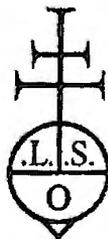VICTOR H. MAIR

ON "TRANSFORMATIONISTS" (*BIANJIA*) AND "JUMBLED TRANSFORMATIONS" (*LAZA BIAN*): TWO NEW SOURCES FOR THE STUDY OF "TRANSFORMATION TEXTS" (*BIANWEN*)

WITH AN APPENDIX ON THE PHONOTACTICS OF THE SINOGRAPHIC SCRIPT AND THE RECONSTRUCTION OF OLD SINITIC

VICTOR H. MAIR

ON "TRANSFORMATIONISTS" (*BIANJIA*)
AND "JUMBLED TRANSFORMATIONS" (*LAZA BIAN*):
TWO NEW SOURCES FOR THE STUDY
OF "TRANSFORMATION TEXTS" (*BIANWEN*)*

WITH AN APPENDIX ON
THE PHONOTACTICS OF THE SINOGRAPHIC SCRIPT
AND THE RECONSTRUCTION OF OLD SINITIC

INTRODUCTION

The most important Chinese genre for understanding the dynamics of Sino-Indian popular Buddhist narrative is *bianwen* ("transformation texts"). Although *bianwen* have been studied in-

---

* My original intention was to undertake a general study of Buddhist narrative literature in Central Asia and China so as to complement Michael Hahn's splendid contribution to this conference. Unfortunately, as I pursued my research, I soon realized that the subject was far too large to encompass in a paper for a conference volume. Thus, I had to abandon that project even before I travelled to Venice. At the conference itself, the paper that I actually delivered was concerned with oral and visual aspects of Sino-Indian Buddhist narratives and concentrated primarily on a single manuscript from Dunhuang (P4524). In the process of revising the paper for publication, it became both too bulky and too convoluted for the present volume. Thus I decided to have the oral-visual narrative paper published in the Sinological journal «Asia Major» under the title "Śāriputra Defeats the Six Heterodox Masters: Oral-Visual Aspects of an Illustrated Transformation Scroll (P4524)." A quite different version of that paper will appear in a book edited by Jean-Pierre Drège about illustrated manuscripts from Dunhuang preserved in the Bibliothèque Nationale de France.

tensively by scholars from around the world since their discovery in a cave storeroom at Dunhuang (far western Gansu province) around the turn of the century,[1] there is still much that is not known about their origins, social context, and oral-visual-literary nature. Because transformation texts belonged to the popular realm and their oral antecedents were even more folkish, the literati not only were uninterested in them, often they ridiculed them or even attempted to suppress them. Consequently, the historical record

---

N.B.: Throughout, I enclose the word "radical" in quotation marks when it applies to a component of sinographs because it is not really the etymological root of a word but rather a semantic classifier or categorizer (which terms I avoid because of their clumsiness). Similarly, I enclose the word "Altaic" in quotation marks because, although it is a convenient designation for referring to Turkic, Mongolian, Tungusic, and other similar languages, it is no longer widely accepted in its entirety as a viable language family by critical linguists. The most recent attempt to resuscitate "Altaic" is that by Joseph H. Greenberg.

[1] Extensive bibliographical references may be found in MAIR, *Partial Bibliography* and in MAIR, *Chinese Popular Literature from Tun-huang.*

concerning *bianwen* is extremely scanty. What we know about this genre must be pieced together from tiny bits of widely scattered evidence. Therefore, whenever new sources pertinent to *bianwen* become available, they are cause for celebration. The purpose of this paper is to introduce and analyze two hitherto unknown sources for the study of *bianwen*.

"TRANSFORMATIONISTS" (BIANJIA)

The late Tang Zen master Zhaozhou (778-897?)[2] was noted for his sharp ripostes. A typical exchange from the records of his teachings goes as follows:

[A monk] asked, "At night he ascended to Tuṣita Heaven,[3] in the day he descended to Jambūdvīpa.[4] While he was doing that, why didn't the *maṇi*-jewel[5] manifest itself?"
The master said, "What did you say?"
The monk asked the same question again.
The master said, "The Buddha Vipaśyin[6] has been mindful of this from a long time ago, but he still hasn't got the knack."[7]

The monk's question is an allusion to the great North Indian Buddhist scholar named Asaṅga[8] (4th-5th c.). Born in the city of Puruṣapura in the region of Gandhāra, he was the oldest son of Kauśika, a Brahmin. His younger brother and one of his most outstanding disci-

---

[2] He gets his name from the place in Hebei province whence he supposedly hailed. *Foguang da cidian*, vol. 6, pp. 5934b-5935a.

[3] The abode of Maitreya, Buddha of the future.

[4] The continent situated to the south of Mt. Meru (*axis mundi*).

[5] A luminous pearl, the symbol of the Buddha and his teachings.

[6] First of the seven Buddhas of antiquity.

[7] SUZUKI, ed., *Chōshū*, scroll B, p. 61, no. 359; Japanese translation by Akizuki on p. 40a.

[8] *Foguang da cidian*, vol. 6, pp. 5126c-5127a.

ples was Vasubandhu. In order to realize the basic meaning of Ma-hāyāna meditation on the unreality of the ego and phenomena, Asaṅga was said to have gone up to the Tuṣita Heaven at night to receive the teachings of Maitreya, then come back down to the human world during the day to write out the *Yogācāryabhūmi-śāstra*. Why, the monk asks, did not supreme Buddhist knowledge manifest itself in Asaṅga's own person while he was engaged in assiduous cultivation? That is to say, why did he have to travel up to heaven to seek it outside of himself from Maitreya?

The master's first response actually implies something like "**Who** are **you** to ask such a question?" That is to say, "Are you at a stage where you are qualified to ask such a question or could you make sense of an illuminating answer to it?" In essence, the master is telling the foolish monk to leave questions about the exalted likes of Asaṅga well enough alone and concentrate on his own improvement. When the monk repeats his vapid question, the master gives him a blast that surely must have left the poor fellow dumbfounded for days.

To the noniniate, the master's second response is much more difficult to handle. Without elucidation, it is impossible to understand what Zhaozhou really meant. Fortunately, the celebrated scholar-monk, Zongmi[9] (780-841) has an annotation in his *Yuanjue jing da shuchao [Major Transcription of Subcommentaries on the* Sūtra *of Perfect Awareness]* which provides the necessary context for coming to grips with Zhaozhou's allusion to Vipaśyin.[10]

The following is from Zongmi's commentary on the concept of

---

[9] Considered to be the Fifth Patriarch of the Avataṃsaka School in China, he hailed from Xichong in Sichuan Province. Deeply involved in the transmission of Zen, he was popularly known as Zen Master of Guifeng (Guifeng chanshi). *Foguang da cidian*, vol. 4, pp. 3158b-3159b. For two valuable book-length studies on Zongmi, see PETER GREGORY, *Sinification* and *Inquiry*. Zongmi's biography may be found in chapter 2 of the former.

[10] The passage in question occurs in *Zoku zōkyō*, vol. 15, scroll 13A.25ab-ba. For Zongmi's special relationship to the *Yuanjue jing*, see GREGORY, *Sinification*, pp. 54ff and 167ff. The *Sūtra of Perfect Awareness* was supposed to have been translated into Chinese by the Kashmiri monk, Buddhatrata, who was in Loyang before 730. The modern scholarly consensus, however, is that it is an apocryphal scripture composed in China.

erye ("two karmas").[11] To provide the necessary context, I shall quote a rather long passage. Our interest, however, is focused on the enigmatic reference to the Buddha Vipaśyin which occurs near the end and which is virtually identical to the concluding clauses of the quotation from the Zen master Zhaozhou above.

"To be skillful at making the six paths of rebirth"[12] means that the shapes and appearances of all the beings on the six paths of rebirth are each [caused to be] different. The different species and types numberless as particles of dust and sand are all due to the power of karma. Therefore, the *Avataṃsaka-sūtra* says, "The mind is like a painter who can paint the various worlds. The five aggregates[13] are attendant upon life; it is impossible not to create them." And, again, it says, "If a person wishes to understand all[14] the Buddhas of the three worlds,[15] he should observe the nature of the dharma-realm.[16] Everything is created by the mind alone,[17] thus Buddhas can be painted too."

As for just the path of humans and, within that, take only the people of Jambūdvīpa,[18] already there are so many faces, tens of thousands of them, each of which is different. And not only are they different at the present horizontal moment in time, vertically in the past and in the future, human faces are each distinct.

---

[11] For the various pairs signified by this term, see SOOTHILL and HODOUS, *Dictionary*, p. 26a.

[12] Skt. *ṣaḍ gati*; these are the paths of hell (*naraka*), hungry ghosts (*preta*), birds, beasts, fish, worms, etc. (*tiryagyoni*), demons (*asura*), humans (*manuṣya*), and heavenly deities (*deva*).

[13] 13. Skt. *pañcaskandha*; these are form (*rūpa*), sensation (*vedanā*), discerning (*sañjñā*), the function of mind concerning dichotomies (*saṃskāra*), and recognition (*vijñāna*).

[14] Skt. *sarva*.

[15] Skt. *try-adhvan, try-adhvaka, trayaḥ kālāḥ, traikālya, adhva-traya, loka-traya*, etc.; these are past, present, and future (Skt. *atīta, pratyutpanna, anāgata*).

[16] Skt. *dharmadhātu*; "things" in general, both numerical and phenomenal, or the underlying and unifying absolute reality of everything.

[17] Skt. *vijñānamātra, cittamātra*; this is the doctrine of idealism, according to which nothing exists apart from mind.

[18] See note 4 above.

2

As for a single person, his great-grandfather was different from him, and his great-great-grandfather and his grandfather were also unlike. The body [19] of his great-grandfather was dissimilar, the body of his grandfather was dissimilar, and even father and son have different bodies. The grandson, again, is different from the son, and we know for certain that an unborn grandson will be distinct.

Since it is thus for one person, it is the same for the whole world, and in each case it is due to the power of karma. Knowing this, things can be skillfully distributed so that each will be distinct. This is indeed difficult to imagine. Therefore, there is a sūtra which states, "Karma can skillfully generate various sorts of shapes."

Just among the people of the southern continent of Jambūdvīpa, it is difficult to describe one form. How much more so for the eastern, western, and northern continents where, again, the shapes and appearances of their faces are even more distinct.

It is said that it is thus for everyone on the path of humans. How much more so for beings on the paths of heavenly deities, ghosts, animals, and hell which are each distinct! [20] Just among animals, there are those that fly and those that run; there are those with feet and those without feet; and so on down to tiny bugs. How can they be described? It's all due to the skillful ability of karma to apportion things. Therefore it is said, "'To be skillful at making the six paths of rebirth' [means] causing each to be different."

The skill of an artist truly cannot compare. Usually when we look at a wall painting with three to five hundred people on it, whether we be donors or connoisseurs, we want to make each and every one of them unalike in terms of being fat, skinny, big and little, swarthy and fair, but it is unavoidable that many of the figures resemble each other. Therefore, Baozhi [21] re-

---

[19] The text, which reads *zengmen wen* and has as a variant *zufu* (a relationship already covered by the colloquial term *aweng* in the sentence just above) is clearly defective. Judging from the following three clauses, this clause should probably be emended to *zengmen shen* (*zengmen* means the same as *zengzu* ["great-grandfather"]).

[20] The author seems to have forgotten *asura* (see note 12 above). For "distinct" at the end of the sentence, the text has *renbie* ("humans are separate"); I follow Iriya in emending to *gebie*.

[21] An eccentric monk who was reputed to be a follower of the Zen patriarch, Bodhidharma. There are many legends about Baozhi who was said to have lived almost a

buked Zhang Sengyou,[22] saying, "The Buddha Vipaśyin has been studying painting from a long time ago, but he still doesn't have the knack." Even though this[23] was composed (*zhuanzuo*) by a transformationist (*bianjia*) who was following popular tradition,[24] it matches a hidden truth.[25]

Let us set aside for the moment the remarkable occurrence of the term *bianjia* which is the actual subject of this investigation. Having read through Zongmi's long and rambling note, we are at least aware that the quotation concerning Vipaśyin has to do with karmic differentation of individuals and the unsuccessful attempts of artists to capture its full range and variety. In addition, we now know that the gibe about Vipaśyin was directed at the sublimely outstanding artist Zhang Sengyou by the mysterious monk Baozhi.

Both Zhang Sengyou and Baozhi were active during the early part of the Liang dynasty, but there is no firm authority for the existence of this enigmatic legend about them in contemporary sources. By the latter part of the Tang period, Baozhi was being venerated as an avatar of the Eleven-faced Guanyin (Ekadaśamukha-avalokiteśvara-bodhisattva). This we learn from an entry, dated to the sixth day of the fourth moon in the year 840, in the diary of the Japanese monk Ennin (794-864), who travelled extensively in Tang China.[26] Furthermore, in a work completed in the year 1333, the Buddhist historiographer Nianchang (1282-1341) recorded a legend according to

---

hundred years (418-514). GILES, *Biographical Dictionary*, #1622; *Foguang da cidian*, vol. 7, p. 6759a-c.

[22] A famous painter of the Liang period who was renowned for his supernatural talent, he was especially favored during the reign of Emperor Wu. Hailing from the area of Wu in the south, he was particularly good at painting landscapes and Buddhist subjects. GILES, *Biographical Dictionary*, #99.

[23] Namely this tale about Baozhi rebuking Zhang Sengyou by referring to Vipaśyin.

[24] In Buddhism, the term *suisu* also has the technical meaning of *saṃketa* ("convention, agreement") or *saṃvṛti* ("dissimulation") and a more general meaning of "follow worldly (as opposed to spiritual / ideal / religious) ways."

[25] More literally, "latency matches principle" (*an yu li fu*).

[26] REISCHAUER, tr., *Ennin's Diary*, p. 202.

which Emperor Wu of the Liang ordered Zhang Sengyou to paint a portrait of Baozhi in the year 503.[27] Thereupon, Baozhi pointed at his own face with his finger and revealed himself in the form of a Twelve-faced Guanyin.[28] Whether this was intended to reveal Baozhi's compassion or awesomeness, since Buddhist iconography does not include a Twelve-faced Guanyin this was so unsettling that Zhang Sengyou could not finish the portrait.

After citing this curious legend about Baozhi and Zhang Sengyou, Zongmi concludes by stating that, even though it was composed by a *bianjia* who was following some popular tale, it still conveys a hidden truth. The expression *bianjia* has not been identified in any other source, but the bare mention of it by Zongmi here provides precious information about transformation texts, some of which is not available elsewhere. In the first place, we now know that the authors of transformation texts were called *bianjia* ("transformationists"). Secondly, the *bianjia* did not make up their tales out of whole cloth but depended on popular stories for their material. Third, Zongmi confirms our previous findings about the decidedly non-elite social background of transformation texts. Fourth, since Zongmi, who lived from the late eighth century to the first half of the ninth century, was familiar with the *modus operandi* of transformationists, this corroborates our earlier determination of the eighth century as the heyday of transformations (despite the ninth and tenth century dates of most of the extant manuscripts). Fifth, the legend recounted by Zongmi indirectly reaffirms our understanding of transformations as having a close association with paintings. Sixth, Zongmi's account reveals beautifully how stories that circulated in the realm of popular culture could be taken up by members of the literate elite like Nian-

---

[27] *Fozu lidai tongzai [Comprehensive Records of the Buddhist Patriarchs in Successive Eras]*, scroll 9 (*Taishō Tripiṭaka* [2036]49.544b).

[28] *Foguang da cidian*, vol. 1, p. 342c. The irony of Baozhi appearing in the guise of a Twelve-faced Guanyin is that there is no such figure in Buddhist iconography. Therefore, either he was engaging in one-upmanship with Guanyin (a mind-boggling proposition to contemplate) or he simply wanted to startle Zhang Sengyou. Or perhaps he was just being mischievous.

chang and become part of the written culture that survived above ground.

It is essential to observe that *bianjia* were **not** themselves transformation performers. As I have pointed out on numerous occasions,[29] transformation performers (both male and female) were of low social status and in most cases must have been illiterate. The verb that Zongmi uses to describe what the *bianjia* did, namely *zhuanzuo*, indicates that they were literate. That is to say, it is the *bianjia* who were actually responsible for writing down the stories told by the transformation performers.[30] The *bianjia* may have composed the written versions we know as *bianwen* ("transformation texts"), but they did not create the stories recounted in them. It was the transformation performers who were the real makers of transformation tales (not the written texts).

Although we now know that the writers of transformation texts (*bianwen*) were called "transformationists" (*bianjia*), we still do not know the name of the transformation performers (i.e., the picture storytellers) themselves. I suspect that, if we ever do find it out, their designation will be something like *yan/shuo/jiangbiannü/nan/zhe/jia* and the name for the performance would be something like *yan/shuo/jiangbian*. We already do know one name for transformation performances, namely *zhuanbian* ("turning transformations"), so perhaps the performers would have been called *zhuanbiannü/nan/jia/zhe*, or the like.

"JUMBLED TRANSFORMATIONS" (*LAZA BIAN*)

The second new item for the study of *bian* to be discussed in this paper is from a mini-essay found among the collected prose works of

---

[29] Most recently and most explicitly in the article to appear in «Asia Major» referred to in the unnumbered first note above.

[30] Nor should we confuse the *bianjia* with the copyists of the later manuscripts who were primarily lay students. See MAIR, "Lay Students."

the celebrated Song period author and statesman, Su Shi[31] (1037-1101). In its entirety, the piece reads as follows:

On a Jumbled Transformation
Shu laza bian

Sima Changqing[32] composed the "Rhapsody on the Great Man." When Emperor Wu (r. 140-87 BCE) read it, he had a feeling of soaring up to the clouds and drifting along. Recently, a scholar composed a "Jumbled Transformation" and called himself [another Sima] Changqing. Even though Changqing admittedly won't scold you, I'm afraid that whoever reads [your transformation] will have a hard time soaring up to the clouds, but will instead doze off and fall into bed.[33]

We shall momentarily skip over the title since it (together with the reoccurrence of the term *laza bian* later in the text itself) is not only the most vital component of the mini-essay for the study of transformations, it is also the most difficult part of the piece to interpret correctly.

