplicably missing from some editions, even in the *Siku quanshu* 四庫全書 [Complete Library of the Four Treasuries])), citing a Sui period (581-618) text, that refers to the Buluoji 步落稽 (associated by some with the Bulgars) as *hutou hanshe* 胡頭漢舌 ‘barbarian-headed Han-tongued’.

<sup>9</sup> It is remarkable how many different vernacular suffixes were available in medieval times for addition to the old classical expression *tianshe* 田舍 (literally, “fields and cottages”), ranging all the way from *gong* 公 and *lang* 郎 to *ren* 人, *zi* 子, *er* 兒, *weng* 翁, *niang* 娘 (N.B.), *nu* 奴, and *han* 漢, with Han evidently being toward the derisive end of the scale.

Before its modern reincarnation as the language of the ethnic Han nation, *Hanyu* had an earlier life as a designation for the language of the people of the Han Dynasty as shown by a search of large data bases such as CHANT (Chinese Ancient Texts) and CBETA (Chinese Buddhist Electronic Text Association). A CBETA search reveals that the term *Hanyu* did occur during the Six Dynasties, but primarily in contrast to *fanyan/you/wen* 梵言/語/文 (Sanskrit), *huyan/you* 胡言/語 ‘West Central Asian language’, etc. to describe the translation of Buddhist scriptures from Indic languages or the familiarization of foreign monks and translators with Sinitic languages. Moreover, terms such as *jinyu* 晉語 ‘language of the Jin people’ occur in similar contexts to distinguish Sinitic translation from Indic original. In other words, *Hanyu* in this early period meant ‘language of the Han Dynasty’ just as *Jinyu* meant ‘language of the Jin Dynasty’. In some instances, *Hanyu* meant ‘tales/stories/accounts of the Han period’, similar to the usage of in the title of the Warring States period work, *Guoyu* 國語 [Discourses of the States {of the Spring and Autumn Period}] or in the Liu Song period work entitled *Shishuo xinyu* 世說新語 (New Account of Tales of An Era) by Liu Yiqing 劉義慶 (403-444) concerning personalities who lived in China between about AD 150 and 420.

The term *Hanzu* did occur rarely before the end of the 19<sup>th</sup> century, but not with its modern meaning of ‘Han nation/people/ethnos’. The earliest occurrence of the term *Hanzu* is in an ironic context, namely, the memorial epitaph of the Son (Zen/Chan) Buddhist master Jin’gam (744-850), written by the famous Silla poet, Choe Chiwon 崔致遠 (857-928) upon the request of the Silla king Heongang: “Yu-Dang Silla’guk Go Jirisan Ssang’gyesa Gyosi Jin’gam Seonsa Bi’myeong” 有唐新羅國故知異山雙谿寺教諡真鑑禪師碑銘 (Stele Inscription for the Late Meditation [Son] Master Whose Posthumous Name Is Jin’gam, of the Ssang’gye Temple on Jirisan in the Tang Tributary State of Silla).<sup>10</sup> The epitaph is preserved at the Ssang’gye temple as Korean National Treasure No. 47. The sentence in question occurs near the beginning of the inscription and states that Jingam’s ‘ancestors were *hanzu*.’ 其先漢族. Choe Chiwon (MSM Cui Zhiyuan) was a Korean, he arrived in China at the age of 12 and passed the *jinshi* (advanced scholar) examination in 874. Several things about this early, isolated occurrence of the term *hanzu* are worth noting: a. it was used by a Korean with reference to another Korean, b. it was not taken up by other writers in the following centuries as a fixed term, and c. it was not intended as a blanket designation for Sinitic ethnicity. After that single medieval occurrence in a most intriguing context, *hanzu* is not in evidence again until near the end of the Qing dynasty when it comes to be employed by ethnically conscious Chinese, first in contrast to the Manchu overlords,

<sup>10</sup> Ssanggyesa, located on the southern slopes of Jirisan in Hadong County of Gyeongsangnam-do, in the southeastern part of South Korea, is the head temple of the Jogye Order of Korean Buddhism.



and then by nationalists like Sun Yat-sen in contrast to the other peoples of East Asia and Southeast Asia (Mengguzu 蒙古族 [Mongols], Malaizu 馬來族 [Malays], Ribenzu 日本族 [Japanese], and Manzu 滿族 [Manchus]).

