Zeitschrift: Asiatische Studien : Zeitschrift der Schweizerischen Asiengesellschaft = Études asiatiques : revue de la Société Suisse-Asie

Herausgeber: Schweizerische Asiengesellschaft

Band: 61 (2007)

Heft: 3

Artikel: "Where have all the prefixes gone?" : iambic prefixes and semanticization in old Chinese

Autor: Boltz, William G.

DOI: https://doi.org/10.5169/seals-147756

"WHERE HAVE ALL THE PREFIXES GONE?": IAMBIC PREFIXES AND SEMANTICIZATION IN OLD CHINESE

William G. Boltz, University of Washington, Seattle

Abstract

Initial consonant clusters in Old Chinese words could have been constituted either of a prefix with a morphological function plus a consonantal root initial or simply of two consonants not morphologically analyzable. While such words can in either case be characterized as **C**$_1$**C**$_2$**v**[**C**], they seem often to have had a sesquisyllabic (**C**$_1$$^{\text{Iç}}$– **C**$_2$**v**[**C**]) or fully bisyllabic (**C**$_1$**ə** – **C**$_2$**v**[**C**]) structure, typically with an iambic stress pattern as would befit a word root in the second syllable and a prefix in the first. The first syllables of such forms are therefore regularly referred to as 'iambic prefixes' whether they can be specified as having a morphological role or not. How often such sesquisyllabic or bisyllabic words were the result of the *dimidiation* of a monosyllable, *i.e.*, were devolved phonetically by some kind of epenthetic process from monosyllables with "true clusters" (**C**$_1$**C**$_2$**v**[**C**]) is not clear. What is clear is that whatever morphological functions the **C**$_1$- prefixes may have had were largely obsolescent by the late Warring States period, and as a consequence direct phonetic reflexes of the prefixes are generally lost in later stages of the language. The thesis set out here is that in many cases the **C**$_1$**ə** – "prefix syllables" were *semanticized*, *i.e.*, came to be interpreted as separate words, independent of the root word represented by the second syllable. The sesquisyllabic and bisyllabic words seem to have undergone a shift to a trochaic stress pattern as a result of which their first syllables, having thus become phonetically fortified, were prone to being re-analyzed lexically as independent, monosyllabic words. This resulted in a complete lexical separation of the first syllable from the second. Typically, the semanticization was based on a fortuitous phonetic similarity of the particular **C**$_1$**ə** – part in question to a pre-existing word in the language, making it essentially a kind of "folk-etymology." Alternatively, many of the newly perceived words were simply adopted into the language as neologisms. Numerous common two-word phrases or binomial expressions, sometimes riming, in the standard classical language when analyzed according to this "semanticization hypothesis" can be seen as preserving vestiges of what were once prefix syllables in Old Chinese sesquisyllabic and bisyllabic forms.

In his book *The Roots of Old Chinese* Laurent Sagart distinguishes iambic from fused prefixes in Old Chinese (hereafter OC), the former having "had some kind of vowel which separated them from the root" and the latter having either a

"reduced vowel or no vowel at all" (Sagart 1999: 17–18). Sagart's view, as expressed in this book, is that OC prefixes, like affixes of all kinds, always played a morphological role, although what the precise meaning of the morphological function was is often not clear. As a consequence of what he calls "a cascade of changes" that affected the language between the Old Chinese and the Middle Chinese stages, what had once been a pattern of productive morphological processes in the form of widespread affixation became increasingly obsolescent and moribund. One specific result of this extended period of what we might call the demorphologicalization of Old Chinese was that the loosely attached, iambic, prefixes were lost (*op. cit.*: 13). They were, Sagart says, "wiped out almost totally, except in the colloquial layer of certain peripheral dialects," and he goes on to say that this "loss of iambic prefixes was a long and gradual lexical process, and not a sound change" (*op. cit.*: 18–19).

One of the mechanisms operating within this "cascade of changes" that led to the "wiping out" of prefixes seems to have been, at least in the literary language, and probably in the colloquial as well, *semanticization*, that is to say, a process wherein the original first syllable that incorporated the prefix was re-analyzed lexically and identified as a separable word distinct from the lexical identity of the root word. The iambic stress pattern of a bisyllabic word, where the first syllable was the original prefix plus epenthetic vowel, became at some point trochaic, as Sagart suggests (*op. cit.*: 18), putting primary stress on the prefix syllable as opposed to the word-root syllable. This first, now stressed syllable, which we can call the *protom*, was often, perhaps typically, apotomically re-analyzed (='rationalized', 'semanticized', 'folk-etymologized') and identified with an independently existing word that was phonetically close and semantically congruent to the root syllable, which we will call the *gonotom*. In some cases it seems that, alternatively, re-analysis of the prefix syllable gave rise to a neologism with a meaning related to but in some way distinct from the word-root syllable. For those prefixes that were in origin morphologically viable, this semanticization process would have been a natural consequence as the morphological process itself ceased to be productive.[1]