The mini-essay itself is fairly straightforward and does not require extensive annotation or commentary. It is clear that Su Shi takes a dim view of whoever it was that wrote the "Jumbled Transformation." We have no record of who this person was, nor has the "Jumbled Transformation" been preserved, at least not under that title. As a matter of fact, the author of the "Jumbled Transformation" may have been a political opponent of Su Shi's, someone like Wang Anshi[34] (1021-1086). Surely, no one would name their own composition a "Jumbled Transformation," particularly if they intended to present it to the emperor (!), so Su Shi must have given the piece in question this title in an effort to disparage it. Thus, it is possible that the "Jumbled Transformation" still exists

---

[31] NIENHAUSER, *Companion*, pp. 729a-730b.

[32] This is the celebrated Western Han rhapsodist, Sima Xiangru (179-117 BCE). NIENHAUSER, *Companion*, pp. 723b-725a.

[33] KONG, ed., *Su Shi wenji*, vol. 5, p. 2062.

[34] NIENHAUSER, *Companion*, pp. 854a-855b.

under another title among the collected works of one of Su Shi's contemporaries.

Whoever the author of the "Jumbled Transformation" was, he must have presented it to the Song emperor in an attempt to curry favor with him. The culprit apparently dared to compare himself publicly (perhaps even in the preface to the work in question or in the body of the work itself) with the premier Han rhapsodist, Sima Xiangru. The latter was much favored by the powerful Han emperor, Wudi, who was an ardent fan of Sima's rhapsodies. It was this outrageous claim (viz., that the "Jumbled Transformation" would garner for its author the same sort of imperial patronage that Sima Xiangru's "Rhapsody on the Great Man" had gained for him) that prompted Su Shi to satirize mercilessly the luckless man with this mini-essay.

In general, the language of this particular mini-essay is more colloquial than most of Su Shi's writings. Aside from the earthy expression *laza*, which we shall examine in detail below, there is also the vernacular term *keshui*[a] ("doze off") which dates from the Song and Jin periods (roughly 11th-13th centuries).[35] Judging from the characters with which it is written, it would seem to mean "thirst for sleep," but it was also written with the nearly homophonous characters *keshui*[b] and still today as *keshui*[c] which are incorporated in the usual vernacular Mandarin expressions for "to doze" in the middle latitudes of China.

Su Shi's adoption of such a relatively highly colloquial diction in this mini-essay was probably intentional and, if so, was undoubtedly meant to show his contempt for the "Jumbled Transformation" and its author. Su Shi also makes plain his disdain by choosing to style the piece in question a "transformation." Although artistic "transformation [tableaux]" (*bianxiang*) – as a genre of elite painting – were respectable, "transformation texts" – as a genre of popular literature – would inevitably have been looked down upon by the literati.[36] Indeed, all references to *bian* as a type of prosimetric folk performance

---

[35] GAO WENDA, *Jindai Hanyu cidian*, p. 452b.
[36] MAIR, *Transformation Tableaux*.

and to *bian(wen)* as a genre of popular literature outside of the *bian-(wen)* texts themselves reveal their definite non-elite status.[37] What is most extraordinary is that, as late as the second half of the eleventh century, Su Shi is still obviously aware of the existence of *bian* as a popular literary genre. (References to *bian* after the tenth century are extremely rare, and even before that time they are very scarce.[38]) It would appear, however, that he had only a vague notion of *bian* as a lowly kind of literature. Su Shi no longer seems to understand *bian* in the specific sense of "prosimetric tale about transformational manifestations (usually Buddhist)" that it conveyed during the Tang period. By Su Shi's time, *bian* as a type of Buddhist picture storytelling performance and a genre of popular literature derived from such performances was already defunct. For Su Shi, only the residual derogatory associations of the term were still resonant in his mind. The few Song period textual references to *bian* that have survived stigmatize it as heretical, cultic, and licentious.[39] By Su Shi's time, in fact, *bian* were no longer remembered as having Buddhist associations. One thing is certain, however: Su Shi did not mean to flatter the author by referring to his work as a *bian*.

It was already bad enough to designate another man's work as a *bian*. We must now devote a great deal of attention to the rest of the title in order to determine just how much worse it was to characterize a text as *laza* ("jumbled"). Upon first glance, one might be tempted to interpret *laza bian* as *la zabian*, i.e., "to pull miscellaneous transformations." There is actually a slight justification for such an initial explanatory impulse, since *la* was used as a verb in this sense during the nineteenth century to describe the action of picture storytellers as they changed scenes by pulling their pictures, which were affixed to strings, through their peepshow boxes. Hence, we encounter the following expressions: *la yangpian* ("Pulling Foreign Picture Cards"), *la dapian(r)* ("Pulling Big Picture Cards"), *la dapian* ("Pulling Big

---

[37] MAIR, *T'ang Transformation Texts*, chapter six.

[38] *Ibid.*

[39] *Ibid.*, pp. 162-166.

Picture Leaves"), *la dahua* ("Pulling Big Pictures"), etc.[40] (All of these genres, however, are much too late to be applicable to the Song dynasty. Furthermore, peepshows were introduced to China from the West. We shall discover other objections to the interpretation of *la* in *laza bian* as "pull" below.)

We might also be tempted to interpret *la* as a verb in the sense of "talk at random," hence *la zabian* might be understood as "tell miscellaneous transformations." Again, there is some justification for such an interpretation, since in colloquial and dialectical parlance *la* by itself or in combination with other terms (in particular in northern Mandarin topolects) can mean "chat / talk leisurely / aimlessly" as in the expressions *la xian, la xianpianr, laguar, lahua, latan,* and *lajiachang*.[41] This is actually a fairly early vernacular usage; we find it already in chapter 95 of the late Ming novel, *Xiyou ji (Journey to the West)*.[42] There are, however, persuasive syntactical, grammatical, and lexical reasons that militate against such a reading (viz., taking *la* in the expression *laza bian* as meaning "pull" or "chat").

Within the essay itself, the expression *laza bian* occurs immediately after the verb *zuo* ("compose, write"). This makes it highly unlikely that *la* could also function as a verb. By comparison with the parallel clause in the first sentence, *zuo "Daren fu"* ("composed the 'Rhapsody on the Great Man'"), we may draw the following conclusions concerning the clause *zuo laza bian*: 1. like *fu, bian* signifies a literary genre; 2. just as *daren* modifies *fu*,[43] so does *laza* modify *bian*.

---

[40] MAIR, *Painting and Performance*, p. 14.

[41] MIN JIAJI, *et al., Changyongci cidian*, p. 294a, def. 2; p. 300b; p. 302a; DUAN KAILIAN, *Zhongguo minjian fangyan cidian*, p. 308a, first entry under *la*, def. 2; *Hanyu da zidian*, vol. 3, p. 1858b, definition 15; *Hanyu da cidian*, vol. 6, p. 497a, definition 15. The expression *la za tan* (in MOROHASHI, vol. 5, no. 4809b, no. 11945.29) is probably better analyzed as *laza tan* ("a conversation about all sorts of things") than as *la zatan* ("chatting about miscellaneous themes"). We will discuss the meanings of *laza* in depth below.

[42] WU SHIXUN and WANG DONGMING, *Xiaoshuo yuci da cidian*, p. 577a.

[43] *Daren* modifies *fu* in the sense that it tells us which *fu* or what kind of *fu* out of the infinity of all possible *fu*, i.e., the *fu* about the Great Man.

The question, then, is this: what exactly does *laza* connote? As we shall see, these two somewhat silly syllables convey a world of meaning. To discover that world, we must embark on an excursion in linguistics.

We could take the easy way out and rely on the authority of Morohashi and the *Hanyu da cidian* which, amazingly, both have entries for *laza bian*. The definition Morohashi gives is *yose atsume no bunshō* ("a heterogeneous composition").[44] The entry goes on to explain that such a composition was written in imitation of Sima Xiangru's rhapsody and that it was the work of someone who was "affecting" airs. The definition given by the *Hanyu da cidian* is "satirizes a work of literature that attempts to imitate the ancients but is actually an accumulation of mishmash."[45] All of this is fair enough so far as it goes, and it probably comes close to conveying Su Shi's intent in employing the expression *laza bian*. Nonetheless, if we wish to discover the full satirical nuances of *laza*, we must dig deeper. We shall start with the present and work backwards in time.

The word *lāzá* is still used today in many modern Sinitic languages, including various types of Mandarin. Dictionaries generally define it as meaning "rambling, jumbled, ill-organized; untidy, confused; all in a heap."[46] What is the derivation of this word? Strange as it may seem, there does not now exist – nor has there ever existed – any etymological dictionary of Sinitic words.[47] Instead, what usually passes for etymology is graphemic analysis in the manner of the *Shuowen jiezi [Explanations of Simple and Compound Graphs]* (100 CE). Never mind that many of the *Shuowen*'s explanations of the shapes of various graphs are incorrect because Xu Shen (c. 58-c. 147), its

---

[44] *Op. cit.*, vol. 5, 4809b, no. 11945.30.

[45] *Op. cit.*, vol. 6, p. 501b.

[46] It is listed, for example, on p. 362a of the outstanding new *ABC Dictionary* edited by JOHN DEFRANCIS. *Zhongwen da cidian*, vol. 4, p. 5687b, no. 12187.137 identifies *laza* as a Wu topolect usage (*Wu yan*) which signifies something that is neither neat nor clean.

[47] I have organized a large, international research project which has as its aim the compilation of such a dictionary within a decade.

author, was ignorant of oracle bones and bronze inscriptions. Nearly two thousand years later, there still is nothing much better to rely upon. Thus, following a Xu Shenian train of thought (as elaborated by his successors),[48] we might begin by saying that *laza* literally means "pull-miscellaneous." (We will ignore how **that** supposedly comes to signify "jumbled.") After that brilliant beginning, we move on to the level of the single graph where – according to the *Shuowen* school – *la* means "breaking something through force (*li*ᵃ) applied by the hand" (hence the hand "radical") and, furthermore, because the phonophore *li*ᵇ symbolizes the form of a man standing on the ground facing forward, it implies standing firmly, a condition which is necessary before one **pulls** something, so that is why *li*ᵇ is used as the phonetic. (!)

If the analysis of the shape, sound, and meaning of *la* is hard to follow, the Xu Shenian style of explanation for *za* is far more abstruse, so I shall have to break up the examination of this graph into several phases and devote a number of whole paragraphs to it. According to the *Shuowen*, *za* has the clothing "radical" *yi* which, combined with the *ji*ᵃ phonophore, affords the graph the basic meaning of "five colors combined together," the idea being that one supposedly uses the five colors when making clothing. Furthermore, according to Xu Shen, *ji*ᵃ originally depicted a flock of birds nesting

---

[48] See, for instance, GAO SHUFAN, *Xing yin yi*, p. 559b, 1984a, 1991b; CHANG HSÜAN, *Xing yi shi*, pp. 332, 821, 822; and WIEGER, *Chinese Characters*, p. 277 no. 119G, which are the chief sources of the Xu Shenian explanations given here, although I have consulted a dozen other premodern and modern grammatological works in Chinese and in Japanese, none of which could shed any rational light on the evolution of the graphs used to write the syllables pronounced *la* and *za* in MSM and roughly pronounced *lap* and *zap* in EMS, meaning respectively "pull" and "miscellaneous." It must be pointed out that the use of the word "etymology" in the English titles of Chang's and Wieger's books, although a very common practice in Sinology, constitutes an egregious error. **Characters do not have etymologies, only words do.** Characters are graphs used to write words. Like letters of the alphabet, Chinese characters are symbols used to record words; they are not the words themselves, hence they cannot have "etymologies." However, like letters of the alphabet, Chinese characters have evolved through time, thus their **derivation** may be studied.

in a tree, hence "gather together, collect" (presumably the source of the "five colors"[?]). The modern form *ji*[a] shows only one bird sitting in a tree, but Xu Shen had the notion that it was an abbreviation of a graph showing several birds perched in a tree, so he concocted a form with three birds *ji*[b] (a couple of birds in a tree would not do because that was already reserved for *shuang* meaning "a pair"). This, unfortunately, is totally fallacious, since the oracle bone and bronze forms of *ji*[a] (both much earlier than Xu Shen's small seal script) clearly depict only a single bird perched at the very top of a tree. It is not obvious why the ancients chose this graph (a bird in a tree) to represent the morphosyllable now pronounced *ji* in Modern Standard Mandarin (hereafter MSM) and meaning "gather, collect." Perhaps there is some yet undiscovered phonological reason, or perhaps one bird (birds ostensibly being gregarious) was meant to stand for a flock of birds. Carving more than one bird in the hard turtle plastrons or bovine scapulas used for divinations would have been avoided if one bird could do the trick.

Already long before Xu Shen compiled his *Shuowen jiezi*, the tree in the graph for *za* had slipped out from under the bird and was mostly to be found under the clothing "radical" (as in the standard form *za*[a]) or sometimes it was found at the bottom between the clothing "radical" and the bird (as variants of the graph on recently unearthed Western Han manuscripts attest).[49] This was remedied by devising another form of the graph which placed the clothing "radical" on the left and anchored the tree securely under the bird (*za*[b]), bringing it squarely in line with Xu Shen's speculations about the derivation of the graph. As for the enigmatic component at the top of the left side of the usual form of *za*[a] (see item *A* in the list of CHINESE CHARACTERS), that is a standard script stylization of the small seal form of the cloth "radical" (see item *B* in the list of CHINESE CHARACTERS).

Unfortunately, although Xu Shen classified *za*[a] under the cloth "radical," it would appear that even his own disciples were reluctant

---

[49] *Hanyu da cidian*, vol. 6, p. 4106a.

to accept the validity of that decision. Therefore, *za*[a] came to be classified under *zhui* ("short-tailed-bird"), where it is now found in virtually all dictionaries arranged by "radicals." Unfortunately, this totally destroys Xu Shen's neat theories about the graph (clothing "radical" *yi* + *ji*[a] phonophore). The top part of the phonophore has now been detached to become the "radical"; this leaves the graph in the disastrous state of being without a functioning phonophore. While there may be some remote logic in making "short-tailed bird" the classifier for a graph meaning "miscellaneous," neither MSM *yi* (Early Middle Sinitic [hereafter EMS] *ʔjei*, "clothing") nor MSM *mu* (EMS *muk*, "tree") can possibly serve as the phonophore for MSM *za*[a] (EMS *dzap*). In any event, if MSM *ji*[a] (EMS *dzəp* or *dzip*) is supposed to have any phonetic purpose in the graph and if it is to express any secondary semantic function ("gather, collect"), then it should be impermissible to sunder it as *zhui* and *mu*.

Still worse for those who hold sacrosanct the memory of Xu Shen, already by the Han period in which he himself lived, stele inscriptions reveal that people were taking shortcuts with the cloth radical on the top left of *za*[a] and were writing something that looked like a cross between a *liu* ("six") and a *jiu* ("nine").[50] Or maybe that is the way they had always written the graph for *za*[a] (which probably had not existed very long in any form by that time since we do not find it among the Shang oracle shell and bone inscriptions [OSBIs] or the Zhou bronze inscriptions) and it was only Xu Shen, due to a preconceived notion that the graph had something to do with variegated cloth, who tried to stabilize it with the proper small seal script form of the cloth radical (see item *B* in the list of CHINESE CHARACTERS, and item *A* for the same form in standard script).

Be that as it may, by the time of Wang Xianzhi[51] (344-388), we find the renowned calligrapher himself unmistakably writing *za*[c] in the standard script (*kaishu*)! In the interest of speed, efficiency,

---

[50] For examples of the forms referred to in this paragraph, see LI LEYI, *Jianhua ziyuan*, p. 293.

[51] GILES, *Biographical Dictionary*, 2176.

and ease, za^c was widely used as a variant[52] of za^a and za^b ("mixed / miscellaneous") through the centuries right up to 1931 when the bird was ejected from the right side, leaving only za^d, the current official form of the graph in the People's Republic of China today. With both the cloth and the bird missing, all that was left were a tree and a nine sitting atop it. A Xu Shenian devotee might claim that za^d means "miscellaneous, variegated," etc. because of the nine birds ("nine" stands for "nine birds") perched atop it. This is an even better solution than having three birds (reduced to one bird) representing numerous birds – two strokes instead of twenty-four strokes (three birds) or eight strokes (one bird).

With the cloth "radical" having vanished, it is no wonder that people desperately went searching for a replacement. Now, when one wishes to look up the character for za, za^a and za^c are to be found under the bird "radical" $zhui$, while za^d (together with another simpified variant za^e) are to be found under the tree "radical" mu. With such an unprincipled approach to the selection of radicals for graphs (and it is pervasive throughout the system), one can put little faith in the semantic explanations of the Xu Shenian School. They afford few reliable clues for the etymological investigation of words. How can one and the same character (za^a) be meaningfully classified under three completely different "radicals" ("cloth," "short-tailed-bird," "tree")? The fact that the *Shuowen* dictionary had 540 "radicals" while the Kangxi dictionary (1710-1716) reduced that number to 214 (less than half!) shows that the division of the semantic fields is arbitrary. The situation has deteriorated markedly within the last thirty to forty years, until now dictionaries are appearing with 181, 186, 189, 190, 242, and other sets of "radicals." It is obvious that the system of "radicals" is utterly bankrupt as a method for determining the origins of sinographs, much less the origins of Sinitic words. At best, the "radical" system is a slow, clumsy, and capricious means

---

[52] In Japan today, za^c is the officially correct form of the sinograph used to write Sinitic and Japanese vocabulary items pronounced $zō$, $zatsu$, ma, and $maji$ ("miscellaneous, mixed," etc.).

for looking up graphs. If agreement on the adoption of one set of radicals (for example, the 214 Kangxi radicals) could be assured, it would – as it was for more than two centuries – be worthwhile for students to memorize that set. With the annoying proliferation of sets of differing magnitude during the last three to four decades, however, even that minimal function is quickly dissipating because often only one or a few reference tools subscribe to a given set. Since agreement is lacking among the different sets of radicals, most dictionary users are simply refusing to learn any of them and have begun to resort more and more to the alphabetical indices at the back (if the editors were sensible enough to include one).