The rarity of ‘Han’ as an umbrella designation for an East Asian state, people, or language after the third century is borne out by the fact that—in pre-modern times—foreign peoples referred to what we now know as “China” by such terms as Cina, Kitai, Seres, Tamghaj, Thin, and so forth, not anything resembling ‘Han’. There is no premodern citation for Hanzu in the *Hanyu da cidian* (HDC).<sup>11</sup> Considering the supreme importance and conspicuous frequency of the term in modern political rhetoric from the PRC, this could hardly be an accident.

iii. The relationships among *Hanyu*, *Hanzu*, and *Putonghua* are neatly—but somewhat perplexingly—summarized in the following formulation by Wang Li 王力 (1900-1986), an ardent proponent of *Putonghua*: ‘*Putonghua* is the standard language of Modern *Hanyu*; it is the common language of the *Han minzu*.’<sup>12</sup> 普通話是現代漢語的標準語，是漢民族共同語 (HDC 5.777a). From around this time, *Hanyu* and *Putonghua* came to replace *Guoyu* ‘National Language’ on the mainland, but were not exactly distinguished, one from the other, nor were they precisely differentiated from the innumerable more or less mutually unintelligible Sinitic lects spoken throughout China.

iv. As with *Hanyu* and *Hanzu*, similar observations may be made concerning the relatively recent appearance of *hanzi* in the sense of ‘sinogram’ or ‘Chinese character’. It is noteworthy that there are no premodern citations for *hanzi* in the HDC. The earliest known occurrences (1610, 1726) of *hanzi* are in the works of Matteo Ricci and his Jesuit successors in reference to Europeans familiarizing themselves with Sinitic languages and the Chinese script (Huang forthcoming). It is both curious and telling that these latter day occurrences of *Hanyu* and the apparently new coinage of *hanzi* are to be found in the context of foreigners becoming acquainted with Sinitic languages and the Chinese script. This parallels exactly what had occurred many centuries earlier when Indian and

<sup>11</sup> The abbreviations used in this paper are:

EAH: East Asian Heartland

EEAH: Extended East Asian Heartland

HDC: *Hanyu Da Cidian* 漢語大詞典 [Unabridged Word Dictionary of Sinitic]

MSM: Modern Standard Mandarin

<sup>12</sup> This statement is from Wang Li’s ‘Tuiguang Putonghua de san ge wenti’ 推廣普通話的三個問題 [Three Questions Concerning the Promotion of Putonghua]. Lu Xun (1881-1936), the greatest writer of the 20<sup>th</sup> century, had earlier used the expression *Putonghua*, but not as a proper noun, merely in the sense of “common language” (HDC 5.777a). It was only after the middle of the century that Putonghua came to be the official designation for the national language of the People’s Republic of China.

Central Asian monks and translators arrived to live and work in East Asia.

v. In premodern times, Sinitic speakers normally referred to their language simply by such terms as *yan* 言, *yu* 語, *hua* 話, and *wen* 文.<sup>13</sup> There rarely was a need to specify it as their own language in contrast to the language of some alien people who spoke a completely different language. It was only in modern times, when the tempo and quantity of linguistic contacts with peoples from abroad (especially those coming from a great distance) increased dramatically, that Sinitic speakers were obliged to clearly identify their own languages as distinct from the languages of the foreigners.

vi. Until the rise of modern nationalistic consciousness, it was usually foreign speakers who felt the need to cap *yan*, *yu*, *hua*, and *wen* with a modifier to distinguish them as different from their own languages. Such, apparently, was the case with the earliest occurrences of *hanhua* 漢話, *zhongyu* 中語, *zhongguoyu* 中國語, *zhongguohua* 中國話, *tangyu* 唐語, *tangrenhua* 唐人話, *qingyu* 清語, *zhinayu* 支那語, *huayu* 華語, and so forth (Huang 2009b).

vii. All of the problems discussed in this section will become much clearer when the papers of the pathbreaking “Critical Han Studies Conference and Workshop” held at Stanford University from April 25-27, 2008 are published.<sup>14</sup>

viii. It has been proposed that *Hanyu* be straightforwardly rendered into English as Hanic or Hannic (Aymat 2009). For all of the difficulties alluded to in this section, however, it would seem prudent to avoid this very recent appellation for the entire language group/family in favor of Sinitic, which has the advantage of being a neutral designation that is easily combined with other group names if necessary (Sino-).