For example, the form given by Sagart as ***^{b}Cə-lat-s** > **ljejH** > *lih* 厲 (var. 礪) 'whet' (Sagart 1999: 127), can be re-reconstructed as ***^{m}bə-lát-s** / ***mə-lát-s**

1 Many readers will likely recognize that the thesis set out here arises directly from the proposal made seventy years ago by the late P. A. Boodberg that he called *dimidiation*, by which he meant generally the bisyllabification of Old Chinese monosyllables with complex initial clusters. See Boodberg 1937: 353–60.

on the basis of word family and *shye sheng* evidence, and must have become, with a shift to a trochaic stress pattern, ***mmə́-lat-s**, the first syllable of which was then identified with the independently existing word ***mmaj** > *mo* 摩 (var. 磨) 'grind', thus becoming semanticized as the bisyllabic "compound" ***mmaj-lat-s** > *mo lih*, written 摩厲, and understood either pleonastically as 'grinding and sharpening' or perissologically as 'grinding on a grindstone'.[2] In either case the protom, that is, the originally meaningless (or morphologically atrophied) first syllable **^{m}bə-** / **mə-**, became invested with a lexical identity as the separate and independent word ***mmaj** > *mo*. The word ***mmaj** > *mo* in the phrase ***mmaj-lat-s** > *mo lih* is not now generally recognized as having an origin other than as an independent lexical item separate from and roughly synonymous with the word *lih* 厲 'whet, grind'; certainly it is not typically recognized as preserving or reflecting in any way the original bilabial prefix of ***^{m}bə-lat-s** / ***mə-lat-s**, yet in this "compound" expression ***mmaj-lat-s** > *mo lih* 摩厲 that is exactly what, diachronically speaking, it appears to be doing.

The expression ***mmaj-lat-s** > *mo lih* 摩厲 is well attested in both pre-Hann and Hann texts, showing that the prefix, if it indeed ever was morphologically productive, had become substantially attenuated by Warring States time and that semanticization had already taken place.[3] Note, *e.g.*:

1 *Tzuoo juann*, Jau 12: 子革曰摩厲以須王出. 吾刃將斬矣.
Tzyy Ger said: "I have been sharpening [my weapon] on the whetstone, waiting in case the king comes out. With my blade I would then cut him down."

2 For convenience I will cite OC and Middle Chinese (MC) forms from Sagart 1999 as they appear there, in spite of subsequent modifications that have been made to that reconstruction. For words or characters not in Sagart 1999, I will posit OC forms based on Behr and Gassmann 2005, but unless a particular word is actually registered with its OC form in the "Systematisches Wörterbuch" there, it should not be assumed that Behr and Gassmann would subscribe to the reconstructions that I propose. In any event, many features of OC reconstructions are uncertain enough to allow only approximate and sometimes alternative proposals, *e.g.*, the phonetic nature of the ***C-** in Sagart's ***^{b}Cə-lat-s** > **ljejH** > *lih* 'whet'. I have posited a bilabial for this consonant, on the basis of the likely phonophoric role of 萬 *wann* < ***man-s** in the written form of the word, *viz.*, 厲 / 礪. But that bilabial might have been a simple **m-** or it might have been instead a prenasalised oral stop, **^{m}b-** (which we could also write as **Nb-**) or it might have been still some other segmental feature with a predominantly bilabial quality; the present understanding of OC phonology does not, as far as I can see, permit a more precise specification than this.

3 The graphic variants already noted for 摩 and 厲 will be disregarded in citing texts.

2 *Gwoyeu*, "Yuehyeu" (shanq): 其達士, 絜其居, 美其服, 飽其食, 而摩厲之于義.
"As for their executive officers, give them pristine living accommodations, give them fine garments, provide them plenty of food, and thus grind and wear them down in regard to their sense of propriety."

3a *Shang shu*, "Fey shyh": 礪乃鋒刃.
"Grind, then, the tip and blade."
3b matching "Koong juann" line: 摩礪鋒刃.
"Grind and sharpen the tip and blade."

The *Shang shu* line from the "Fey shyh" chapter, which is a *jin wen* 今文 text, while difficult to date very precisely, is likely to be relatively early in the history of transmitted classical literature. It uses the single word 礪. The Hann period Koong An-gwo commentary, by contrast, uses the bisyllabified form of the same verb, 摩礪, now understood as a two-word phrase.[4]

4 *Luenn herng*, "Shuay shinq": 消鍊五石, 鑄以為器, 磨礪生光.
"Having reduced and refined five minerals, out of these is cast a vessel; polished and ground to a fresh gleam."