Of course, all of these perturbations over the derivation and construction of the individual graphs *la* and *za*[a] are of no value whatsoever in apprehending the etymology of the bisyllabic word *laza*; in terms of trying to come to grips with the etymology of the word *laza* they are nonsensical. I shall refrain from further commenting on whether or not these Xu Shenian explanations of *la* and *za*[a] do justice to the monosyllabic lexemes *la* ("pull") and *za*[a] ("miscellaneous"). We must adopt an entirely different strategy for analyzing bisyllabic Sinitic words.

Some may argue that the meaning of the bisyllabic word *laza* ("jumbled, rambling, ill-organized," etc.) can be extracted from the combined meanings of the two graphs used to write it, thus "pull" + "miscellaneous" = "jumbled." Such an explanation is, to say the least, forced, if not entirely ludicrous. I shall now proceed to demonstrate that it is patently false because *laza* (as do all words) possesses an underlying etymon. As such, it is possible to write that etymon (basically a fixed pattern of phonemes plus a semantic core) and its cognates with various combinations of homonphonous and near-homophonous graphs. The question we are faced with, then, is this: if the derivation of the word *laza* cannot be adequately explained by the combined connotations of *la* and *za*[a], what is its true etymology?

In his epoch-making dictionary[53] of premodern words entitled

---

[53] In my estimation, conceptually speaking, the three most important dictionaries

Citong, Zhu Qifeng lists *laza*[a] together in a group with the following words: 1. *lasa*, 2. *laza*[b], 3. *laza*[c], 4. *esa*, and 5. *laji* (for the six terms listed together, see item *C* in the list of CHINESE CHARACTERS). No. 5, *laji* or *lese* ("garbage"), is still used daily throughout much of China; I shall devote special treatment to it below. No. 1, *lasa* (usually glossed as "debauched, disorderly, jumbled," etc.) is still known to many and, among Wu speakers south of the lower Yangtze Valley, it is equivalent to no. 5. The other three terms are no longer current and are rarely encountered even in early sources. All five terms are vernacular words and thus extremely valuable for the data they provide concerning the living, spoken Sinitic languages as opposed to the moribund book language (Literary Sinitic).

It must be observed that these five terms are not recent inventions but share a long past. No. 1 appears first in the "Wuxing zhi [Treatise on the Five Elemental Phases]", scroll B of the *Jin shu [History of the Jin]* (646; the quotation is in reference to c. 395). No. 2 is from the *Guangyun [Expanded Rhymes]* (1008). No. 3 is from the *Nü lunyu [Women's Analects]* (820). No. 4 is from the Zen text *Wudeng huiyuan [Combined Original Texts of the Five Lamps]* (1252). No. 5 is from the record of daily life in Hangzhou during the Southern Song entitled *Mengliang lu [Record of the Millet Dream]* by Wu Zimu (fl. 1300).

Laza itself occurs quite early, being found in a famous Music Bureau ballad entitled "You suo si [There's Someone Whom I'm Think-

---

in the history of Sinitic lexicography to date are *Citong*, Axel Schuessler's *Dictionary of Early Zhou Chinese*, and the *ABC Chinese-English Dictionary* edited by John DeFrancis because all three of these dictionaries possess a clear concept of word in distinction to graph. In addition, *Citong* comprises the best collection of well-organized data for determining the etyma of Old Sinitic and Middle Sinitic, Schuessler's dictionary has a clear notion of morphology (all the more remarkable for such an early period as the Western Zhou), and the *ABC Dictionary* not only properly makes explicit the primacy of sound over script but also establishes a workable set of orthographical word boundaries that are essential for linguistic analysis and electronic information processing and transmission. The next major revolution in Sinitic lexicography will come during the decade 2000-2010 with the completion of the first genuine etymological dictionary.

ing of]." This poem was included in *Yuefu shi ji [Collection of Music Bureau Poems]*, compiled during the twelfth century by Guo Mao-qian. The *Yuefu shi ji* contains poems dating to the Tang period and earlier. Guo places "There's Someone Whom I'm Thinking of" in the Han period. The poem is about a girl who is thinking of her lover far away to the south. She wants to send him a gift, a tor-toiseshell hatpin with a pair of pearls and a bejewelled cord. Then she hears that he has another lover so she declares that she will "smash and burn it, leaving it in a jumble" (*laza cui shao zhi*).[54]

Of the sextuplet (Nos. 1-5 plus *laza*[a]) under discussion, though dating back as much as two thousand years or more, Nos. 1, 5, and *laza*[a] are still very much alive in common parlance. Although Nos. 2, 3, and 4 are seldom encountered today and were rare already in pre-modern Literary Sinitic texts, it is evident from the contexts in which these expressions occur that they all mean essentially "rubbish, garbage, filthy refuse." For example, No. 3 (from *Women's Analects*) is found in the following couplet:

> Sweep away the ashes and dust[55] with water,
> Gather up and remove the refuse.
> *sasao huichen*
> *cuochu laza*

Not only do all the members of this sextuplet share the same basic meaning, what is equally intriguing is that – despite the many differ-

---

[54] Even such excellent translators as Burton Watson and Anne Birrell are seduced by the dictionary gloss for *la* ("to break") to render this line as "I will break it, smash and burn it" (BIRRELL, *Ballads*, p. 147) and "So I break, I smash, I burn them" (WATSON, *Poetry*, p. 80). Such a reading fails to account for the fact that the syllable *la* is part of a bisyllabic word and that the surface signification of the character used to write it, *la* ("break; pull"), is irrelevant. Stephen Owen's translation, "I broke it and burned it in a pile" (*Anthology*, p. 228) shows that he recognizes *laza* as a single word meaning "in a jumbled heap." Owen, incidentally, dates this poem to the Western Han (*op. cit.*, p. 227).

[55] *Huichen* has long meant just "dust" but, as used in this early ninth-century quotation, the discrete significance of the *hui* syllable was probably still effectual.

3

ent graphs (some of very low frequency) with which they are written – all the members bear a close phonological resemblance to each other. This is apparent even in their MSM pronunciations (*lasa, laza*[b], *laza*[c], *esa, lese/laji, laza*[a]). It is particularly striking that, in EMS, all twelve sounds end with the entering tone *-p*. The EMS vowels are mostly some variant of *-a-* and the initials, too, coincide so that the first syllable begins with an alveolar lateral *l-* and the second with a fricative (generally dental or alveolar). Even the one exception, the initial EMS velar of the first syllable of No. 4, is phonologically explainable when we consider that a variant graph for the first syllable is *ke*.[56] MSM *ke* belongs to a class of words that had an initial velar-lateral (*kl-, gl-*) cluster in Ancient Sinitic.[57] Indeed, while most of the words in this class yielded some sort of initial velar in EMS and MSM, a few words actually developed an initial *l-* in EMS and MSM (e.g., *luo*[a] and *luo*[b] [both meaning "bare, naked"]) while others (e.g., *luo*[c]) occurred only in binoms that preserved the velar-lateral sequence as the initials of separate syllables, hence *guoluo*[a] ("wasp") and *guoluo*[b] ("gourd"),[58] cf. Guoluo (a tall, fair-skinned, big-nosed people living in Sichuan, Guizhou, Yunnan, and northern Vietnam).[59] *E*, on the other hand, belongs to a class of words that had an initial velar.[60] I suspect, therefore, that when No. 4 (*esa*) first came to be written down, like all of the other members of the set, it had an initial *l-* and that a graph from the *guo* group which retained an *l-* (perhaps originally *ke* itself) was used to write it. Later, phonological interference from the bulk of the group caused the word to be read with an initial *k-* and this could subsequently have led to the use

---

[56] IRIYA and KOGA, *Zengo jiten*, p. 138a. See also YUAN BIN, *Chanzong cidian*, pp. 591b-592a and *Zengaku daijiten*, vol. 1, p. 317d which provide several quotations proving that No. 4 literally means "rubbish, garbage."

[57] KARLGREN, *Grammata Serica Recensa*, no. 351; SHEN JIANSHI, *Guangyun shengxi*, pp. 101a-102b; MAIR, *Bottle-Gourd Myths*, p. 190.

[58] KARLGREN, *Grammata Serica Recensa*, no. 351c, g.

[59] *Gwoyeu tsyrdean*, s.v.

[60] KARLGREN, *Grammata Serica Recensa*, no. 642; SHEN JIANSHI, *Guangyun shengxi*, pp. 309b-310b.

of *ke* to write the syllable. Or maybe No. 4 arose in a topolect area where the initial velar-lateral cluster survived well into the medieval period and corresponded in some cases to a simple lateral initial in most other topolects. Regardless of the ultimate explanation, it seems certain that No. 4, with its apparent velar initial, is actually a phonological variant of the same etymon as that which constitutes the root of the other five members of the set.

What is more, a variant of the second graph used to write No. 4 is *ji*.[61] This is manifestly the identical graph as that used to write the second syllable of No. 5 which still today is a very common word for "garbage, refuse, rubbish."[62] By now it is becoming plainer and plainer that all six of the items under discussion are merely intimately linked cognates of a single etymon meaning primarily "garbage, rubbish, refuse," secondarily "untidy, unkempt, jumbled," and – by further extension – "debauched, filthy."

If there is a single etymon for all six of the members of the set, although this is not a customary exercise in Sinology, perhaps there is a way to recover it. Let us look a bit more closely at No. 5 which seems to convey most directly the basic meaning of the common etymon, viz., "garbage, refuse, rubbish." Both of the characters used to write the term have the earth "radical" *tu* which makes us think of "dirt(y)." (The hand "radicals" in most of the other graphs belonging to the set were most likely chosen because "garbage / refuse / debris / rubbish" is what you **sweep** or **gather up** and **carry out**, i.e., actions done with the hands.) In mainland Putonghua (Common Mandarin), the pronunciation of No. 5 is *lājī*, but this is an anomalous and relatively recent phonological change in the northeast, perhaps under the influence of "Altaic" languages. The MSM pronunciation of No. 5 is *lèsè*.[63] Whatever the reason for its unusual phonological transforma-

---

[61] IRIYA and KOGA, *Zengo jiten, loc. cit.*

[62] The first occurrence of *laji/lese* cited in *Hanyu da cidian* (vol. 2, p. 1087b) dates to c. 1300. There is no doubt, however, that the word *lapsap*, which the characters now pronounced *laji/lese* transcribe, already existed before that time and was represented in writing by other members of the sextuplet under discussion.

[63] See *Gwoyeu tsyrdean*, p. 1384b.

tion, we need not pay much attention to *laji* in our efforts to recapture the lost etymon of the sextuplet. Central Chinese topolects have largely lost the entering tone finals of the two syllables, although some (e.g., Suzhou) do optionally retain final glottal stops as relics of the entering tones that closed each of the constitutent syllables. Most southern topolects (e.g., Meixian, Canton, Yangjiang), which preserve older phonetic features of Sinitic, still pronounce No. 5 as something like *lapsap*.[64]

We are now in a position to begin to formulate a reasonable hypothesis concerning the common etymon for the sextuplet of old vernacular words meaning "garbage, rubbish, refuse." I believe that the common root for all six members of the set was something like *slap*. (Pronounced with a back *a*, it would sound almost exactly like the English word "slop.") How did I arrive at this etymon?[65]

My hypothesis, based on a preliminary examination of the words in *Citong* and half a dozen other collections of bisyllabic terms in pre-twentieth-century texts, is that a significantly large percentage of the pre-modern multisyllabic vernacular vocabulary of Sinitic (e.g., *daolu* ["road"], *cuiwei* ["lofty, towering"], *weiyi* ["winding, meandering"], *qilin* ["unicorn"], *kulong* ["cave"], *Kunlun* [name of a mountain range in the west], and so forth)[66] is the result of the bisyllabicization of words that were originally one syllable in length. Regardless of whether such single syllable terms which were divided into bisyllabic words were indigenous (i.e., originally Sinitic) or were borrowed into Sinitic from other language groups, they were all subjected to a particular bisyllabicizing process that we may refer to as dimidiation.[67]

---

[64] *Hanyu fangyan cihui* (1st ed.), p. 242a; *Hanyu fangyan cihui* (2nd ed.), p. 35b.

[65] South Coblin, in a personal communication of July 27, 1997, suggested to me the possibility that *laza* ("jumbled") and *lesi / laji* ("garbage") might actually be variants of the same word.

[66] Most of these words can be written with different combinations of homophonous or nearly homophonous graphs. Again, this underscores the primacy of the phonemically determined etymon over the graphs used to write its derivatives.

[67] For Boodberg's concept of dimidiation (splitting into halves), see the two articles by him listed in the bibliography (esp. p. 409 of "Iconography" and pp. 401-402

Some would argue that this process was the result of phonetic nivellization (levelling or drying up), that is, a reduction in the phonemic inventory of Sinitic languages which made it impossible to pronounce consonant clusters and other complicated phonetic configurations. Supposedly, this would have forced speakers to break up those clusters and configurations into separate syllables.

I am not in a position to determine the veracity of such an assertion, although I would suspect that phonetic nivellization was indeed operative in the bisyllabicization of Sinitic vocabulary. (Whether phonetic nivellization was a cause or an effect of bisyllabicization is another question altogether.) I would further surmise that the syllabic nature of the script also contributed greatly to phonetic fission because, unlike an alphabetic script, it did not allow for the joining of phonemes through spelling. Instead, the script required the manipulation of whole syllables at a time. For example, faced with the monosyllabic name "Türk" (perhaps from the Old Turkic word meaning "strong"), the person who is wielding the sinographic script as a device for transcription would write *tujue*. In so doing, not only has the originally monosyllabic word become bisyllabic, it has also lost its medial -*r*-. Examples of this sort could be multiplied endlessly from all periods of the history of Sinitic languages after the adoption of the Sinographic writing system. They are to be found both externally, as with "Türk" > *tujue*, and internally, as with the dimidiate words mentioned two paragraphs above.

There is, of course, an enormous component of the multisyllabic vocabulary of Sinitic that was created by other processes such as the following: 1. affixation (to account for morphological elements that were formerly probably indicated chiefly by segments rather than whole syllables), 2. reduplicative, alliterative, onomatopoeic, and rhyming binoms (frequently encountered already in the *Shi jing* [*Poetry Classic*]), 3. the direct translation or transcription of Buddhist

---

of "Proleptical Remarks"). Since the Song period, traditional Chinese language studies has had the concept of *lianmianzi* ("connected graphs") which recognized the existence of alliterative, rhymed, and other types of bisyllabic terms.

terms and terms from other foreign sources (*saṃyak-sambodhi, deng-zhengjue* ["complete and correct awareness"], *pusa* ["Bodhisattva"]), 4. the joining of synonymous or nearly synonymous graphs (*quqian* ["drive off"]), 5. the joining of semantically opposite or nearly opposite graphs (*benmo* ["the fundamental and the incidental"]), 6. the fusion or binding into a single lexical unit of a modifier and the noun or other element that it modifies (*baojian* ["precious sword"], *keyang* ["deeply admire"]), of a verb and its complement or object (*xunwei* ["abdicate"], *jiuzheng* ["redress"]), or of components standing in other grammatical relationships to each other, including particles used as conjunctions (*tangruo* ["if, supposing, in case"]), 7. the invention of entirely new words through the combination of two or more graphs of appropriate meaning (*dimsun / dianxin* ["appetizer, dessert, snack"], *jixin* ["chicken heart, heart-shaped ornament"], examples are endlessly diverse), 8. the insertion of fillers and nonsense syllables (*shabulengdengde* ["daffy, doltish, foolish"], *pangdundun* ["chubby, plump"], *nahar* ["where"], 9. the creation of calques and neologisms to account for new devices, terms, and concepts (*shengwuxue* ["biology"], *jingji* ["economics"], *huochetou* ["engine"], *daziji* ["typewriter"]) (many of the items in this category were borrowed from Sino-Japanese vocabulary established in Japan to cope with the flood of European technology and terminology that inundated East Asia during the latter part of the 19th century and the early part of the 20th century), 10. and so forth. I would argue that these processes were: 1. by and large later than (and perhaps even stimulated by) the type of dimidiation discussed above, 2. largely (but not entirely) semantically or grammatically / syntactically, rather than phonetically, driven as was dimidiation, and 3. except for no. 2 and no. 8, initially more characteristic of book (i.e., purely written) vocabulary in contrast to dimidiate words which were all along more characteristic of spoken vocabulary and, indeed, were often not writable in sinographs. That is to say, I believe that the phonological process of splitting or fission of monosyllables which we may, for convenience, refer to as dimidiation, was more primitive than that of the other types of script-oriented multisyllabicization listed above. It is "primitive" in the sense of being earlier and more closely tied to

speech (which I view as a primary linguistic phenomenon) in contrast to writing (which I view as a secondary linguistic phenomenon).

Whatever the cause of the splitting process (dimidiation), it appears to have begun already during the Zhou period (1045-771 BCE), to have picked up momentum during the Spring and Autumn period (770-476 BCE), to have surged during the Warring States period (475-221 BCE), to have reached a peak during the Qin and Han periods (221 BCE-220 CE), and to have been largely (but by no means entirely) succeeded by the more script-driven types of multisyllabicization from the Wei and Jin periods (220-420 CE) onward. This scheme is intended only to give a very rough outline of the evolution of Sinitic multisyllabic vocabulary; it is not meant to be exhaustive or definitive. If there is any truth to the summary of the development of the multisyllabic lexicon presented here, what it tells us is that: 1. the most archaic Sinitic vocabulary was chiefly (perhaps even exclusively) monosyllabic, 2. the earliest stage of multisyllabicization of the lexicon was dimidiation triggered by simultaneous phonetic nivellization and the impact of the monosyllabic script, 3. the overwhelmingly multisyllabic nature of the Sinitic lexicon cannot be adequately accounted for solely by such simplistic notions as compounding.