---

<sup>13</sup> In periods of intense interaction between Sinitic and non-Sinitic speakers, already in antiquity distinguishing modifiers might be added to self-referential attributes. For example, when the Xiongnu (Huns) impinged heavily upon the EAH, terms such as *Qinren* 秦人 ‘Qin person/people’ and *qinyu* 秦語 ‘Qin language’ arose. Such terms, however, typically did not last much beyond the time of most intense interaction. It was only in modern times, when massive, sustained, cultural encroachment was operative, that the need was felt for a distinguishing set of autonomous terms.

<sup>14</sup> Among the most important papers delivered at the conference and discussed at the workshop is Mark Elliott’s revelatory “*Hushuo* 胡說: The Northern Other and Han Ethnogenesis,” which independently corroborates many of the points made in this paper. The title of Elliott’s paper speaks for itself.

## 4. Protocols

i. With regard to linguistic classification, “Chinese” contradictorily means at least two quite different things:

- a. the national language of the People’s Republic of China (i.e., *Putonghua*)
- b. the entirety of all non-Tibeto-Burman, non-Austronesian, non-Austro-Asiatic, non-Hmong-Mien (Miao-Yao), non-Tai-Kadai (Kra-Dai), non-Indo-European, non-Turkic, non-Mongolic, non-Tungusic lects in China

ii. To avoid confusion, in this study “Chinese” (when used at all) is reserved for the modern national language of the People’s Republic of China (i.e. MSM), whereas “Sinitic” shall be used to refer to the mass of hitherto inadequately differentiated (yet presumably genetically related) lects specified in 4.i.b.

iii. Sinitic may be divided temporally (all ranges are rough approximations) into Modern Sinitic (MS, 13<sup>th</sup> c.-present), Middle Sinitic (MS, 3<sup>rd</sup> c.-12<sup>th</sup> c.), Old Sinitic (OS, 10<sup>th</sup> c. BC-2<sup>nd</sup> c. AD), and Archaic Sinitic (AS, 11<sup>th</sup> c. BC and before), with the following subdivisions: Early Modern Sinitic (EMoS, 13<sup>th</sup> c.-19<sup>th</sup> c.), Late Modern Sinitic (LMoS, 20<sup>th</sup> c.-present); Early Middle Sinitic (EMiS, 3<sup>rd</sup> c.-6<sup>th</sup> c.), Late Middle Sinitic (LMiS, 7<sup>th</sup> c.-12<sup>th</sup> c. AD); Early Old Sinitic (EOS, 10<sup>th</sup> c.-3<sup>rd</sup> c. BC), Late Old Sinitic (LOS, 2<sup>nd</sup> c. BC-2<sup>nd</sup> c. AD); Early Archaic Sinitic (EAS, 14<sup>th</sup> c. BC and before), Late Archaic Sinitic (LAS, 13<sup>th</sup> c.-11<sup>th</sup> c. BC).

iv. Our comprehension of the historical development of the regional subgroupings of Sinitic is still very limited, but eventually we should aim to divide them into Old, Middle, and Modern phases, as well as to determine the relationship of these stages to the overall development of the SLG/F as a whole.

v. In this study, “dialect” is divorced from *fangyan*; the English word and the Chinese word are not considered as coterminous.

## 5. Propositions

### i. The SLG/F is not an undifferentiable monolith.

I adopt SLG/F as a neutral designation for the mass of genetically related lects specified in §4.i.b and remain agnostic about whether the SLG/F is actually a family unto itself or whether it is more or less closely linked to some other group(s)—such as Tibeto-Burman or Austronesian—in a family (see 5.ii). As a matter of fact, tremendous progress has been made in the analysis of the branches of the SLG/F by Jerry Norman, W. South Coblin, Anne Yue-Hashimoto, Richard Van Ness Simmons, David Prager Branner, Laurent Sagart, William H. Baxter, and others. Their work, unfortunately, is partially vitiated by inadequate terminology, since they refer to all cladistic levels of the SLG/F

as “dialects”, making it difficult to distinguish clearly the hierarchical relationships among the constituent branches, languages, and genuine dialects and sub-dialects of the group.

*Corollary:* The SLG/F is internally classifiable.