A part of the circumstantial evidence for what I might call this "semanticization hypothesis" comes from the fact that the two separate words *mo* and *lih* when occurring together with each other always appear as *mo lih*, never in the reverse order as #*lih mo*. Each occurs by itself, of course, numerous times in the classical literature, but when the two occur together as a so-called "compound" the order is invariably *mo lih*. This suggests that the order of the two relative to each other is not random or free, but that there was some constraint, perhaps prosodic, on the coordinate compound form. My suggestion is that this constraint arises from the two-word compound's origin as the semanticization of a trochaically stressed single OC bisyllabic word ***^{m}bə-lat-s** / **mə-lat-s**.

Following up on an observation made by Wolfgang Behr (Behr 1994), Sagart recognizes a somewhat marginal form of this phenomenon when he discusses instances of the common verbal negative *buh* 不 occurring before a word in classical text passages where there can be no sense of negation (*op. cit.*: 88–

4 Notwithstanding the uncertain relation of the "Koong juann" to Koong An-gwo himself, it is all the same undoubtedly a Hann period text reflecting Hann period language.

9). The word for 'wildcat', for example, normally 貍 *lii* < ***mə-rɨ̆***, is attested bisyllabically as *bu-lai* < ***pə-rrɨ̆*** written 不來 in the *Jyi jiee* commentary to the occurrence of the word 貍 *lii* in *Shyy jih* 28. Ode 179 of the *Shy jing* contains two examples of this prefix: *twu yuh buh jing, dah paur buh yng* 徒御不驚, 大庖不盈 which Sagart translates as "if the footmen and charioteers are attentive, the kitchen will be filled." As he points out, a negative meaning for these lines, such as "if the footmen and charioteers are not attentive, the kitchen will not be filled," although easily construed on the basis of a straightforward reading, is explicitly rejected by the Mau commentary for *buh jing* and *buh yng*. Instead, the commentary says specifically 不驚驚也 "*buh jing* means *jing* 'attentive'" and 不盈盈也 "*buh yng* means *yng* 'filled'." The only plausible way to understand this, as Behr and Sagart suggest in their respective discussions, is to recognize that the *buh* must be a protomic syllabic vestige of the original prefix ***p-***.[5]

The *buh* 不 in all of the preceding cases has not been semanticized or lexicalized in any conventional way, but merely "explained away" by the commentators who apparently knew the traditionally correct meaning of the lines, but did not know how to explain why this was the meaning. As a consequence the passages remain puzzling when read as everyday classical Chinese. Once it is recognized that the *buh* has nothing to do with the negative particle and is instead a vestige of an original bilabial iambic prefix on the next word, how the passage can be understood to mean what it is traditionally known to mean becomes clear. The use of *buh* in this fashion is a distinctly uncommon way of preserving some trace of the original prefix. The predominant pattern by far was instead for the first syllable to have been semanticized such that it took on life as a lexically independent, separable form. The result of this development left no grammatically anomalous residue such as the "non-negative *buh*" and precisely for that reason does not attract any special attention or demand any *ad hoc* exegesis in order to explain its meaning. Note, for example, this line from the *Juang tzyy*:

5 *Juang tzyy*, "Shiau-yau you": 摶扶搖而上.
"turning and spiraling upwards" (descriptive of the flight of the *perng* 鵬 bird.)

5 What meaning, if any, this prefix contributes to the sense of the line or of the individual words *jing* and *yng* is not clear.

The two-character expression 扶搖 *fwu-yau* < ***ba-lew** is glossed in the *Eel yaa*, "shyh tian" section as 焱 (var. 飆) 也 *biau* < ***pew** 'whirlwind'; *cf.* also *piau* < ***p^{h}ew** 飄 glossed in the *Shuo wen* as 回風也 'whirlwind' (*SWGL* 6056).[6] It is hard to imagine that all of these lexical and graphic forms do not reflect the same word, in some cases with the so-called "fusing prefix," in others with an iambic prefix. It is the latter that gives rise to the bisyllabic form which in turn comes to be semanticized as 扶搖 *fwu-yau*, "literally" understood (I suppose) as 'lifted on a whirlwind', a forced meaning at best. In fact I suggest that this is again nothing more than a bisyllabic representation in origin of the iambic form ***bə-lew**, which had become trochaic, resulting in strong syllabicity of the protom, and subsequent semanticization as 扶 *fwu* < ***ba** 'support, lift'.