Having established that dimidiate bisyllabicization played an early and crucial role in the multisyllabicization of the Sinitic lexicon, I shall now adumbrate an experimental methodology for the restoration of monosyllabic etyma. We shall begin with the *laza* sextuplet which forms the centerpiece of this section of the paper. I have proposed that *lapsap* (or *lapdzap*) was originally *slap* (or *dzlap*). Initials like *ls-* and *ldz-* are rare throughout the world but virtually impossible in Sinitic.[68] Conversely, it is not unusual to find recon-

---

[68] On August 8, 1997, I posted a question about the existence of *ls-* and *lz-* (i.e., a liquid plus a fricative in that order) configurations on the Linguist List electronic bulletin board. I stated that such articulations seemed to me to be phonologically improbable and that they might naturally metathesize to *zl-*, *sl-*, etc., or that, if they did occur, they would be highly marked. Among those who kindly replied to my query were Victor Peppard via Jacob Caflisch (University of South Florida), Mark

structions of Old Sinitic (e.g., Baxter and Schuessler) that posit initial *sl-* consonant clusters. Hence *lapsap* and *lapdzap* most likely were due to metathesis from *slap* and *dzlap*, not directly from phonologically improbable **lsap* and **ldzap*. To make it easier to identi-

---

Liberman (University of Pennsylvania), Peter Chew (Oxford University), David Robertson (tincan), John E. Koontz (Boulder), Subhadra Ramachandran (cyantic), Robert Beard (Bucknell University), Sondra Ahlen (cmu), James Giangola (General Magic), Christopher Miller (University of Quebec), Colin Whiteley (Barcelona), Ronald Cosper (Saint Mary's University, Halifax), Alain Theriault (University of Montreal), Jakob Dempsey (Yuan-ze University, Taiwan), Kimmo Huovila (kielikone, Finland), Michael Betsch (Tübingen University), Sandra Paoli (University of York, England), Mark Donohue (United Kingdom), David Gohre (Indiana), Geoffrey Sampson (University of Sussex), James Kirchner (no address or affiliation), Olga Shaumyan (University of Sussex), Steve Seegmiller (Montclair State University), Manaster (probably Alexis Manaster Ramer, Michigan), Paul Boersma (Instituut voor Fonetische Wetenschappen, Amsterdam), Wolfgang Behr (Frankfurt University), Keith Goeringer (University of California at Berkeley), Heli Harrikari (University of Helsinki), Charles Gribble (OSU), Elena Andonova (Bulgaria[?]), Lance Eccles (Macquarie University, Australia), and F. Gladney (cso.uiuc). Several graduate students at the University of California (Los Angeles) and elsewhere requested that their names not be listed in my response because they did not want to get in trouble with their advisers for spending too much time on the Internet. I hope that I have not inadvertently forgotten any others. My profound gratitude is due to each and every one who responded.

The gist of the information which the above-named individuals provided to me is that there certainly do exist *ls*, *lz*, and similar configurations, even in English (e.g., "else," "holster," "also," "balsam," "pulse," "calcium," "dulcimer," "bells," "pulls," "files," and "celsius"), but note that these are all internal or final. Other languages with internal *-ls-*, *-lz-*, etc. (often separated in two adjoining syllables) cited in the responses include Coast Salish, Malayalam, Bulgarian, French, Spanish, Portuguese, Italian, and Finnish. It was reported that some Athapaskan languages may have such clusters in final position. As indicated by the dashes, however, I was thinking of syllable initial *ls-*, *lz-*, etc.; it would appear that such articulations are quite rare throughout the world.

Levantine and Western dialects of Arabic were mentioned among the replies I received. Lance Eccles cited Maltese *lsien* ("tongue") and *lsir* ("slave"), but said that he was fairly sure that a native speaker pronouncing these words in isolation puts an epenthetic *i* at the beginning. Also mentioned was the mysterious language Lvova, said to be from the Santa Cruz Islands, Solomons, and written about by Wurm in articles for numerous Pacific linguistics publications. The languages of the Caucasus were noted as being particularly rich in initial consonant clusters, but *ls-* and *lz-* were not cited specifically.

The overwhelming preponderance of the citations for such configurations were

fy, I shall also refer to the *slap / dzlap* etymon as the "slop(py)" etymon.

One may object that, even if we do accept the possibility of the derivation of a set of dimidiate words from a hypothetical *slap / dzlap*

---

from Slavic languages, in which some of my correspondents declared that virtually any combination of consonants is possible! (For example, there is a Russian word, *vzbzdnut'*, which you will not find in any dictionary, that means "to emit a silent but very smelly fart." And Czech, amazingly, even has whole words that are spelled without any vowels, although out of physiological necessity a kind of epenthetic schwa is used when they are pronounced. Geoffrey Sampson cites the Czech word *vlh* ["wolf"] which consists wholly of an *-l-* sound surrounded by fricatives on both sides [the *-h* in this word is actually a voiceless velar fricative, IPA [x]]!) As Victor Peppard put it, "One of the reasons Slavic has so many complex consonant clusters is that in about the ninth century Common Slavic lost a pair of semi-vowels, one back and one front, precipitating in a lot of places, to put it colloquially, a tremendous collision of consonants." Nonetheless, even in Slavic, *ls* or *lsh* and *lz* or *lzh* are usually found intervocalically, but are much less common (and **harder to pronounce**) in initial position (cf. *lzh-* ["false"], *lze* ["possible"], etc.). Often, as with Russian *l'stit* ("to flatter") and *l'viny* ("lion's"), an initial *l-* in such combinations tends to become palatalized. The rule, according to F. Gladney (also noted by John Koontz and Axel Schuessler for consonants other than *l* as well), is that loss of *-i-* led to soft *l'-* and loss of *-u-* led to hard *l-*. Both *l'-* and *l-* before C cause the same sort of syllabification challenge.

The difficulty of pronouncing syllable initial *ls-*, *lz-*, was commented upon by Sondra Ahlen as follows: "In that case I would not be surprised to see some phonological process occur since as I recall syllable initial sequences tend to involve increasing levels of sonority as you get closer to the nucleus, with the common exception of fricatives before stops as in *str-*. Metathesis is one of several phonological processes that might affect an underlying syllable initial (or potentially syllable initial) [liquid plus fricative] such as *lz-*, *ls-*. Other options might include vowel epenthesis, consonant deletion, syllabification of the liquid, etc."

Paul Boersma cited one instance of metathesis in Czech: *ml-ha* ("fog," two syllables, the /l/ being syllabic) from an older *mgla* which still exists in Polish.

A check of all the roots beginning with *l-* in the *Etimologicheskii Slovar' Slavyanskikh Yaz'ikov*, vols. 15-17, revealed that whenever the *l-* was not followed by a vowel (i.e., when it was followed by something other than a vowel), the letter to be found was either the hard or the soft sign. My interpretation of this pattern would be that it reflects a phonological process designed to ease the pronunciation of the following consonant (including *-s-* , *-z-*, and *-zh-*) after the *l-*.

English-Russian dictionaries list under the word "slop" Russian *luzha* ("puddle, pool") and the idiom *sest' v luzhu* ("to get into a mess; slip up"). The vowels are problematic and the lack of a *-p* at the end is troubling, but the sequence of consonants is very close to the unmetathesized EMS form of our *lapsap* (or *lapdzap*) set of words

etymon, would we not expect them to look like *saplap* (or *dzaplap*), rather than *lapsap* (or *lapdzap*)? If we are starting with a hypothetical √*slap*("slop[py]") etymon and positing dimidiation, why should we not end up with *saplap* (or *dzaplap*) instead of *lapsap* (or *lapdzap*)?

---

and the meaning partially coincides with that of some members of the set. The superficial similarities encourage us to look further. However, Russian etymological dictionaries (FASMER, vol. 2., pp. 529-530 and SHANSKII, *et al.*, p. 248) identify *luzha* as deriving from an old Slavic root *lug* ("morass, swamp, marsh; lake") with final palatalization (TRUBACHEV, *op. cit.*, vol. 16, pp. 169-170). Therefore, we can reject *luzha* as a possible source for the *lapsap* (or *lapdzap*) sextuplet.

Jakob Dempsey provided extremely valuable data from Tibetan which lends support for the possibility of metathesis being involved in the evolution of *lapsap*, etc. from a hypothetical √*slap*("slop[py]") etymon: "Old Tibetan 'moon' was *sla* which assimilated to *zla* in the classical period, but in the western dialects this underwent initial-cluster metathesis (seen in many examples of western Tibetan): *zla* > *lza*. That form remains in the extreme west (Balti), but in central Tibet we have: *nda* < *lda* which in turn seems to come from *lza*. It has been proposed that *lce* ('tongue') came from *sle* (via *lse*), but since there are still dialects in Tibetan which preserve *cle*, this is yet another example of that metathesis, with the *c*- in *cle* probably a palatization of earlier *tle* which in turn may be from *ple*, cf. Drung *p-lai* (Drung has many old loans from Tibetan). 'Tongue' in many other Tibeto-Burman languages is from *ble*." There are other possible historical explanations of voiceless initial *l*-.

Wolfgang Behr observed that "Qiangic [a Tibeto-Burman language found in Sichuan Province of China] allows *rp*-, *rk*-, *rt*-, *rb*-, *rg*-, *rts*-, *rdz*-, *rtsh*-, *rdzh*-, *rdzh*-, *rm*-, *rng*-, *rl*- [!!], *rw* (with distinctive syllabic and non-syllabic *r*-), but no *ls*- or *lz*- (neither *rs*- or *rz*-). Jiarong [another Tibeto-Burman language from the same area of southwest China], although equipped with one of the most curious initial cluster systems known (> 170 types), has such things as *ltsh*-, *ldz*-, *ldzh*-, *lj*-, but again, no *ls* or *lz*-." As for the anomalous distribution of preinitial resonants in Written Tibetan (e.g., <*rts*> but not *<*lts*>, etc.), this phenomenon has apparently never been explained in the literature. It is not known for sure whether these clusters were ever pronounced as they were written in the Old Tibetan and Pre-Tibetan periods (we may notice the great variation of written cluster representations in the Dunhuang documents), or if they were pronounced sesquisyllabically, or if the preinitials came into being as mere graphical conventions marking tone. Similar clusters, violating not only basic sonority hierarchy restrictions but even such notions as Hjelmslev's "resolvability principle" (i.e., every language L that allows C1C2C3- initials of a given shape in its phonotactics must allow for all adjacent subsets of the cluster, viz., C1C2-, C2C3), have been set up for Old Sinitic by "proto-form stuffers" (to use James A. Matisoff's term). Those who have done so, again quoting Matisoff, lack an adequate "Proto-Sprachgefühl."

Finally, Wolfgang Behr also offered some very interesting theoretical perspectives,

We may conjecture that the metathesis was invoked by a principle of euphony that preferred the sequence *lapsap* over *saplap, lateral–alveolar rather than alveolar–lateral in separate, sequential syllables each of which ended in a voiceless bilabial stop. Mechanically, it does seem easier to say *lapsap* than it does to say *saplap. Furthermore, so long as they were comfortably separated from each other by the syllabic vowel (and its accompanying entering tone stop), there was no danger of running afoul of a prohibition against *ls or *lz. Regardless of the reason for choosing *lapsap* over *saplap, we can be fairly sure that the *lapsap* / *lapdzap* ("refuse, rubbish, garbage") sextuplet derived from a hypothetical √*slap*("slop[py]") etymon, not from a phonologically unlikely *lsap or *lzap.

The next question to pose is this: where did this productive, prolific √*slap*("slop[py]") etymon come from? Naturally, the first place to look would be within Sinitic. Are there any monosyllabic Old Sinitic words pronounced roughly *slap* that mean "garbage, refuse, rubbish; filthy, untidy," etc. and that could have split up into two syllables by Han times? I do not know of any.[69] Nor can I find such a root in lexicons of Sino-Tibetan[70] or Tibeto-Burman,[71] which might have been expected to harbor cognates if √*slap* ("slop[py]") were part of Sinitic vocabulary from the beginning. It would seem, then, that

---

complete with an extensive bibliography, concerning the "sonority sequencing principle" (SSP) and its violations. A basic assumption of the SSP is that the least sonorant segments occur toward the margins of a syllable. Among the finer differentiations of the sonority scale are those proposed by Th. Vennemann in his *Preference Laws for Syllable Structure* (Berlin, Mouton, 1988). According to the sonority restrictions applying to the distribution of segments in a syllable on Vennemann's fine-grained scale, predictions may be made about statistical frequencies or markedness properties. By these standards, *ls*- and *lz*- would have to be classified as marked.

[69] A check of SCHUESSLER, *Early Zhou Chinese* and BAXTER, *Old Chinese Phonology* did not yield such a word. CHANG, *Indo-European Vocabulary*, p. 7, lists *xi* ("slip into"), which he reconstructs as *zleup*, as having a connection with √*sleub(h)-*. If this is really so, about which I am rather dubious because of the poor semantic and phonological fit, it would have been a separate – and probably much earlier – borrowing.

[70] I checked COBLIN, *Handlist* and other standard lists.

[71] See *Zang-Mianyu yuyin he cihui*, p. 843, no. 473 and HUANG BUFAN, ed., *Zang-Mian yuzu yuyan cihui*, p. 181, no. 542.

√*slap* ("slop[py]") may have entered Sinitic from some other language family.

Although "Altaic" languages appear too late in history to deserve serious consideration as the source of a word that was already in Sinitic from at least the time of the Western Han dynasty (206 BCE-23 CE), a survey of the standard sources[72] fails to turn up a suitable candidate for √*slap* ("slop[py]"). Nor can I find a good candidate for √*slap* ("slop[py]") in Austroasiatic, Austronesian, or any other likely source of an early borrowing into Sinitic. Consequently, I would conclude that it is not altogether impossible that √*slap* ("slop[py]") has an Indo-European (henceforth IE) source (IE √*sleub(h)* > Germanic √*sleup-*).[73]

Numerous old Indo-European words (e.g., those for "magus," "goose," "dog," "lion," *pipa* ("pear-shaped lute"), "coral," "honey," "wheel," "wheat," "track," etc.) have been found in Sinitic.[74] Hence, it would not be entirely out of the question for "slop" to have been borrowed by Sinitic languages. I wish to emphasize that, if √*slap* ("slop[py]") did have an IE source, I do not know specifically from which of the daughter languages it might have entered Sinitic. My only point in raising the possibility of an IE source for √*slap* ("slop[py]") is that it provides a conceivable origin for this etymon which corresponds to it quite closely both semantically and phonetically. One objection that may be raised against the connection of √*slap* ("slop[py]") and IE √*sleub(h)* ("slip") is that the relevant intermediary cognates (if, indeed, the IE term was actually borrowed into Sinitic) are to be found chiefly in Germanic.[75] Skeptics would want to

---

[72] E.g., CLAUSON, *Pre-Thirteenth-Century Turkish*, pp. 825, 828.

[73] *American Heritage Dictionary*, p. 2125b.

[74] See, for some examples and for additional references, MAIR, *Language and Script*.

[75] POKORNY, *Wörterbuch*, p. 963; LEHMANN, *Gothic*, S103, p. 315b: "PIE *(s)lewb(h)-* **glide, slip**: outside Gmc only in Lat *lūbricus* **slippery**...." Under LSB, Clauson (*Pre-Thirteenth-Century Turkish*, p. 764b) gives the Tocharian A word *lešp* ("mucous, phlegm") which worked its way into Turkic. *Lešp* most likely cannot account for the Sinitic √*slap* ("slop[py]") etymon; the sound and the meaning are both too far

know how, when, and where such a transmission could have oc-
curred.

My response to this objection would be that we still have not de-
termined the ethnic and linguistic identity of the thousands of Euro-
poid individuals dating to the Bronze Age and Iron Age whose re-
mains have recently been found on the western borders of China.[76]
While research is still going on, we should not prematurely dismiss
the possibility that some of them may have been speakers of Germa-
nic or Celtic (just as others were likely to have been Tocharian speak-
ers, Iranian speakers, and so forth). The textile evidence points in
that direction, as does the abundant physical anthropological evi-
dence and the limited genetic data that are available. We also need
to take into consideration the proposals of Tsung-tung Chang for ex-
tensive borrowing of Germanic words in Old Sinitic.[77] So far, his the-
ories have not received a serious, scholarly rebuttal but have only en-
countered sarcasm and sheer disbelief.

---

removed. Nor is it phonologically possible to derive Tocharian *leśp* from PIE
√*sleub(h)*- ("slip" → "slippery") for the following reasons: 1. The palatal sibilant
in Tocharian A never reflects a PIE *ś. 2. Tocharian A *e* can reflect PIE *ai* or *oi*
(or *ei*, if the preceding *l* were palatalized to *ly*), but it cannot reflect any *u*-diphthong.
3. There are insufficient grounds for explaining the metathesis. (Personal communi-
cation, Donald Ringe, August 17, 1997.)

[76] HADINGHAM, *Mummies*; MAIR, *Mummies*; MAIR, *Language and Script*, MAIR,
ed., *Mummified Remains*; MAIR, ed., *Bronze Age and Early Iron Age Peoples*; etc.

[77] CHANG, *Indo-European Vocabulary in Old Chinese*. To disprove Chang's the-
ories, it will require more than mere assertions of impossibility or simple statements
to the effect that his phonological reconstructions are inadequate. To date, all recon-
structions of Old Sinitic are grossly inadequate. To disallow any validity for Chang's
proposals concerning IE elements in Old Sinitic, what must be shown precisely and
explicitly are how his reconstructions are inconsistent (if, indeed, they are so) and how
they fail to support the linkages he asserts. Even if grave errors are discovered in a
substantial proportion of Chang's correspondences, this does not mean that all of
them are to be dismissed out of hand. Eventually, among those who disagree with
Chang, someone will have to take the initiative to carry out a responsible critique,
rather than attempt to denigrate his ideas with peevish, emotional outbursts. After
all, Chang has spent many years assembling and presenting his evidence. A few mo-
ments of indignation will not suffice to demolish the totality of his work. Petulant dia-
tribes convince no one and only bring disrepute to those who unleash them.