*Caveat:* There is no consensus on the internal classification of the SLG/F.

**ii. The SLG/F is not isolated from all other language groups.**

Proposals for linking the SLG to Tibetan, Tibeto-Burman, Tai, Austronesian, Austro-Asiatic, Caucasian, Na-Dene, and other families and groups have been made (Wang 1995, Handel 2008, Behr 2005, van Driem 2007, Sagart 1994, Sagart, Blench & Sanchez-Mazas 2005, Schuessler 2007, Thurgood 2008, and Thurgood & LaPolla 2003), but none of these proposals has been argued with sufficient rigor and supported by adequate data to gain broad acceptance. I maintain that, until the internal classification of the SLG/F is worked out, it is both premature and impractical to speculate on the external classification of the SLG/F. The fact that the proposed external genetic relationships for Sinitic vary so wildly is a good indication of how far we are from achieving a convincing solution to the problem of the affinities of the SLG/F. Whether the Sinitic languages turn out to constitute a family of their own or a group belonging to a larger family, they surely have some sort of relationships—be they genetic or contact—with languages in other families or groups, but it is futile to attempt to determine the external affinities of Sinitic until the tasks outlined in §4.iv are resolved.

*Corollary:* The SLG/F is externally classifiable.

*Caveat:* There is no consensus on the external classification of the SLG.

## 6. Proposals

i. The significance of mutual (un)intelligibility should not be ignored in discussions of the classification of the SLG/F.

*Corollary:* If two lects are completely or largely mutually unintelligible, it makes little sense to refer to them as dialects of each other or of a higher-level entity.

ii. The scientific classification of languages should not be held hostage to extra-linguistic political and cultural prejudices.

*Corollary:* Dogmatic assertions of identity or minor, insignificant variations are unacceptable in the face of a high degree of lexical, phonological, grammatical, and other types of dissimilarity.

iii. In the scientific classification of modern Sinitic languages, as much as possible, data should be drawn from the strikingly different spoken varieties, not from standard written forms. Writing is a second-order linguistic phenomenon. Since most speakers of

Sinitic throughout prehistory and history have been illiterate, the nonessentiality of writing for the existence of the SLG/F is self-evident.

iv. A better Sinitic translation of ‘dialect’ than *fangyan* 方言 would be *xiangyan* 相言 ‘mutual lect’ or *tongyan* 通言 ‘intelligible lect’, since the basic meaning of the English word derives from Greek *dialegesthai* ‘to converse’ < *dia-* ‘one with another’ + *legesthai*, the middle voice of *legein* ‘to tell’.

v. In traditional usage, perhaps *xiangtan* 鄉談 ‘village talk’<sup>15</sup> may be considered as a rough, functional equivalent to ‘dialect’. Other somewhat comparable terms are *suyu* 俗語 ‘common/popular saying’, *liyu* 里語 ‘neighborhood speech’, *liyu* 俚語 ‘rustic/vulgar speech’, *tuhua* 土話 or *tuyu* 土語 ‘patois’, and so forth. However, like *fangyan*, none of these terms are suitable for use in modern, scientific, linguistic classification.

vi. The neologisms *yuxi* 語系 ‘family’, *yuzu* 語族 ‘group’ and *yuzhi* 語支 ‘branch’ are perfectly serviceable for modern, scientific analysis.

vii. It is becoming increasingly common for Western linguists to make the following sort of disclaimer in their publications: “Although ‘Chinese’ is commonly thought of as a single language made up of countless ‘dialects,’ linguistically it is more accurate to describe ‘Chinese’ as an extremely complex and diverse group of related *languages* that is comparable to Romance or Germanic in time, depth and scope. These languages that make up the ‘Chinese’ group may be classified into various subgroups; in the modified traditional scheme to which I adhere, these subgroups are called ‘major dialects.’ Languages belonging to the different ‘major dialects’ are mutually unintelligible, and even within the ‘major dialects’ there may be a high degree of internal diversity. A few scholars refer to ‘Chinese’ as Sinitic because they rightly recognize it as a group (or family) of languages rather than a single language. However, in this book/article/paper, I shall continue to follow common convention—based on Chinese cultural and political beliefs—that the various languages of the group are ‘dialects.’” While this represents a definite improvement over the time when scholars were so cowed by tradition that they would not dare to question the dogma that, for China alone, language and dialect are taxonomically undifferentiable, it is still tantamount to saying that, although one realizes languages are not conterminous with dialects, one will doggedly persist in the false assumption that they are identical. One of the chief purposes of the present study is to help my colleagues confront linguistic reality more directly and unabashedly.

viii. If Sardinian and Italian—which are written with the same script, found in the same country, and belong to the same group (Romance)—are considered to be two different languages, all the more so Mandarin and Cantonese—one of which is written

<sup>15</sup> This expression is found already in Zhu Xi’s *Zhu Zi yulei* [Classified Conversations of Master Zhu], 140 and in the great Ming vernacular novel, *Shuihu zhuan* 水滸傳 [Water Margins], 61. See *HDC*, 10.670b.

with Chinese characters and the other is normally not written (but may be with a combination of the regular character set, over a thousand special characters, and conspicuous use of letters from the Roman alphabet), are found in the same country, and belong to the same group (Sinitic)—should be considered as two different languages. The differences between spoken Cantonese and spoken Mandarin are greater than those between spoken Sardinian and spoken Italian. Similar criteria obtain for Catalan and Spanish, and for Scots and English.

## 7. Clarifications

i. A pervasive fallacy asserts that “all the [Sinitic] ‘dialects’ are the same when written down.” Nothing could be further from the truth. As a matter of fact, the vast majority of the countless varieties of Sinitic have never received a written form, and, indeed, are difficult—if not impossible—to write with the standard set of characters (even in its largest incarnation). When, as a sort of tour de force, they exceptionally do get written down in their full, unadulterated form, as has happened with Cantonese and Taiwanese from time to time, the results are intelligible only to individuals who are literate according to the special conventions of the relevant language.

ii. The tremendous differences between the various Sinitic lects and the standard written language is borne out by the absence of suitable characters for writing numerous basic morphemes, even among the so-called Mandarin dialects. The traditional concept of *benzi* 本字 (“original characters”) for missing morphemes constitutes a flawed policy that fails to take into account borrowings from non-Sinitic origins, substrate languages, and other unknown sources.

iii. Throughout the history of the SLG/F, there have basically been only two types of written language:

- a. Classical/Literary (divinatory, inscriptional, documentary, etc.—starting from the 13<sup>th</sup> c. BC and continuing up to modern times)
- b. Vernacular (beginning with snatches of the *koine* in medieval times and culminating with full acceptance in the early 20<sup>th</sup> c.)

Neither of these types of written Sinitic directly and fully reflects any particular speech form tied to a specific time and place, although Vernacular Sinitic (VS) is closer to certain forms of spoken language than is Classical Sinitic (CS) and Literary Sinitic (LS), which tend to be highly elliptical and aggressively terse.

iv. The oft-quoted quip that the difference between a language and a dialect is that the former has an army and a navy while the latter does not is simply silly and should not be adduced in any serious discussion of linguistic classification. There are countless counterexamples that may be brought forward to negate it: French in Quebec; French,

German, Italian, and Rumantsch in Switzerland; French and Flemish in Belgium; Amerindian languages in the United States—all are recognized as independent languages but all lack armies and navies of their own; English and French as the languages of numerous sovereign states; Portuguese in Portugal and Brazil; Spanish in Spain and in many Central American and South American nations—all are dialects of major languages, yet all belong to countries that possess armies and navies of their own.

v. “Missionaries in the 19<sup>th</sup> century produced Bible translations in Amoy, Foochow, Kienow, Kienyang, and Shaowu (perhaps others as well), showing that one could not use a single ‘Min’ translation that would work for all these different speech areas. Actually, there are possibly twenty or thirty non-mutually intelligible forms of Min alone” (Norman 2009a). North of Fuzhou in Fuan 福安 county, the Catholic missionaries had a romanized form of the local language, which they called in Spanish *Foganes* (Norman 2009b). Church Romanization has been widely and successfully utilized for Hokkien (Amoy and Taiwanese) (A. Lin 1999, and C. Lin 1999). If Sinitic is accepted as a family, then Southern Min, Northern Min, and Eastern Min would count as groups within it, and they would be divisible into branches; if Sinitic is accepted as a group, then Southern Min, Northern Min, and Eastern Min would count as branches within it, and they would be divisible into languages. Similar observations would hold for Wu and other major clusters of Sinitic languages.