The √*slap* ("slop[py]") etymon that I have proposed almost certainly would have been linked to other cognates. For example, another colloquial word still widely used today is *lata* ("sloppy, dirty, untidy, unkempt, slovenly"). Although this is pronounced *lata* in MSM, it may be reconstructed roughly as *lapt^hap* in EMS. The same word is also written with the graphs *lata*[a] which also have the same MSM and EMS pronunciations as *lata*. *Lata* occurs with the meaning "not neat, untidy" in Yuan drama written both as *lata*[a] and *lata*[b].[78] By this time, the EMS pronunciation of the latter (*latt^hap*), while close to that of the former, is irrelevant because both pairs of graphs had lost their final entering tone stops and would have sounded almost exactly alike (*lāta* versus *làta*).

We may note that several of the definitions for English "slop(py)" coincide uncannily closely with the range of meanings for the extended √*slap* ("slop[py]") set of Sinitic words:

slop "to move in an idle, lazy, casual, or slovenly manner"
"the dirty water, liquid refuse, etc. of a household or the like"
"kitchen refuse; swill"

sloppy "careless; loose: *sloppy writing*"
"untidy; slovenly: *sloppy clothes; a sloppy eater*"
"(of clothes) loose-fitting; baggy: *a big, sloppy sweater*"[79]

cf. **slip, cowslip, oxslip**, all of which mean "dung"

It is essential to observe that, as in the lines from the *Women's Analects* quoted above and in most of the vernacular Zen texts where the √*slap* ("slop[py]") cognates occur, the reference is specifically to piles of garbage, rubbish, etc. that are to be swept or carried out.[80] In old Chinese houses, and still today even in the finest traditional Japanese houses, the swill and refuse from the kitchen is **washed down with water** and then **swept up** (sometimes into a hole sunk in the

---

[78] WANG and YE, *Yuci da dian*, p. 826b.

[79] *Random House Dictionary*, p. 1800ab.

[80] See references cited in note 56 for relevant quotations.

floor, from which it is then **gathered up** and **carried out**) to be deposited in the street, whence it may be collected.

Similarly, *laluo* (EMS *lapla*),"jumbled", which seems to occur only from the Qing period on,[81] is probably a late variant of *laza*, etc. in which the sibilant is lost altogether.

The expression *zata* (EMS *dzapdap*), which occurs already in the *Shi ji [The Grand Scribe's Records]* (c. 90 BCE) and the *Han shu [History of the Han]* (82 CE), and is taken up by such illustrious mainstream authors as Liu Xie (465-520), Han Yu (768-824), and Du Fu (712-770),[82] as well as occurring numerous times in the celebrated *Wen xuan [Literary Selections]* (c. 526-531), has a meaning close to that of *laza*, *lata*, etc. and is written with some of the same graphs used for the latter series, actually belongs to a separate set of words. We may extrapolate the basic etymon of this sextuplet meaning "crowded / jammed / close together" as √*stap* ("to stop (up), stuff").[83]

Another superficially related series is *sata*[a], *sata*[b], *sata*[c], *sata*[d], and *sata*[e] (EMS *saptap*), of which there are many instances in *Wen xuan*, although the expression continued in use through the Song, Ming, and Qing periods.[84] The *sata* quintuplet, again, uses many of the same phonophores and even some of the same characters as the previously discussed sets. The underlying etymon conveys flurried numerousness, which might tempt us into connecting it with the √*slap* ("slop[py], jumbled, variegated") series. However, as is obvious from all but two of the graphs chosen to write the word, it also is conceived of as having to do with horses rushing or stamping rapidly. Thus it should be treated as a separate series whose etymon I would reconstruct roughly as √*stap* ("stampede").[85] That the phonetically and semanti-

---

[81] *Hanyu da cidian*, vol. 6, p. 502a.

[82] *Hanyu da cidian*, vol. 11, p. 872a; *Citong*, p. 2766c.

[83] BARNHART, *Dictionary of Etymology*, pp. 1071b-1072a, 1080a. Both "to stop up" and "to stuff" come from Old High German *stopfon*. Cf. Old Saxon *stuppon* which basically means "to fill, stuff, cram, stop up."

[84] *Citong*, p. 2766c; *Hanyu da cidian*, vol. 12, pp. 800b-801a. Zhu Qifeng adds *sajia* to this series, but I reject it both on semantic and phonetic grounds.

[85] The Germanic origin of "stampede," related to "stamp" (from *stampen*

cally determined *sata* < *saptap* < √*stap* ("stampede") etymon is primary and the graphs used to write it are secondary is evident from the plethora of different characters involved.

Members of the extended √*slap* ("slop[py]") family are often defined by the expression "not *lisuo* ('neat, clean, orderly; agile, nimble, dexterous')."[86] A variant of this word is to be found in *lisou.* It is curious that the meaning of the northern topolecticism *lisuo* is almost exactly the opposite of *laza, lata,* etc. Even more interesting is the fact that *lisuo* (EMS *li*ʰ*sak*) appears to be another dimidiated term. We may reconstruct a possible √*slik* ("slick"). The word *lisuo* appears to have surfaced in writing only within the past century, but it may have existed in speech for a much longer period before that.[87]

In all of the above cases where I have given an English parallel for a reconstructed Sinitic etymon, viz., √*slap* ("slop[py]"), √*stap* ("stop [up], stuff"), √*slik* ("slick"), and √*stap* ("stampede"), the English word is intended as a mnemonic gloss. I do not necessarily maintain that in each of these cases (and in dozens of others that could be easily adduced) there is any direct connection between the Sinitic etymon and the English gloss. On the other hand, I would not rule out altogether the possibility of Germanic borrowings into Sinitic at various times in history and prehistory, just as I would not ignore the countless borrowings from English and other Germanic languages into Sinitic that have occurred during the last several centuries. Still less would I rule out the possibility of borrowings into Sinitic from other IE branches such as Indian, Iranian, Tocharian, etc. We should note, however, that the IE roots of the mnemonic glosses

---

["pound, beat, crush, mash"], c. 1200) is noted in BARNHART, *Dictionary of Etymology*, p. 1058b-1059a and in *American Heritage Dictionary*, p. 1752a. Cf. Old High German *stampfon* (Modern German *stampfen* ["to stamp with the foot, pound"]) < Proto-Germanic **stampojanan.*

[86] *Hanyu da cidian*, vol. 2, p. 637b; DUAN KAILIAN, *Fangyan cidian*, p. 337a; *Hanyu da cidian*, vol. 1, p. 1401ab.

[87] A close synonym of *lisuo* is *liluo* (EMS *li*ʰ*lak*). Its range of meaning and area of usage are almost identical. *Citong*, p. 2766c. It is possible that *liluo* is a secondary derivation from *lisuo.*

for the four etymons identified in this paper are located specifically within Germanic.

Having dissected *laza* and related terms, we are at last ready to ask: just how bad would it have been to call someone's composition a *laza bian*? Aside from the decidedly negative (to the mind of a Song literatus) connotations of *bian* itself as a semi-vernacular, popular literary genre, characterizing someone's work as *laza* would have amounted to calling it a heap of refuse.

While *laza* itself no longer explicitly and specifically signified a heap of refuse, its membership in the √*slap* ("slop[py]") series of words would still have been alive for a clever, phonologically sensitive writer in Su Shi's time. With the loss of entering tones by the Yuan period in the north, referring to something as *laza* may only have implied that it was jumbled or heterogeneous.[88] In this sense, the offal (pun intended) associations of *laza* were undoubtedly even further diluted (pun intended) for northerners since the Yuan period. Still today in the south, however, it would be hard to avoid the connection with swill, scourings, dregs, scrapings, and lees. When Song period texts mention *lapsap*, it is frequently in tandem with night soil and other fecal matter. Together, these foul substances were to be swept or washed out of the house where they would be collected by individuals who were paid to remove them. When Su Shi styled the would-be Sima Xiangru's composition as *lapdzap*, he knew exactly what he was doing and the effect it would have. The choice of words would

---

[88] In chapter 52 of the mid-eighteenth-century novel, *Rulin waishi (Unofficial History of the Literati)*, written by Wu Jingzi (1701-1754) of Anhui, we find the expression *lalazaza*, meaning "in great disorder." See TIEN, *Colloquial*, p. 211a and TIAN, *Yongyu*, p. 556a. In such a construction where the individual syllables of *laza* have each been reduplicated, it would have been difficult any longer to catch the resonances with √*slap* ("slop[py]"). A different sort of lengthening has occurred with *lata*, such that we now also have *lalilata*, where *li* is merely a filler syllable. The meaning remains the same as the shorter form ("unkempt, slovenly; baggy; lackadaisical"), except perhaps for a slight intensification. As with *lalazaza*, the extension of syllables (particularly when spoken in a northern topolect) further removes us from the nuances of the √*slap* ("slop[py]") etymon (if, in fact, *lata* properly belong among the series of words deriving from it.).

4

have been calculated to achieve maximum effect without exposing himself to charges of outright slander.

What would a *laza bian* have been like? Su Shi was most likely prompted to brand the object of his disgust as *laza* because it consisted of piled-up allusions and accumulated borrowings that had not been forged into an integral piece of literature. Sima Xiangru's rhapsodies were distinguished by their seemingly endless catalogs of flora, fauna, hunting activities, and other realia, but they were always molded into astonishingly effective works of art. What must have irked Su Shi so much about the "Laza bian (Jumbled Transformation)" was its indiscriminate heaping up of odds and ends. In short, it was a pile of garbage.

CONCLUSION

Both of the new sources introduced here present us with information about *bian* ("transformations") that was hitherto completely unknown. At the same time, they also reinforce our understanding of other aspects concerning *bian* that were only partially surmised from the totality of previously known sources.

Only by patiently mining all Tang and later texts (including especially Buddhist works), and by keeping our eyes open for relevant materials made available through archeology and archival research, will we slowly recover a fuller understanding of China's past, including that of its unlettered and partially lettered masses who were fascinated by the Sino-Indian Buddhist narrative tradition in which *bian* played a pivotal role.

Appendix

The Phonotactics of the Sinographic Script and the Reconstruction of Old Sinitic

In my estimation, Chinese historical linguistics is at a crossroads. As Zhu Qingzhi, my friend and colleague from Sichuan University has repeatedly expressed, the best and brightest of China's young linguists (especially those who have been exposed to Western linguistics) feel "lost." (He himself uses the English word to characterize their feeling.) They recognize the inadequacies of traditional Chinese linguistics but do not know which way to turn. Yet this sense of crisis not only afflicts historical linguists in China, it also applies to scholars in the West who have devoted their lives to the history of Sinitic languages. It was due to this intense frustration and dissatisfaction with conventional methods of analysis that Jerry Norman and South Coblin recently issued their important manifesto suggesting new approaches. I hasten to point out that I fully endorse the positions advocated by Norman and Coblin.

Quite by coincidence, on the very day (August 19, 1997) when I was scheduled to mail this paper to the editor in Venice, my distinguished senior colleague, Tsu-Lin Mei, sent me a letter discussing the problem of cognates in Old Sinitic. Since much of the second half of this paper is devoted to that very issue, I shall summarize some of the points Professor Mei made in his letter.[89]

First of all, there is a large body of secondary literature in Chinese on the subject of cognates, most of which is not very reliable and hence must be used with extreme caution. Only when words are properly transcribed into Old Sinitic can we determine which ones are cognates with each other. Old Sinitic phonology is a new discipline; the most important advances were

---

[89] It is incumbent upon me to state that I am responsible for the exposition of Professor Mei's views as summarized here. The wording of this passage is mine. Although closely based on information provided by Professor Mei, the passage is not a direct quotation from his letter.

made during the 1960s.[90] All previous scholarship is unsound so far as Old Sinitic sounds are concerned. For example, at least half of the cognate sets in Wang Li's *Tongyuan zidian* [*Dictionary of Graphs Having Common Origins*] are incorrect or unproven.

This tradition of uncritical and unsound scholarship goes all the way back to the great Wang Niansun (1744-1832) who more or less advocated the principle that "synonyms are likely to be cognates" (cf. *Guangya shuzheng* [*Documented Subcommentaries on* Expanded Elegance]). Wang Niansun's influence extends to the present day. Thus traditional Chinese scholarship is really not of much help in trying to establish cognation in Old Sinitic.

There are several ways to remedy this sad state of affairs: 1. pay more attention to Old Sinitic phonology, 2. utilize epigraphical evidence more fully (e.g., if two words are written with the same graph in OSBIs or in bronze inscriptions, there is *prima facie* evidence that the two words may be cognates), 3. Tibetan morphology may have preserved certain Sino-Tibetan morphophonemic processes which can be used to explain why two words phonologically dissimilar in Middle Sinitic may actually be related.

I totally agree with all of the points made by Professor Mei and believe that his proposals for rectifying the situation – in combination with those of Norman and Coblin – constitute a workable solution to the crisis faced by Chinese historical linguistics today. In other words, we do not have to be pessimistic about the future of our subdiscipline, but do need to adopt radical new strategies in order to overcome the genuine crisis that it is facing.

The only contribution that I might be able to offer in addition to those of Professors Norman, Coblin, and Mei is to focus more squarely on the difficulties created by the nature of the Sinographic

---

[90] Professor Mei's description of the state of the field in Old Sinitic phonology is not to be confused with the situation in Middle Sinitic phonology, where the most important advances occurred in the first half of this century. The various proposed reconstructions of Middle Sinitic are much more accurate and constitute a greater degree of consensus than those for Old Sinitic. Yet even the study of Middle Sinitic is plagued by disagreement over the status of the reconstructed language, such as whether it is an artificial, homogenized construct or whether it approximates a real language spoken in a specific place at a specific time and, if the latter, where and when it was spoken by whom.

script. In numerous articles[91] and lectures, I have alluded to the limitations that the Sinographic script imposes on the study of the history and nature of Sinitic languages. However, since this Appendix is the strongest statement I have made on the subject, it can stand by itself.

The sinographic script is an enormous syllabary[92] and syllabaries inevitably exercise severe constraints upon phonemic representation. Syllabaries are normally composed of symbols representing a CV-type of syllable and a smaller number of symbols representing V-type syllables. Occasionally (the Sinographic script is unusual in this regard), a syllabary may also possess symbols representing CVS/N-type syllables (where C = consonant, V = vowel, S = stop [e.g., -p, -t, -k], and N = nasal). There are also glides (G) which precede vowels in various positions, thus we may have GV, CGV, and so forth. It is extremely rare to find syllabaries with the following types of syllables: VC, CVC, etc. Rarer still are syllabaries that include configurations such as CC (i.e., consonant clusters) in any position.[93] Furthermore, the syllable final stops tend first to collapse into an indeterminate ʔ final and then to disappear entirely, whereas the syllable final nasals are rather volatile and may be dropped or added at different stages in the evolution of a given syllable.

The most serious limitation of syllabic scripts is their inability to represent consonant clusters, complex or otherwise. This is a phenomenon that I have long observed in Sumerian, so I shall illustrate it here with reference to the syllabic cuneiform representation of that

---

[91] Most notably in various articles and reviews published in *Sino-Platonic Papers*.

[92] This is precisely what John DeFrancis proved in his *Chinese Language*. Because each and every unit of the script is neither a picture, an idea, nor a word, we may not specify it as pictographic, ideographic, or logographic. The closest we can come to designating the type of script we are dealing with is to call it "morphosyllabic," for which see DEFRANCIS, *op. cit.*, pp. 88, 125-126, 147, 187, and 196.

[93] Compare the remarks of Peter T. Daniels, an authority on the history of writing systems and a specialist on syllabaries: "...virtually every extant syllabary represents syllables comprising (besides a vowel alone) a consonant (C) followed by a vowel (V), rather than VC or CVC syllables." See his *The Study of Writing Systems*, p. 4.

language. Let us take, for instance, the Sumerian word <gu-ru-un> = <gurun> ("fruit"). Now, the same three syllables (<gu-ru-un>) which are used to transcribe this word could potentially also stand for /kur-un/, /krun/ (realized as [krun] or even [kurun] with a purely phonetic *svarabhakti* [vowel insertion]), /kurn/ (realized as [kurn] or even [kurun] with a purely phonetic *svarabhakti*, /krn/ (realized as [kr̩n] with a syllabic r̩), /skurun/, /skurn/, /skrun/, /skrn/ (realized as [skr̩n]), and so forth and so on. It is apparent that one may readily move from the actual pronunciation to its representation in the syllabic script, but the reverse process is inevitably fraught with irrecoverable loss of phonological data for Sumerologists. I submit that exactly the same process has been going on with Sinitic during the last 3,200 years.

As Boisson has remarked, when syllabaries

are used for languages with phonemic consonant clusters, such as the Indo-European languages, they create a graphic image which cannot do justice to the real prouniation, as is only too obvious in the treatment of Greek in Linear B and the Cypriote syllabary, or in the distortions imposed by the cuneiform syllabary on Hittite or the "Hittite hieroglyphic" syllabary on Luwian.[94]

Consequently,

one should not exclude the possibility that, due to the distortions imposed by the [syllabic] writing system, apparently disyllabic [Sumerian] words with total vowel harmony are in fact monosyllabic words with consonant clusters.[95]

I have no doubt whatsoever that such is the case with Old Sinitic,

---

[94] BOISSON, *Phonotactics of Sumerian*, p. 32. Boisson's article is the source of other information in this paragraph. I have also consulted the well-known grammatical treatise on Sumerian by Thomsen, who makes many of the same points as Boisson equally forcefully. An example of this phenomenon in Hittite is the word for "star", *hasterz* (cf. Greek ἀστηρ, Latin *astēr*, Armenian *astł*), which appears as *ha-as-te-ir-za* in syllabic hieroglyphs and in Mycenean is the word for "quadriped", $*k^w$etropodp^{h}i ($*k^w$ετ-ροποδϕι, which appears in the syllabic script as *qe-to-ro-po-pi*.

[95] *Ibid.*, p. 44.

and the last statement quoted from Boisson could apply just as well to the sets of dimidiate Sinitic words discussed in the second half of this paper and to thousands of others like them.

With Old Sinitic, however, we are actually much better equipped to restore the lost phonological data than with Sumerian. In both cases, the primary means for the restoration of phonological information are two: 1. finding the same words transcribed syllabically in cuneiform or sinographs transcribed at approximately the same time with an alphabet or an alphasyllabary, 2. closely comparing the syllabically transcribed words with the same words in cognate languages, some of which may still be alive and thus capable of being recorded with the precise symbols of the International Phonetic Alphabet. Because Sumerian came to be written down at an early stage in the history of writing when alphabets and alphasyllabaries had not yet been invented, and died out as a living language long before alphabets and alphasyllabaries were in widespread use, the first method for the recovery of lost phonetic information is not applicable except, as it were, "at a distance" (i.e., by comparing Sumerian words with the same words recorded much later in alphabets or alphasyllabaries). Furthermore, since Sumerian is still generally considered to be a linguistic isolate, it is impossible to compare it with any other languages. Thus, unless current efforts to link up Sumerian phyletically with other languages are successful, the prognosis for the full restoration of its phonology may be declared to be truly dismal. In contrast, Sinitic first came to be written down at a time when alphabets had already been invented. What is more, hundreds of cognates with words in Tibeto-Burman languages have been identified.[96] Hence, the prognosis for the eventual recovery of Old Sinitic phonology is good.

We can be deeply grateful to the wise Tibetan king Srong btsan

---

[96] Most, but by no means all, historical linguists working on Sinitic languages maintain that they are genetically linked to Tibetan. There is no doubt that Sinitic is at least partially related to Tibeto-Burman, but it may also have arisen in a complicated cultural context where Austroasiatic, Austronesian, Indo-European, and other language groups exerted a significant impact upon it. In other words, Sinitic may well not be purely and simply a group within the Tibeto-Burman family.

sgam po who, under the guidance of his minister Thon mi Saṃbhoṭa, decided in the first half of the seventh century to select an Indian-derived alphasyllabary rather than a Sinographic script to write the language of his people.[97] Had he opted for the latter, Tibetan historical linguistics would doubtless be in the same sorry situation that Sinitic historical linguistics now finds itself. Instead, we are fortunate to possess all of the vital phonological data preserved in the nuanced orthography of the Tibetan script. Indeed, the rich assemblage of phonological information found in Tibetan, as observed by Tsu-Lin Mei, is one of the most important sources for the recovery of the sounds of Old Sinitic.

Aside from the evidence provided by Tibetan and other cognate languages, there is a great deal of direct transcriptional evidence from Sanskrit, Iranian, Tocharian, and other languages indicating that Old Sinitic must originally have possessed a vastly richer phonology than the modern Sinitic languages. There is also little doubt that it had a greater variety of initial consonant clusters and finals than Middle Sinitic. This recognition of the phonological complexity of Old Sinitic is reflected in the increasingly complicated reconstructions that have been proposed in recent years (e.g., Li Fang-kuei, Axel Schuessler, Edwin G. Pulleyblank, S. A. Starostin, and William H. Baxter) in contrast to earlier proposals such as those of Bernhard Karlgren, Tung T'ung-ho, and Chou Fa-kao.

Despite the great progress that has been made in the reconstruction of the phonology of Old Sinitic, I believe that a fundamental methodological flaw still governs all such efforts. Namely, instead of taking the phonetically determined Sinitic **languages** as the basis for reconstruction, the parameters of the search for Old Sinitic have essentially been stipulated by strict adherence to the closed system of the Sinographic **script** (especially as it was codified in the *Qieyun [Tomic Rhymes]*). The fallacy of this approach is that the conceptualization of the entire enterprise of the reconstruction of phonologically complex Old Sinitic has been determined by the rhyme classes of a relatively impoverished

---

[97] The precise origins of the Tibetan script are contested, but its systematic similarities to North Indian scripts cannot be denied. See VAN DER KUIJP, *Tibetan Script*.

system. To this day, all researchers who have made specific proposals for the reconstruction of Old Sinitic have proclaimed that they are working backword from the phonological standards of Middle Sinitic as classified by the *Qieyun* system. This will not do.

I believe that the Sinographic system, particularly in its later, developed stages, is a mortal trap for those who wish to reconstruct Old Sinitic. Unless it is abandoned in favor of the phonological systems of the Sinitic languages themselves, a satisfactory reconstruction of Old Sinitic will never be achieved. Instead of such heavy and primary reliance upon the sinographs and their *Qieyun*ian categorizations, I propose the following steps toward the recovery of the sound system of Old Sinitic. They are listed in order of importance:

1. Direct recording and analysis of the sounds of all the living Sinitic topolects.

2. Reconstruction of the Old Min, Old Cantonese, Old Wu, Old Mandarin, and other branches of Sinitic.

   a. Through comparison of the totality of the still living dialects (from step 1).

   b. From data found in old texts, including the records of missionaries and other early visitors to China.

3. Identification and analysis of very early loan-words in Sinitic, such as those for "milk," "magus," "honey," "river (*jiang*)," "crossbow," "wheat," "lion," and so forth.

4. Thorough analysis of Sinitic languages at diverse periods of history as transcribed in Brahmi, Tibetan, Tangut, Khitan, Latin, Arabic, Runic, and other scripts.

5. Thorough analysis of Sanskrit, Sogdian, Khotanese, Turkic, Uyghur, Khitan, Tibetan, Tangut, Thai, Mongolian, and other terms from diverse periods of history as transcribed in Sinographs.

6. Thorough analysis of the vernacularisms of the late Classical and Medieval periods.

7. Phonological analysis of Sino-Tibetan and other Tibeto-Burman cognates.

8. Examination of the rhyming, assonance, consonance, and other prosodic features of all extant literature starting from the OSBIs.[98]

---

[98] The recent dissertation of Wolfgang Behr, in which he analyzes the rhymed pas-

9. Paying more strict attention to variant orthographies and miswritings as evidenced in manuscripts (e.g., from Mawangdui and Dunhuang), in stele inscriptions (e.g., from the Northern Wei), and as recorded in the vast commentarial and lexicographical traditions of China.

10. Utilization of the *duruo* ("read as"), *fanqie* ("countertomy"), and other traditional methods for indicating the sounds of sinographs. This would include the 30 (later increased to 36 by Song period scholars) "letters" (*zimu*) of the nascent, but unfortunately also abortive, spelling system devised by the late Tang monk Shouwen as inspired by the principles of Indian-based alphasyllabaries.

11. Tentative establishment and analysis of "word families" in Sinitic.

12. Analysis of the meager and imprecise phonological and semantic data available from the Sinographic script itself ("radicals," phonophores, etc.).

13. Consultation of the *Qieyun* and its successors.

All of these techniques are already being employed for the purpose of the reconstruction of Old Sinitic, except no. 8 which has only been used in a limited fashion. Unfortunately, until the manifesto of Norman and Coblin, they have almost always been applied in exactly the reverse order!

It is a tragic irony that the enormous efforts which have been expended to reconstruct the sounds of Old Sinitic have been and still are to this day delimited chiefly by the least useful tool for that purpose. So long as Old Sinitic reconstructions are governed primarily by the artificial and abstract *Qieyun* system, they can only amount to "anti-reconstructions" (to use David Prager Branner's brilliant term). I refer to the *Qieyun* system as "artificial and abstract" in the sense of its not being demonstrably grounded upon any real language that was ever alive (i.e., spoken by living human beings in the midst of their day-to-day activities) in a certain place and at a certain time. The *Qieyun*

---

sages of bronze inscriptions, is an exciting initiative that bodes well for the future of efforts to reconstruct Old Sinitic. Nonetheless, even though he directly examines the rhymes in 197 bronze inscriptions dating from the late 11th through the 3rd centuries BCE, he still feels compelled to measure his otherwise sophisticated findings against the traditional Middle Sinitic rhyming categories.

is a handy compendium of temporally and spatially homogenized Middle Sinitic as codified in 600 CE. As such, it is a sort of bench mark for phonological research on Middle Sinitic. Nonetheless, when it comes to the reconstruction of Old Sinitic or the description and analysis of regional and local languages in China, it harbors grave deficiencies. I am not opposed to utilizing the phonological data embodied in the *Qieyun* and its successors, only to making them the sole standards of correctness in research on Sinitic historical phonology.

The *Qieyun* served Karlgren (and those who followed him) well in the preparatory stages of the reconstruction of Middle Sinitic, but we must now move on to a more nuanced description. No longer should we consider the *Qieyun* to be *primum inter pares*, not even for Middle Sinitic languages. (I use the plural advisedly, because I believe that it is now possible to embark upon reconstructions for the various branches of Middle Sinitic. We should not continue to imagine that Sinitic was utterly monolithic throughout time and space. Not only do we have to contend with Literary Sinitic and Written Vernacular Sinitic, which are already very different grammatically, syntactically, and lexically between themselves. We also must accept the reality that the host of spoken Sinitic languages, which collectively were [and are] of a quite disparate nature than either Literary Sinitic or Written Vernacular Sinitic, differ markedly among themselves and each experienced a unique historical development of its own.) Furthermore, since the main aim of the *Qieyun* and its successors is classification and categorization, the data they present must be rigorously and critically scrutinized by testing them against data more directly derived from living languages. The success of any scientific enterprise is decided by how well the typologies it employs explain the data rather than by how well the data fit the typologies. That is to say, the data are primary and the typologies are secondary. Therefore, I have inverted the list of priorities for the reconstruction of Old Sinitic. I am not advocating the total rejection of the *Qieyun* system, only a readjustment of priorities. It is more a matter of emphasis than one of utter exclusion or abolition. The *Qieyun* should not be worshipped as the *sine qua non* for the reconstruction of Old Sinitic. It should be consigned to the position that it deserves.

In addition to the thirteen legitimate and tested methods for the reconstruction of Old Sinitic listed above, I would like to advocate the adoption of a new device. That is the restoration of Old Sinitic words and etyma by reversing the process of dimidiation as illustrated in the second section of this paper. Through the application of this method ("reverse dimidiation"), we may obtain undimidiated forms independent of the sinographic script (no characters should be expected to exist for writing the resultant roots or etyma). I would rank reverse dimidiation eighth or higher on the list of methods for the reconstruction of Old Sinitic. Of course, due to different possible reconstructions of monosyllables for the same pair of syllables (e.g., *mat* + *lat* = *mlat*, *malt*, etc.), and also depending upon the reliability of the EMS or other reconstructions on the left side of the equation, we must always check our results against reconstructions obtained by other means. Such, however, is true with all of the available methods for Old Sinitic reconstructions: the results obtained by one method must be checked against those obtained by all other methods.

There are thousands of dimidiate words dating from the 1st millennium BCE and the 1st millennium CE. The largest group of such words stem from the Han period (roughly the two centuries before and the two centuries after the beginning of the Common Era). Since they come in sets that consist of an average of a bit less than four members, reversing the process of dimidiation would yield a substantial body of roots or etyma. This would represent an extremely valuable fund of phonological and semantic data for the reconstruction of Old Sinitic.

If (and only if) historical linguists engaged in the study of Old Sinitic jettison their overdependence on the sinographically determined Middle Sinitic rhyme categories and employ the other thirteen methods (twelve plus reverse dimidiation) outlined above, there is no reason why their reconstructions should not be at least as accurate as those for Proto-Indo-European which, after all, is at a time depth approximately 3,000 years earlier. On the other hand, no amount of heroic contortions will overcome the limitations of the script, so long as scholars regard its 7th-century and later codifications in rhymebooks as the final arbiters of phonological rectitude.

The sooner historical linguists abandon overreliance on the

closed system of sinographs (i.e., "internal reconstruction" as determined exclusively by the rhyme classes of the *Qieyun* and its successors), the sooner will they achieve a satisfactory and convincing recovery of the phonology of Old Sinitic. Ultimately, once the phonology of Old Sinitic has been recovered, I believe that the following overall picture of the development of Sinitic languages will emerge:

1. At the time of the invention of the Sinographic script around 1300 BCE, Sinitic words were mostly (perhaps almost exclusively) monosyllabic, while the phonology was complex, with a rich assortment of consonant clusters and other configurations, and the morphology was by no means simple.

2. When the syllabic script became more firmly established as well as politically and culturally influential, the complexity of the phonological system and morphological structure of Sinitic was gradually and correspondingly reduced. The entrenchment of the script and the phonological reduction were not necessarily causally linked.

3. Dimidiation was employed as a device for the ersatz preservation of consonant clusters and other complex phonological configurations in the face of a reduced phonetic inventory.

4. Bisyllabicization of different sorts and polysyllabicization of greater lengths increased in proportion as the phonology was levelled. This compensatory phenomenon was inevitable for two reasons: a. to prevent misunderstandings in oral communication due to hyperhomophony at the syllable level, b. to keep up with the flood of scientific, technological, and cultural innovations.

Jerome L. Packard of the University of Illinois is engaged in a project to clarify the multisyllabicization of Sinitic in stages 3 and 4. It is necessary to note that, whereas the sinographic script may be largely (nearly entirely) monosyllabic, Sinitic languages have certainly been far from monosyllabic during the last one and a half millennia. Already during the first millennium BCE, it is easy to point to many truly bisyllabic words (words in which the syllables are tightly bound and often have no meaning when isolated, e.g., *hudie* ["butterfly"], *zhizhu* ["spider"], *jiejue* ["wriggler; the larva or pupa of the mosquito"], and *qiuyin* ["earthworm"], to cite only a few examples from entomology and annelids).

I shall conclude this Appendix with a few explanatory comments on Fig. 1 ("The graphs and words of Sinitic correlated with literary versus vernacular forms of writing"). First, a few caveats: whereas the solid line indicating the proliferation of sinographs during the last 3,200 years is for the most part quite precise, the broken line showing the increase in the number of bisyllabic and polysyllabic words is based on estimates of the contents of dozens of dictionaries and databases compiled during the last two millennia (some of which deal with materials dating back to the known beginning of writing in China). The line of dots and dashes indicating the divergence of Literary Sinitic and Vernacular Sinitic written languages is determined by comparison of the proportion of literary and vernacular elements (including lexical items, morphological structures, and grammatical / syntactical features) in the totality of all written texts from the last 3,200 years (based on representative samples, historical grammars, etc.). The degree of precision of the latter two lines being less than that of the first line, they are shown as smoothed-out curves rather than as having shifting angles. This graph is an initial, tentative attempt to visualize the relationship between the development of the Sinographic script and the Sinitic lexicon. It obviously needs to be refined by future researchers, yet even at this early stage of research we may draw several preliminary conclusions:

1. Because the script begins with such a large number of components, either the notion of writing must have been imported to China or previous stages have not been archeologically recovered. The isolated marks on pots, shards, and other objects (similar to those found elsewhere in the world from preliterate times) dating to before the era of the OSBIs does not constitute **writing**, which I define as **the representation of language on a surface.**

2. The most striking discovery yielded by the creation of Fig. 1 is the very close correlation which it reveals among script, lexicon, and type of written language. (After I had connected the dots and drawn in the lines, I literally nearly fell out of my seat with astonishment when I saw how tightly the three lines coincided.) Even at this precursory stage of research, it is clear that these fundamental components of writing (at least for Sinitic) are intimately linked. This conclusion should be tested against writing in other cultures.

3. We may draw no hard and fast conclusions from Fig. 1 regarding possible divergence between spoken Sinitic languages and various writing styles (literary, vernacular, and mixed). Fig. 1 tells us nothing specifically or directly about speech, only about writing. In other words, we may not conclude from Fig. 1 that spoken language and written language coincided fairly closely until about 400 CE, nor may we conclude that spoken Sinitic languages before the beginning of the Common Era were largely monosyllabic.[99] Such propositions may be true, but they will have to be determined by other means. Fig. 1 deals strictly with written forms of language.

---

[99] There is undoubtedly a much closer correlation between spoken Sinitic languages and Written Vernacular Sinitic than there is between any of the spoken languages and Literary Sinitic. The difficulty in drawing firm conclusions in this regard for pre-modern times is due to the fact that there are so few reliable records of spoken Sinitic before this century. Written Vernacular Sinitic itself should not be confused with any variety of spoken Sinitic. For example, even colloquial Pekingese is simply not writable in sinographs, as the most famous author of that idiom, Lao She, bitterly complained. For detailed evidence, see the many reviews of Pekingese dictionaries that I have written for *Sino-Platonic Papers*. Since this is true of Pekingese, which is the nearest model for modern Written Vernacular Sinitic, one can well imagine what the situation must be like when someone tries to write down Cantonese, Amoy, Shanghainese, or any of the other many non-Mandarin languages. Most people merely assume that if something can be spoken in one of the Sinitic languages it can be written in Sinographs. This is far from true, hence many individuals who are trying to devise a workable script for written Taiwanese advocate the adoption of Romanization or partial Romanization. For all practical purposes, the non-Mandarin topolects (*fangyan*) simply are not written down. Except as a kind of *tour de force* employing many nonce characters and intelligible only to an extremely limited circle of initiates, speakers of non-Mandarin Sinitic languages either must write in Literary Sinitic (rare nowadays) or must learn to write Written Vernacular Sinitic (based upon Modern Standard Mandarin).

The problems with writing spoken languages in a morphosyllabic script are numerous and beyond the scope of the present article. Suffice it for the moment just to mention the following: 1. semantic distortion when the symbols of the script are used to "spell" new words whose morphemes were not previously written down (whether arising internally or borrowed from an external source), 2. lack of recognized conventions for phonetic or phonemic representations, 3. inability to convey distinctions and variations through time and across space (topolects). There is inevitably a gap between spoken and written forms of language, no matter what type of script is used. However, the gulf between spoken and written forms of Sinitic is far greater than that between spoken and written forms of languages using alphabetic scripts. Hence the ubiquitous use of subtitles in China (even in Peking and surrounding areas) in situations where they would not be needed in countries using alphabetic

4. Since approximately the beginning of the Common Era, the number of bisyllabic and polysyllabic words has exceeded the number of Sinographs by a multiple of between 5 and 10. So long as the Chinese people continue to use the Sinographic script as their main system of writing, ratios of this magnitude will likely persist. This means that, within another few centuries, dictionary-compilers, font-makers, and Internet-designers will conceivably have to contend with upwards of 100,000 characters.