vi. The contentious, non-scientific nature of the debate over the SLG/F is manifest in the circumlocutions used to designate its constituent members: “speech forms,” “varieties,” “styles,” “regionalects,” “dialects” (no matter how far up or down the taxonomic scale one may go), and so forth. At the same time, scholars openly admit that the main reasons why they do not use normal linguistic terminology (family, group, branch, language, dialect) in dealing with the SLG/F are due to sociopolitical and cultural factors. The fallacy of such a bizarre approach is evident when one considers that all nations have special sociopolitical and cultural circumstances, yet an impartial analytical outlook does not allow such circumstances to interfere with pure linguistic research.<sup>16</sup>

---

<sup>16</sup> Of course, it will be objected that “there is no such thing as ‘pure linguistic research’” or an impartial, objective outlook. Granted, but I maintain that we should strive to our utmost to come as close to these desiderata as possible, not to throw all standards to the winds because the phenomena we study are complicated and challenging. Think of how complex the human genome is (not to mention the genomes of many other organisms). Nonetheless, dedicated scientists have mapped the entire genomes of human beings and other organisms. We should be no less dedicated in our own research on Sinitic languages, their history, their affinities, and their nature. The microscopic existence of continua does not negate the macroscopically classifiable units (comparable to phyla, genera, and species).

## 8. Conclusions

i. The classification of the SLG/F, both internally and externally, is still in the beginning stages; much difficult work remains to be done.

ii. ‘Chinese’, ‘dialect’, and other terms in broad popular usage should be employed with extreme caution in technical discussions of the countless varieties of speech forms that currently exist and that have existed at various periods and places during the past in the East Asian Heartland (EAH) and Extended East Asian Heartland (EEAH) (Mair 2005a).

iii. The approximately 1.4 billion people of China do not all share the same language. Rather, they speak a complex mixture of many different tongues, not all of which belong to the SLG/F.

iv. The billion or so members of the *Hanzu* do not all share the same language. Rather, they speak a combination of numerous different languages and dialects that are more or less mutually (un)intelligible.

v. The 867,200,000 individuals who are claimed to be speakers of Mandarin (when examined on the ground in specific locales—not in the abstract as imagined speakers of MSM) actually speak a Babel of tongues that are more or less mutually (un)intelligible. No attempt has been made to determine the number of speakers of MSM, but it is unlikely to constitute as much as a third or half of the total number claimed for Mandarin.

vi. The linguistic landscape of China is every bit as diverse and complex as that of India. Both are modern nation states with national languages (MSM and Hindi) and both have rich cultural histories, including extensive development of their classical languages (Sanskrit and LS). Politically, although one is the world’s largest democracy and the other is the world’s largest country ruled by a single party, they are both sovereign states. The whole world acknowledges that there are many Indic languages in India (Hindi, Bengali, Marathi, Gujarati, Oriya, Assamese, Dogri, Konkani, Maithili, Nepali, Punjabi, Sindhi, and Urdu), not to mention numerous non-Indic languages (Tamil, Telegu, Kannada, Malayalam, Bodo, Manipuri, Santali, and Kashmiri), just as there are numerous non-Sinitic languages in China. The differences among the separate languages of the SLG/F (Cantonese, Taiwanese, Shanghainese, Sichuanese, Hakka, etc.) are every bit as great as those among the Indic group of languages, if not more so, yet scholars and laypersons alike persist in calling the various Sinitic languages ‘dialects’, while readily admitting that Hindi, Bengali, Marathi, etc. are separate languages. There is no linguistic justification for this singularly peculiar usage.

vii. It is unthinkable that an allegedly genetically related linguistic assemblage of such great geographical scope, chronological depth, and burgeoning number of speakers as Sinitic would have only a single form throughout time and space, with but minor



variations ('dialects').

viii. There is an urgent need for the classification of the SLG/F, but this cannot be accomplished satisfactorily without precise, linguistically justifiable terminology.

ix. The people of China have a right to conceive and speak of the languages of their country however they wish; linguists of the world have a duty to study the languages of China according to universal principles. If linguists abandon their scientific duty, the current chaos and lack of consensus concerning the nature of Sinitic will continue, much to the detriment of our understanding not only of the languages of China, but to linguistics as a whole.

x. It is time for linguists to openly declare that there exists a multiplicity of Sinitic languages (not just dialects) and to act upon this reality in their analysis and classification, both internally and externally to the language group/family.