· 5. That the divergence factor is keyed so closely both to the number of Sinographs and to the number of bisyllabic and polysyllabic words reveals a profound psycholinguistic verity about the inescapable linkages among homophony, multisyllabicity, and the capacity of the brain to process written language.

6. Until about the 16th century, it is evident that the number of bisyllabic and polysyllabic words lagged behind the ratio of ten to one more than the number of Sinographs, but after that date they exceeded that ratio slightly. Likewise, the rise in the divergence factor has all along lagged slightly behind the rise in the number of bisyllabic and polysyllabic words. These lag times disclose two things:

a. Until about 1500, the increase in the number of bisyllabic and polysyllabic words may be considered to be at least partially due to the increase in the number of Sinographs. Similarly, the rise in the divergence factor may be explained as at least partially the result of the increase in the number of bisyllabic and polysyllabic words.

---

scripts (e.g., films, reading the evening news or weather report), the frequent need for repetition or restatement, the constant requests for differentiation of homophones (which *yi*? which *shi*?), and so forth.

In terms of their ability to represent spoken languages with facility, morphosyllabaries have all of the drawbacks of syllabaries mentioned above. In addition, they have the liability of constant semantic interference. They also tend to have extraordinarily multitudinous discrete symbols numbering in the thousands or tens of thousands (in contrast to syllabaries which usually have between about fifty to several hundred symbols and alphabets which normally have only a score or two of separate symbols), each of which is complex, consisting of an average of twelve or more strokes in the case of the Sinographs but equally exacting to draw in the other morphosyllabaries (in contrast to syllabaries which customarily have three or four strokes and alphabets which mostly have about two strokes per symbol). On the other hand, morphosyllabaries (e.g., the Sinographs, Egyptian hieroglyphics, and Mayan glyphs) have the virtues of being esthetically beautiful, monumental scripts perfected and treasured by elite priests, scholars, and scribes for centuries.

*b.*The 16th century, when the rate of increase in the number of Sinographs was exceeded by the rate of increase in the number of bisyllabic and polysyllabic words and by the rate of change in the divergence factor, witnessed the beginning of a massive explosion of written vernacular fiction and drama. There were even halting efforts to write in some of the topolects (e.g., Wu and Min).

7. The phenomena described in 6a. and 6b. are purely linguistic and grammatological. There were also historical forces involved. Ultimately it is the historical (i.e., political, economic, and cultural) forces that were actually the determining factors in the changes experienced in the development of writing. The four main stages in the development of writing in China, all of which are salient in Fig. 1, are:

a. Its beginning around the 13th century BCE.

b. A marked increase in dimidiate and other types of bisyllabic vocabulary starting in about the middle of the first millennium BCE and the unification of the script during the third century BCE.

c. The rise of Written Vernacular Sinitic during the medieval period. This is the period when Sino-Indian Buddhist narrative prospered.

d. The eventual triumph of the written vernacular over Literary Sinitic which is already a foregone conclusion by the 16th century.

All four of these monumental changes in writing in China were precipitated by the impingement of external agencies. These were:

a. Around the middle of the 2nd millennium BCE, Eurasian nomads who brought with them chariots, bronze metallurgy, and other vital components of culture.[100]
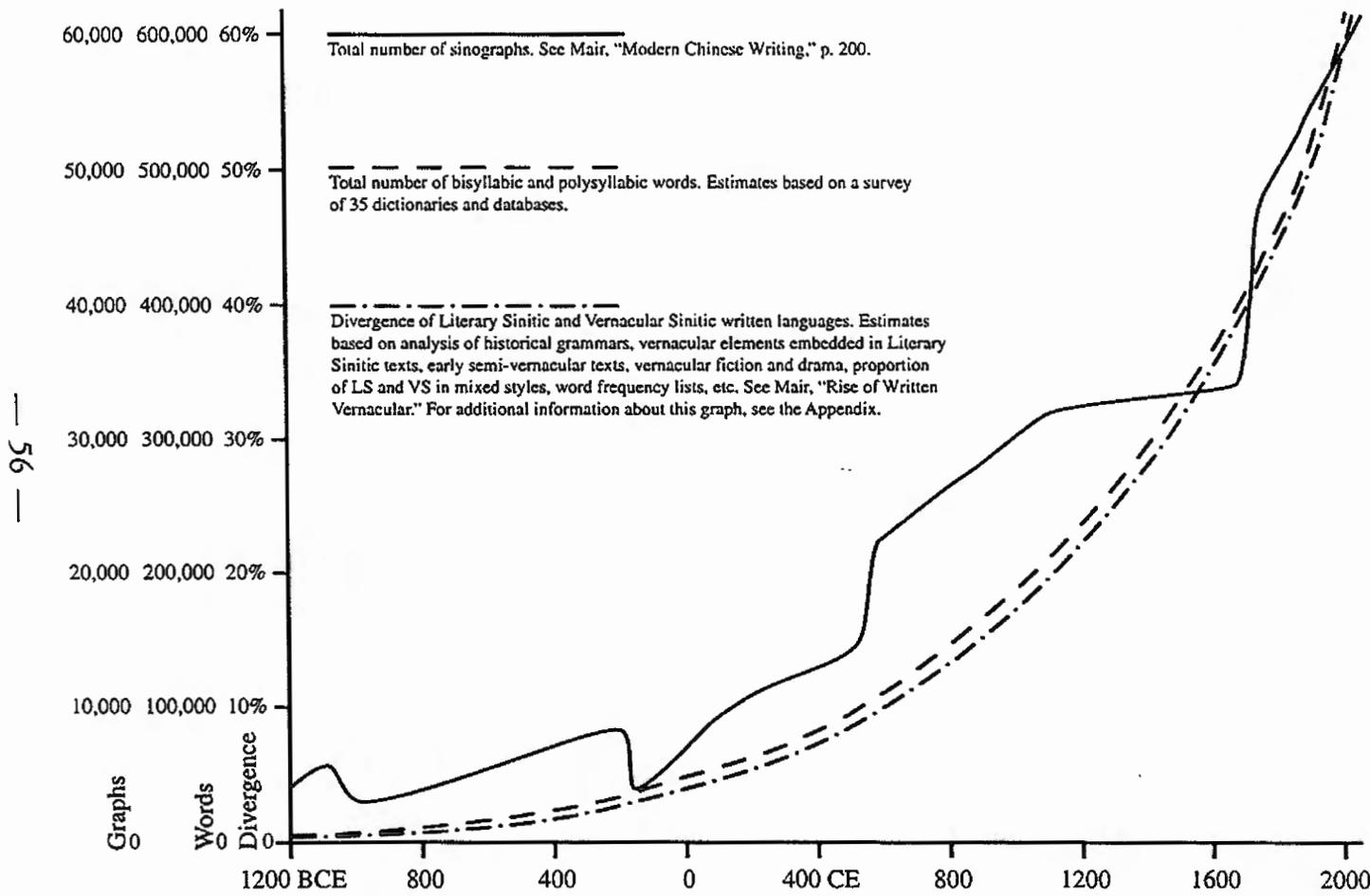
b. The introduction of iron, mounted warfare, and advanced weaponry, together with stimulating intellectual currents flowing into China from the northwest, west, southwest, and perhaps also from the southeast via the ocean. This is the golden age of Chinese thought known as the Warring States period.[101]

c. Buddhism (and, to a lesser extent, other foreign religions such as Manicheism, Zoroastrianism, and Nestorian Christianity) transmitted to Chi-

---

[100] See the works cited in note 76.

[101] Here I wish to pay tribute to the Warring States Project which is headquartered at the University of Massachusetts. Although underfunded, this pathbreaking academic endeavor led by Bruce Brooks is totally reshaping our understanding of the formative age of Chinese politics and culture.

5

Figure 1: The graphs and words of Sinitic correlated with literary versus vernacular forms of writing.

Total number of sinographs. See Mair, "Modern Chinese Writing," p. 200.

Total number of bisyllabic and polysyllabic words. Estimates based on a survey of 35 dictionaries and databases.

Divergence of Literary Sinitic and Vernacular Sinitic written languages. Estimates based on analysis of historical grammars, vernacular elements embedded in Literary Sinitic texts, early semi-vernacular texts, vernacular fiction and drama, proportion of LS and VS in mixed styles, word frequency lists, etc. See Mair, "Rise of Written Vernacular." For additional information about this graph, see the Appendix.

©V.H.Mair 97/8/22

na by Indians, Iranians, Tocharians, and other peoples from the "Western Regions."

d. The huge numbers of West Asians and Central Asians (Persians, Arabs, Uyghurs, and other so-called *semu* ["colored eyes"][102]) brought to China by the Mongols and installed in positions of power, followed closely by European adventurers, commercial entrepreneurs, and missionaries. There seems to have been a lull of sorts after the arrival of the former, at least in the production of new Sinographs, but the rise in all three lines approaches the perpendicular after the arrival of the latter.

It is apparent from Fig. 1 that vital encounters with external groups led to fundamental transformations in the development of writing. In this sense, Chinese history is no different from the history of other countries: interaction with neighboring cultures leads to changes in society and culture (including language).

8. It is apparent that the Sinographic script is reaching some sort of peak (one might almost say that we are witnessing a concatenation of spikes). It is impossible to predict with assurance what will happen in the coming decades and centuries, but the current trajectories cannot continue indefinitely. In the first place, at the present rate of change, there will be 100% divergence between Literary Sinitic and Written Vernacular Sinitic within a couple of centuries at most. When that happens, Literary Sinitic will be relegated to classical studies; the demands voiced by reformers at the beginning of this century for the replacement of Literary Sinitic by written vernacular will have been fulfilled. Secondly, word production has outstripped the capacity of the Sinographic script to carry the burden. The multiplication of Sinographs cannot go on unchecked indefinitely. Eventually, limitations (probably from the government, as in Japan) will be imposed on the number of permissible Sinographs. Third, the requirements of electronic information processing and transmission, the burgeoning lexicon, and the potential flourishing of written topolects (comparable to the situation in Europe or India) will probably result in the collapse of the Sinographic system for written vernacular Sinitic languages and its replacement by an alphabetic script. The Si-

---

[102] These individuals are now innocuously and misleadingly referrred to as "[people of various] categories." I am preparing a separate paper, complete with references to Mongolian texts, which will show that the term *semu* was actually based on a Mongolian term that means exactly what the Sinographs designate: "[people having unusual] colored eyes" (viz., not like the majority of Mongolians and Chinese).

nographic script will continue in use, of course, in classical studies, calligraphy, ornamentation and decoration, and the like.

Some may question why a paper on the history of Chinese literature has devoted so much space to what appears to be a purely linguistic topic. My response would be that the study of literature cannot be separated from the study of linguistics (just as the study of literature cannot be divorced from the study of religion, art, history, and numerous other relevant disciplines). Certainly, we cannot fully understand what Su Shi meant by *laza* unless we tackle head on the problem of cognates in Sinitic, and that is a problem which can only be solved by linguistics.

BIBLIOGRAPHY

The American Heritage Dictionary of the English Language, 3rd ed. Boston, New York, London, Houghton Mifflin, 1992.

BARNHART, ROBERT K., *The Barnhart Dictionary of Etymology*, n.p., H. W. Wilson, 1988.

BAXTER, WILLIAM H., *A Handbook of Old Chinese Phonology*, (Trends in Linguistics. Studies and Monographs, 64), Berlin, New York, Mouton de Gruyter, 1992.

BEHR, WOLFGANG, *Reimende Bronzeinschriften und die Entstehung der chinesischen Endreimdichtung*, 2 vols., Ph.D. dissertation in Sinology. Frankfurt a.M., Johann Wolfgang Goethe-Universität, 1996-1997. Extended abstract in «Cahiers de Linguistique Asie Orientale», 26,1, 1997, pp. 165-170.

BIRRELL, ANNE, *Popular Songs and Ballads of Han China*, London, Sydney, Wellington, Unwin Hyman, 1988.

BOISSON, CLAUDE PIERRE, *The Phonotactics of Sumerian*, in IRÉN HEGEDŰS, PETER A. MICHALOVE, and ALEXIS MANASTER RAMER, ed., *Indo-European, Nostratic, and Beyond: Festschrift for Vitalij V. Shevoroshkin*. Being «Journal of Indo-European Studies», Monograph Number 22, Washington, Institute for the Study of Man, 1997, pp. 30-50.

BOODBERG, PETER A., *'Ideography' or Iconolatry?*, in COHEN, ed., *Selected Works of Peter A. Boodberg*, pp. 407-429. Originally in «T'oung Pao», 35, 4, 1940, pp. 266-288.

— *Some Proleptical Remarks on the Evolution of Archaic Chinese*, in COHEN, ed., *Selected Works of Peter A. Boodberg*, pp. 363-406. Originally in «Harvard Journal of Asiatic Studies», 2, 1937, pp. 329-372.

CHANG HSÜAN. See ZHANG XUAN.

CHANG, TSUNG-TUNG, *Indo-European Vocabulary in Old Chinese: A New Thesis on the Emergence of Chinese Language and Civilization in the Late Neolithic Age*, «Sino-Platonic Papers», 7, January 1988, 56 pages.

CLAUSON, GERARD, *An Etymological Dictionary of Pre-Thirteenth-Century Turkish*, Oxford, Clarendon, 1972.

COBLIN, WELDON SOUTH, *A Sinologist's Handlist of Sino-Tibetan Lexical Comparisons*, «Monumenta Serica Monograph Series», 18, Nettetal, Steyler, 1986.

COHEN, ALVIN, ed. *Selected Works of Peter A. Boodberg*, Berkeley, Los Angeles, London, University of California Press, 1979.

DANIELS, PETER T., *The Study of Writing Systems*, in PETER T. DANIELS and WILLIAM BRIGHT, ed., *The World's Writing Systems*, New York, Oxford, Oxford University Press, 1996, pp. 3-17.

DEFRANCIS, JOHN, *ABC Chinese-English Dictionary*, Honolulu, University of Hawaii Press, 1996.

— *The Chinese Language: Fact and Fantasy*, Honolulu, University of Hawaii Press, 1984.

DUAN KAILIAN, comp., *Zhongguo minjian fangyan cidian [Dictionary of Chinese Folk Topolecticisms]*, Haikou, Nanhai chubanshe, 1994.

FASMER, MAX, *Etimologicheskii Slovar' Russkogo Yaz'ika*, 4 vols., Moscow, Progress, 1986-1987. Translation from the German *Russisches Etymologisches Wörterbuch* (Heidelberg, 1950-1958) by O. N. TRUBACHEV.

Foguang da cidian [Buddha Light Encyclopedic Dictionary], CIYI, ed., 8 vols., Kaohsiung, Foguang chubanshe, 1989.

FU DINGYI, *Lianmian zidian [Dictionary of Bisyllabic Terms]*, 4 vols., 1954, 2nd ed.; 1983, 3rd printing.

GAO SHUFAN, ed., *Zhengzhong xing yin yi zonghe da zidian [Zhengzhong Large, Comprehensive Dictionary of the Shapes, Sounds, and Meanings of Single Characters]*, Taipei, Zhengzhong shuju, 1971; rev. and enlgd. ed., 1974.

GAO WENDA, chief ed., *Jindai Hanyu cidian [Dictionary of Recent Sinitic]*, Peking, Zhishi chubanshe, 1992.

GILES, HERBERT A., *A Chinese Biographical Dictionary*, London, B. Quaritch; Shanghai, Kelly & Walsh, 1898.

GREENBERG, JOSEPH H., *Does Altaic Exist?*, in IRÉN HEGEDŰS, PETER A. MICHALOVE, and ALEXIS MANASTER RAMER, ed., *Indo-European, Nostratic, and Beyond: Festschrift for Vitalij V.* Shevoroshkin. Being «Journal of Indo-European Studies, Monograph Number 22», Washington, Institute for the Study of Man, 1997, pp. 88-93.

GREGORY, PETER N., tr. and annot., *Inquiry into the Origins of Humanity: An Annotated Translation of Tsung-mi's* Yüan jen lun *with a Modern Commentary*, Kuroda Institute, Honolulu, University of Hawaii Press, 1995.

— *Tsung-mi and the Sinification of Buddhism*, Princeton, Princeton University Press, 1991. .

Gwoyeu tsyrdean [National Language Dictionary], 4 vols. WANG YI, ed., Shanghai, Shangwu yinshuguan 1927, 1st ed.; *Chorngbian Gwoyeu tsyrdean [Recompiled National Language Dictionary]*, 6 vols. Compiled by JIAOYUBU CHORNGBIAN TSYRDEAN BIANYI WEIYUANHUI *[Ministry of Education Committee for the Recompiled National Language Dictionary]*, Taipei, Taiwan Shangwu yinshuguan, 1982, 6th ed.

HADINGHAM, EVAN, *The Mummies of Xinjiang*, «Discover», 15, 4 (April, 1994), pp. 68-77.

Hanyu da cidian [Great Word Dictionary of Sinitic], LUO ZHUFENG, chief ed., 12 vols, plus index vol. (1995), Shanghai, Hanyu da cidian chubanshe and Hong Kong, Sanlian shudian, 1987-1994.

Hanyu da zidian [Great Character Dictionary of Sinitic], XU ZHONGSHU, chief ed., 8 vols., Wuhan, Hubei cishu chubanshe and Sichuan cishu chubanshe, 1986-1990.