## References

- Aymat, Ivan. 2009. Personal communication. March 9.
- Beck, B. J. Mansvelt. 1986. The fall of Han. *The Cambridge History of China*, Vol. 1: *The Ch'in and Han Empires, 221 B.C. - A.D. 220*, ed. by Denis Twitchett & Michael Loewe, 317-376. Cambridge: Cambridge University Press.
- Behr, Wolfgang. 2005. Genealogie, Phonologie und Morphologie des Antikchinesischen. *Antikchinesisch: Ein Lehrbuch in drei Teilen*, by Robert H. Gassmann & Wolfgang Behr, 408-445. Bern & New York: Peter Lang.
- Broughton, Jeffrey L. 1999. *The Bodhidharma Anthology: The Earliest Records of Zen*. Berkeley & Los Angeles: University of California Press.
- Campbell, Lyle. 1997. *American Indian Languages: The Historical Linguistics of Native America*. Oxford & New York: Oxford University Press.
- Casad, Eugene H. 1974. *Dialect Intelligibility Testing*. Norman, Oklahoma: Summer Institute of Linguistics at the University of Oklahoma.
- Chao, Yuen Ren. 1983. *Tongzi Fang'an [A Project for General Chinese]*. Beijing: The Commercial Press.
- Chappell, Hilary. (ed.) 2001. *Sinitic Grammar: Synchronic and Diachronic Perspectives*. Oxford & New York: Oxford University Press. Republished in 2004 by Oxford University Press as *Chinese Grammar: Synchronic and Diachronic Perspectives*.
- Chen, Sanping. 2008a. Personal communication. April 14.
- Chen, Sanping. 2008b. Personal communication. April 10.
- Cheng, Chin-Chuan. 1982. A quantification of Chinese dialect affinity. *Studies in the Linguistic Sciences* 12.1:29-47.

- Cheng, Chin-Chuan. 1991. Quantifying affinity among Chinese dialects. *Languages and Dialects of China*, ed. by William S-Y. Wang, 78-112. Berkeley: Journal of Chinese Linguistics.
- Cheng, Chin-Chuan. 1998. Extra-linguistic data for understanding dialect mutual intelligibility. Paper presented at the 1998 Annual Conference of the Pacific Neighborhood Consortium (PNC), May 15-18, 1998. Taipei: Academia Sinica.
- Coblin, W. South. 2009. Personal communication. January 17.
- DeFrancis, John. 1984. *The Chinese Language: Fact and Fantasy*. Honolulu: University of Hawai'i Press.
- Doak, Kevin M. 2007. *A History of Nationalism in Modern Japan: Placing the People*. Leiden & Boston: Brill.
- Driem, George van. 2007. The diversity of the Tibeto-Burman language family and the linguistic ancestry of Chinese. *Bulletin of Chinese Linguistics* 1.2:211-270.
- Egerod, Søren. 1988. Sino-Tibetan Languages. *The New Encyclopaedia Britannica*, Vol. 22, 721-731. Chicago: Encyclopaedia Britannica.
- Fogel, Joshua A. (ed.) 2005. *The Teleology of the Modern Nation-State: Japan and China*. Philadelphia: University of Pennsylvania Press.
- Gassmann, Robert H., and Wolfgang Behr. 2005. *Antikchinesisch: Ein Lehrbuch in drei Teilen*. Bern & New York: Peter Lang.
- Handel, Zev. 2008. What is Sino-Tibetan? Snapshot of a field and a language family in flux. *Language and Linguistics Compass* 2.3:422-441.
- Huang, Heqing. 2009a. Personal communication. February 4.
- Huang, Heqing. 2009b. Personal communication. February 10.
- Huang, Heqing. (forthcoming). 'Hanren,' 'Hanzu,' 'Hanyu,' 'Hanzi' ciyuan luekao [A brief study on the etymology of 'Han Person,' 'Han Nationality,' 'Han Language,' and 'Hanogram'].
- Lass, Roger. 1997. *Historical Linguistics and Language Change*. Cambridge & New York: Cambridge University Press. See especially Chapter 3: "Relatedness, Ancestry, and Comparison."
- Lin, Alvin. 1999. Writing Taiwanese: The Development of Modern Written Taiwanese. *Sino-Platonic Papers* 89.
- Lin, Christine Louise. 1999. The Presbyterian Church in Taiwan and the Advocacy of Local Autonomy. *Sino-Platonic Papers* 92.
- Lu, Xun. 1934[2005]. An outsider's chats about written language, trans. by Victor H. Mair. *Hawai'i Reader in Traditional Chinese Culture*, ed. by Victor H. Mair, Nancy S. Steinhardt & Paul R. Goldin, 617-641. Honolulu: University of Hawai'i Press.
- Mair, Victor H. 1991. What is a Chinese 'dialect/topolect'? Reflections on some key Sino-English linguistic terms. *Sino-Platonic Papers* 29:1-31.

- Mair, Victor H. 1994. Buddhism and the rise of the written vernacular in East Asia: the making of National Languages. *Journal of Asian Studies* 53.3:707-751.
- Mair, Victor H. 2001. Introduction: the origins and impact of literati culture. *The Columbia History of Chinese Literature*, ed. by Victor H. Mair, 1-15. New York: Columbia University Press.
- Mair, Victor H. 2005a. The north(west)ern peoples and the recurrent origins of the 'Chinese' state. *The Teleology of the Modern Nation-State: Japan and China*, ed. by Joshua A. Fogel, 46-84. Philadelphia: University of Pennsylvania Press.
- Mair, Victor H. 2005b. Introduction: a constantly shifting mosaic of peoples and cultures. *Hawai'i Reader in Traditional Chinese Culture*, ed. by Victor H. Mair, Nancy S. Steinhardt & Paul R. Goldin, 1-7. Honolulu: University of Hawai'i Press.
- Norman, Jerry. 1988. *Chinese*. Cambridge & New York: Cambridge University Press.
- Norman, Jerry. 2009a. Personal communication. January 29.
- Norman, Jerry. 2009b. Personal communication. March 3.
- Ramsey, S. Robert. 1987. *The Languages of China*. Princeton: Princeton University Press.
- Sagart, Laurent. 1994. Proto-Austronesian and Old Chinese evidence for Sino-Austronesian. *Oceanic Linguistics* 33.2:271-308.
- Sagart, Laurent, Roger Blench, and Alicia Sanchez-Mazas. (eds.) 2005. *The Peopling of East Asia: Putting together Archaeology, Linguistics and Genetics*. London & New York: RoutledgeCurzon.
- Sanders, Robert M. 1987. The four languages of 'Mandarin'. *Sino-Platonic Papers* 4:1-14.
- Schuessler, Axel. 2007. *ABC Etymological Dictionary of Old Chinese*. Honolulu: University of Hawai'i Press.
- Tang, Chaoju, and Vincent J. van Heuven. 2009. Mutual intelligibility of Chinese dialects experimentally tested. *Lingua* 119.5:709-732.
- Thurgood, Graham. 2008. Han-Zang yuxi: puxi chuancheng, wenhua jiechu, yiji lishi bianhua [The Sino-Tibetan languages: genetic inheritance, external contact, and change]. *Minzu Yuwen [Minority Languages of China]* 2008.2:3-15.
- Thurgood, Graham, and Randy J. LaPolla. (eds.) 2003. *The Sino-Tibetan Languages*. London: Routledge.
- Wang, William S-Y. (ed.) 1991. *Languages and Dialects of China*. Berkeley: Journal of Chinese Linguistics.
- Wang, William S-Y. (ed.) 1995. *The Ancestry of the Chinese Language*. Berkeley: Journal of Chinese Linguistics.
- Wilkinson, Endymion P. 2000. *Chinese History: A Manual*. Revised and Enlarged. Cambridge: Harvard University Asia Center.
- Xu, Wenkan. 2009. Personal communication. February 11.

Victor H. Mair

Yao, Dehuai. 2009. Personal communication. February 11.

Zhao, Gang. 2009. *A Marginal Identity and its Modern Transformation: Inquiring into Han Identity*. Manuscript.

Zhou, Youguang. 2003. *The Historical Evolution of Chinese Languages and Scripts*. Bilingual edition of *Zhongguo Yuwen de Shidai Yanjin* (trans. by Zhang Liqing). *Pathways to Advanced Skills*, Vol. VIII. Columbus: National East Asian Languages Resource Center, The Ohio State University.