Hanyu fangyan cihui [Lexicon of Sinitic Topolects], BEIJING DAXUE ZHONGGUO YUYAN WENXUEXI YUYANXUE JIAOYAN SHI [Peking University Chinese Language and Literature Department, Linguistics Instructional and Research Section], ed., Peking, Yuwen chubanshe, 1995. The first edition was published by Wenzigaige chubanshe (Peking, 1964).

HUANG BUFAN, Editor-in-Chief, *Zang-Mian yuzu yuyan cihui (A Tibeto-Burman Lexicon)*, Peking, Zhongyang minzu xueyuan chubanshe, 1992.

IRIYA YOSHITAKA, *Henbun nisoku* (Two Materials of *Bian Wen*), in *Torii Hisayasu sensei kakō kinen ronshū – Chūgoku no gengo to bungaku (Studies in Chinese Literature and Language Dedicated to Professor Torii Hi-*

sayasu on His Sixty-First Birthday), Nara, Tenri daigaku, 1972, pp. 143-150.

— and KOGA HIDEHIKO, eds., *Zengo jiten [Dictionary of Zen Terms]*, Tokyo, Shibunkaku, 1991.

KANAOKA SHŌKŌ, *Tonkō no bungaku [Literature of Tun-huang]*, («Okura sensho [Okura Selected Works]»), Tokyo, Okura shuppan, 1971.

— *Tonkō no bungaku bunken [Dunhuang Literary Documents]*, («Koza Tonkō [Lectures on Dunhuang]», 9), Tokyo, Daito, 1990.

— *Tonkō no emonogatori [Illustrated Tales from Tun-huang]*, Tokyo, Tōhō shoten, 1981.

KARLGREN, BERNHARD, *Grammata Serica Recensa*. Reprinted from «Bulletin of the Museum of Far Eastern Antiquities», 29 (Stockholm, 1957).

KOMAZAWA DAIGAKU NAI ZENGAKU DAIJITEN HENSANSHO *[Compilation Office for the Great Dictionary of Zen Buddhism]*, *Zengaku daijiten [Great Dictionary of Zen Buddhism]*, 3 vols., Tokyo, Taishukan shoten, 1978.

KONG FANLI, ed., *Su Shi wenji [Collected Prose of Su Shi]*, 6 vols., Peking, Zhonghua shuju, 1986.

KUIJP, LEONARD W. J. VAN DER, *The Tibetan Script and Derivatives*, in PETER T. DANIELS and WILLIAM BRIGHT, ed., *The World's Writing Systems*, New York, Oxford, Oxford University Press, 1996, pp. 431-441.

LEHMANN, WINFRED P., *A Gothic Etymological Dictionary* (Based on the third edition of *Vergleichendes Wörterbuch der Gotischen Sprache* by SIGMUND FEIST), Leiden, E. J. Brill, 1986.

LI LEYI, *Jianhua ziyuan [The Origins of Simplified Chinese Characters]*, Peking, Huayu jiaoxue chubanshe, 1996.

MAIR, VICTOR H., *Buddhism and the Rise of the Written Vernacular: The Making of National Languages*, «Journal of Asian Studies», 53, 3 (August, 1994), pp. 707-751.

— *Chinese Popular Literature from Tun-huang: The State of the Field (1980-1990)*, in ALFREDO CADONNA, ed., *Turfan and Tun-huang, the Texts: Encounter of Civilizations on the Silk Route*, («Orientalia Venetiana», IV), Florence, Leo S. Olschki, 1992, pp. 171-240.

— *East Asian Round-Trip Words*, «Sino-Platonic Papers», 34 (October, 1992), pp. 5-13.

— *Language and Script: Biology, Archaeology, and (Pre)history*, «International Review of Chinese Linguistics», 1, 1, 1996, pp. 31-41, 47-50.

— *Lay Students and the Making of Written Vernacular Narrative: an Inventory of Tun-huang Manuscripts*, «Chinoperl Papers», 10 (1981), pp. 5-96.

— *Modern Chinese Writing*, in PETER T. DANIELS and WILLIAM BRIGHT, ed., *The World's Writing Systems*, New York, Oxford, Oxford University Press, 1996, pp. 200-208.

— *Mummies of the Tarim Basin*, «Archaeology», 48, 2 (March-April, 1995), pp. 28-35.

— ed., *The Mummified Remains Found in the Tarim Basin*. Being «The Journal of Indo-European Studies», 23, 3-4 (Fall/Winter, 1995), pp. 279-444.

— *The Narrative Revolution in Chinese Literature: Ontological Presuppositions*, «Chinese Literature: Essays, Articles, Reviews», 5, 1 (July, 1983), pp. 1-27.

— *Painting and Performance: Chinese Picture Recitation and Its Indian Genesis*, Honolulu, University of Hawaii Press, 1988; paperback ed., 1996.

— *A Partial Bibliography for the Study of Indian Influence on Chinese Popular Literature*. Being «Sino-Platonic Papers», 3 (March, 1987), pp. IV + 214.

— *Records of Transformation Tableaux* (pien-hsiang), «T'oung Pao», 72, 1-3 (1986), pp. 3-43.

— *Southern Bottle-Gourd* (hu-lu) *Myths in China and Their Appropriation by Taoism*, in *Zhongguo shenhua yu chuanshuo xueshu yantaohui lunwenji [Proceedings of the Conference on Chinese Myth and Legend]*, («Hanxue yanjiu zhongxin congkan [Center for Chinese Studies Research Series]», no. 5), Vol. 1 of 2, Taipei, Hanxue yanjiu zhongxin, 1996, pp. 185-228.

— *T'ang Transformation Texts: A Study of the Buddhist Contribution to the Rise of Vernacular Fiction and Drama in China*, («Harvard-Yenching Institute Monograph Series», 28), Cambridge, Massachusetts, Harvard University Council on East Asian Studies, 1989.

— *Tun-huang Popular Narratives*, («Cambridge Studies in Chinese History, Literature and Institutions»), Cambridge, Cambridge University Press, 1983.

— tr. *What Is* Pien-wen?. See PAI HUA-WEN.

MIN JIAJI, CHAO JIZHOU, and LIU JIEMING, comp., *Hanyu fangyan changyong-*

ci cidian [Dictionary of Frequently Used Words in Sinitic Topolects], n.p., Zhejiang jiaoyu chubanshe; 2nd printing, 1995.

MOROHASHI TETSUJI, *Dai Kan-Wa jiten [Great Sinitic-Japanese Dictionary]*, 13 vols., Tokyo, Taishūkan shoten, 1955-1960.

NAKANO TATSUEI, *et al.*, comp., *Dai Nippon zoku zōkyō [Great Japanese Continuation of the Tripiṭaka]*, 750 vols., Kyoto, Kyōto zokkyō shoin, 1905-1912.

NIENHAUSER, WILLIAM H., Jr., *The Indiana Companion to Traditional Chinese Literature*, Bloomington, Indiana University Press, 1986.

NORMAN, JERRY L and W. SOUTH COBLIN, *A New Approach to Chinese Historical Linguistics*, «Journal of the American Oriental Society», 115, 4 (1995), pp. 576-584.

OWEN, STEPHEN, *An Anthology of Chinese Literature: Beginnings to 1911*, New York, London, W. W. Norton, 1996.

PAI HUA-WEN, *What Is Pien-wen?*, (translated by Victor H. Mair), «Harvard Journal of Asiatic Studies», 44, 2 (December, 1984), pp. 493-514.

PULLEYBLANK, EDWIN G., *Lexicon of Reconstructed Pronunciation in Early Middle Chinese, Late Middle Chinese, and Early Mandarin*, Vancouver, University of British Columbia Press, 1991.

The Random House Dictionary of the English Language, 2nd ed., unabridged., New York, Random House, 1987.

REISCHAUER, EDWIN O., tr. and annot., *Ennin's Diary: The Record of a Pilgrimage to China in Search of the Law*, New York, Ronald Press, 1955.

SCHUESSLER, AXEL, *A Dictionary of Early Zhou Chinese*, Honolulu, University of Hawaii Press, 1987.

SHANSKII, N. M., V. V. IVANOV, T. V. SHANSKAYA, comp., *Kratkii Etimologecheskii Slovar' Russkogo Yaz'ika: Posobie dlya Uchitelya*, Moscow, Prosveschenie, 1971, 2nd rev. and enlgd. ed.

SHEN JIANSHI, chief ed., *Guangyun shengxi [The Sound System of the Expanded Rhymes]*, Taipei, Dahua shuju, 1977.

SKEAT, WALTER W., *An Etymological Dictionary of the English Language*, Oxford, Clarendon, 1909 (?), rev. and enlgd. 3rd ed.; 1st ed. 1879-1882.

SOOTHILL, WILLIAM EDWARD and LEWIS HODOUS, *A Dictionary of Chinese Buddhist Terms with Sanskrit and English Equivalents and a Sanskrit-Pali Index*, London, Kegan Paul, Trench, Trubner Co., 1937.

SUZUKI DAISETZU, ed., AKIZUKI RYŪMIN, tr., *Kōtei kokuyaku Chōshū zenji goroku [Recorded Sayings of the Zen Master Zhaozhou, Collated and Translated into Japanese]*, Tokyo, Shunjūsha, 1964; rpt. of Kamakura, Matsumatsuoka bunko, 1962.

THOMSEN, MARIE-LOUISE, *The Sumerian Language: An Introduction to Its History and Grammatical Structure*, («Copenhagen Studies in Assyriology», 10), Copenhagen, Akademisk Forlag, 1984.

TIAN ZONGYAO, comp., *Zhongguo gudian xiaoshuo yongyu cidian [Dictionary of Expressions in Chinese Classical Fiction]*, Taipei, Lianjing chuban shiye gongsi, 1985.

TIEN, T. Y., comp., *A Dictionary of Colloquial Terms and Expressions in Chinese Vernacular Fiction*, Taipei, Shin Wen Feng, 1984; rev. and enlgd. from the original ed. of, 1982.

TRUBACHEV, O. N., editor-in-chief., *Etimologicheskii Slovar' Slavyanskikh Yaz'ikov: Praslavyanskii Leksicheskii Fond*, vols. 15-17, Moscow, Nauka, 1988-1990.

VASMER, MAX. See FASMER, MAX.

WANG GUIYUAN and YE GUIGANG, comp., *Shi ci qu xiaoshuo yuci dadian [Large Dictionary of Vernacular Terms from Poetry, Lyrics, Cantos, and Fiction]*, Peking, Qunyan chubanshe, 1993.

WANG JITING, ZHANG SHAOCHENG, and WANG SURONG, eds., *Changjian wenyan shumianyu [Frequent Written Expressions in Literary Sinitic]*, Chongqing, Chongqing chubanshe, 1990.

WATSON, BURTON, ed. and tr., *The Columbia Book of Chinese Poetry: From Early Times to the Thirteenth Century*, New York, Columbia University Press, 1984.

WIEGER, L., *Chinese Characters: Their Origin, Etymology [sic], History, Classification and Signification. A Thorough Study from Chinese Documents*, (translated from the French by L. Davrout), New York, Paragon and Dover, 1965. 2nd ed., enlgd. and rev. according to the 4th French ed. published in 1927; 1st ed. 1915.

WU, HONG, *What is Bianxiang? – On the Relationship between Dunhuang Art and Dunhuang Literature*, «Harvard Journal of Asiatic Studies», 52, 1 (June, 1992), pp. 111-192.

WU SHIXUN and WANG DONGMING, chief eds., *Song Yuan Ming Qing baibu xiaoshuo yuci da cidian [Dictionary of Vernacular Expressions in More than*

a Hundred Novels of the Song, Yuan, Ming, and Qing Periods], Xi'an, Shaanxi Renmin Jiaoyu chubanshe, 1992.

YUAN BIN, chief ed., *Chanzong cidian [Dictionary of the Zen Sect]*, Wuhan, Hubei renmin chubanshe, 1994.

ZANG-MIANYU YUYIN HE CIHUI BIANXIE ZU *[Tibeto-Burman Phonology and Lexicon* Editorial Group] (Chinese Academy of Social Sciences, Nationalities Research Institute, Linguistics Section), ed., *Zang-Mianyu yuyin he cihui [Tibeto-Burman Phonology and Lexicon]*, n.p., Zhongguo shehui kexue chubanshe, 1991.

Zengaku daijiten. See KOMAZAWA DAIGAKU.

ZHANG XUAN, *Zhongwen changyong san qian zi xing yi shi (The Etymologies [sic] of 3000 Chinese Characters in Common Usage)*, Hong Kong, Hong Kong University Press, 1968.

Zhongwen da cidian (The Encyclopedic Dictionary of The Chinese Language). Editor in Chief, CHANG CHI-YUN (ZHANG QIYUN). 10 vols., Taipei, Zhonghua xueshuyuan and Huagang chubanbu, 1973.

ZHU QIFENG, comp., *Citong [Interchangeable Orthographies]*, 2 vols., Shanghai, Kaiming shudian, 1934.

Zoku zōkyō. See Nakano Tatsuei.

CHINESE CHARACTERS

A 立

B 仐

C　　(1) *lasa* 拉颯　　(2) *laza*[b] 擸擨　　(3) *laza*[c] 擸摣

　　　　(4) *esa* 搕摨　　(5) *laji/lese* 垃圾　(6) *laza*[a] 拉雜

an yu li fu 暗與理符

aweng 阿翁

baojian 寶劍

Baozhi 寶誌/寶志/保志/保誌

benmo 本末

bianjia 變家

bianwen 變文

bianxiang 變相

cuiwei 崔巍

daolu 道路

daziji 打字機

dengzhengjue 等正覺

dimsun / dianxin 點心

Du Fu 杜甫

Dunhuang 敦煌

duruo 讀若

e 搕

Ennin 圓仁

erye 二業

esa 搕摨

fangyan 方言

fanqie 反切

Fozu lidai tongzai 佛祖歷代通載

gebie 各別

Guangya shuzheng 廣雅疏證

Guangyun 廣韻

Guifeng chanshi 圭峰禪師

guo 果

Guo Maoqian 郭茂倩

guoluo[a] 蜾蠃

guoluo[b] 果蠃

Guoluo 猓玀

Han shu 漢書

Han Yu 翰愈

hudie 蝴蝶

huochetou 火車頭

ji 圾
ji[a] 集
ji[b] 霽
jiang 江
jiejue 孑 孓
Jin shu 晉書
jingji 經濟
jiu 九
jiuzheng 糾正
jixin 鷄心
kaishu 楷書
Kangxi 康熙
ke 堁
keshui[a] 渴睡
keshui[b] 瞌 睡
keshui[c] 瞌 睡
keyang 渴仰
kulong 窟窿
Kunlun 崑崙
la 拉
la dahua 拉大畫
la dapian 拉大篇
la dapian(r) 拉大片儿
la xian 拉閑
la xianpianr 拉閑篇儿
la yangpian 拉洋片
la za tan 拉雜談
laguar 拉呱 / 瓜儿
lahua 拉話
laji/lese 垃圾
lajiachang 拉家常
lalazaza 拉拉雜雜
lalilata 邋裹邋邊
laluo 拉邏
Lao She 老舍
lapdzap 拉雜
lapsap 垃圾
lasa 拉颯
lata 拉邊
lata[a] 邋遢
lata[b] 剌塔
latan 拉談
laza[a] 拉雜

laza[b] 擨攡
laza[c] 擒摛
laza bian 拉雜變
laza cui shao zhi 拉雜/襍摧燒之
li[a] 力
li[b] 立
lianmianzi 聯綿字
liluo 利落
lisou 利 颼/嗖
lisuo 利/俐 索
liu 六
Liu Xie 劉勰
luo[a] 倮
luo[b] 裸
luo[c] 臝
Mawangdui 馬王堆
Mengliang lu 夢梁錄
mu 木
nahar 那口儿
Nianchang 念常
Nü lunyu 女論語
pangdundun 胖墩墩
pusa 菩薩
Qieyun 切韻
qilin 麒麟
qiuyin 蚯蚓
quqian 驅遣
renbie 人別
Rulin waishi 儒林外史
sajia 颯/颰 攃
sasao huichen, cuochu laza 洒掃灰塵, 撮除擒摛
sata[a] 颯沓
sata[b] 駁沓
sata[c] 駁迊
sata[d] 駁踏
sata[e] 駁駘
semu 色目
shabulengdengde 傻不楞登的
shengwuxue 生物學
Shi ji 史記
Shi jing 詩經
Shouwen 守溫
Shu laza bian 書拉雜變

shuang 雙
Shuowen jiezi 說文解字
Sima Xiangru 司馬相如
Su Shi 蘇軾
suisu 隨俗
tangruo 倘若
Tongyuan zidian 同源字典
tu 土
tujue 突撅
Wang Anshi 王安石
Wang Li 王力
Wang Niansun 王念孫
Wang Xianzhi 王獻之
weiyi 逶迤
Wen xuan 文選
Wudeng huiyuan 五燈會元
Wu Jingzi 吳敬梓
"Wuxing zhi" 無行志
Wu yan 吳諺
Wu Zimu 吳自牧
xi 襲
Xichong 西充
Xiyou ji 西遊記
Xu Shen 許慎
xunwei 遜位
yan/shuo/jiangbian 演/說/講變
yan/shuo/jiangbiannü/nan/zhe/jia 演/說/講變女/男/家/者
yi 衣
"You suo si" 有所思
Yuanjue jing da shuchao 圓覺經大疏鈔
Yuefu shi ji 樂府詩集
za[a] 雜
za[b] 襍
za[c] 雜
za[d] 杂
za[e] 桨
zata 雜/襍沓/遝
zengmen 曾門
zengmen shen 曾門身
zengmen wen 曾門文
zengzu 曾祖
Zhang Sengyou 張僧繇
Zhaozhou 趙州

zhizhu 蜘蛛
Zhou Yukai 周裕鍇
Zhu Qingzhi 朱慶之
zhuanbian 轉變
zhuanbiannü/nan/jia/zhe 轉變女 / 男 / 家 / 者
zhuanzuo 撰作
zhui 隹
zimu 字母
Zongmi 宗密
zufu 祖父
zuo 作
zuo "Daren fu" 作 大人 賦