In his introductory discussion laying out his view of the typological structure of Old Chinese Sagart says that a given word can occur in any of three different forms: (a) unprefixed monosyllabic form, (b) iambic form, *i.e.*, with loosely attached prefix, and (c) a form with fused prefix. (*op. cit.*: 14–15) In this case of 扶搖 *fwu-yau* < ***ba-lew** I would suggest that ***lew** 搖 is the unprefixed monosyllabic form, ***bə-lew** is the iambic form, and ***p^{h}ew** 飄 and ***pew** 焱 / 飆 are fused forms.[7] In addition to the bisyllabic 扶搖 *fwu-yau* the iambic form ***ba-lew** seems to have had an alternative development as a riming binome 漂搖 *piau-yau* < ***p^{h}ew-lew**, which is found in the transmitted text of the *Shy jing*, ode 155, meaning something like 'tossed about by a whirlwind', 'wind-tossed' (Baxter 1992: 532–33). In either case, the semanticized phrase 扶搖 *fwu-yau* or the riming binome 漂搖 *piau-yau*, the salient feature of the development is semanticization, that is to say, a two-character form arises in which each character is understood as standing for a separate, meaningful lexical item. The original prefix has, in other words, been transformed into a distinct word independent of the root word. The riming binome "option" seems to have been a very common development in the late Warring States and Hann periods and is especially apparent in the Hann *fuh* 賦 literary genre. It likely had a stylistic or literary dimension to it that was associated in particular with its riming character, and to be sure the proliferation of such binomes in Hann *fuh* and similar texts is likely

6 Behr 1994 also includes this example and these data.

7 A variant unprefixed form of the word may be reflected in 飂 *liou* < ***riw** (also *liow* < ***riw-s** and *liau* < ***rew**), glossed in the *Shuo wen* as 高風也 'a high wind'. (*SWGL* 6056) The 翏 *liow* / *liaw shye sheng* series includes characters with Middle Chinese initial *m-*, *e.g.*, 謬 and 繆, both *miow* (or *miaw*) < ***mriw-s**, which may be a reflection of the same bilabial element as in ***ba-lew**.

due in many cases to analogic formation and not to instances of genuine semanticization of an earlier iambic prefix.[8]

Prefix **p-** is not the only prefix to undergo this kind of semanticization. Note the following:

6 *Tzuoo juann*, Cherng 11: 吾不能死亡.
"I am unable to suffer death (on your behalf)."

7 *Mencius* 1A.7: 樂歲終身飽. 凶年免於死亡.
"In favorable years, always to eat to one's fill; in times of bad harvests, to escape from suffering demise."

In both of these passages the two-word expression *syy-wang* 死亡 clearly means 'to suffer death'; to render it as 'die and perish' or some similarly redundant way as if the two words *syy* and *wang* were genuinely independent and coordinate does not conform very well to the naturalness of the language. I suspect that in origin the "compound" phrase *syy-wang* is a reflection of an original iambic ***sə-mang** for *wang* 亡 'to perish', which, as is widely recognized, seems to have had an **s-** prefixed form. Sagart discusses this word in *Roots* (p. 65) in connection with its obvious cognate *sang* < ***smang** 喪 'suffer the loss of', in which *wang* 亡 'to perish' seems to be both graphically the phonophoric component and etymologically the root.[9] The three expected forms in this case are: ***sang** (fused), ***sə-mang** (iambic), and ***mang** (unprefixed).

8 See Knechtges 1998 for a thorough discussion of the nature of these kinds of binomes and the challenge that they present to a translator.

9 I discussed this same pair in pretty much the same way in a short paper nearly forty years ago, never questioning Sheu Shenn's analysis in the *Shuo wen* indicating that 亡 ***mang** was phonophoric in 喪 ***smang**. (Boltz 1968). It is now fairly widely recognized that the pre-Hann epigraphic data do not unambiguously support the *Shuo wen* analysis for the early history of the character 喪. In particular there is no evidence of the graph 亡 ***mang** in the Shang inscription form or in most bronze inscription forms of 喪. This of course means that 亡 ***mang** could not have served as a phonophoric in 喪 ***smang** in those early cases. It does not mean, it is important to note, that we are compelled to conclude that the word *sang* did not have an ***sm-** initial cluster or that it was not cognate with ***mang**. There are enough late Warring States period forms of the character attested where the 亡 component is present to suggest that 亡 was introduced secondarily as a phonophoric component at some relatively late point in the character's history. Thus we do not have to throw Sheu Shenn's analysis out as useless; the very fact that he analyzed 喪 *sang* as having 亡 *wang* as a phonophoric suggests that he was aware of an etymological and graphic link between these two words,

The same ***s-** / ***sə-m-** / ***m-** pattern can be seen in the word 滅 *mieh* < ***met** 'extinguish, exterminate, wipe out'. The unprefixed form is ***met**, the iambic form was ***sə-met**, and the fused form was ***smet** > **swit** > *shiu* 戌. The fused form has become obsolete in the meaning 'extinguish' and exists generally only as the eleventh of the twelve terrestrial rames (the *dih jy* 地支). But it is all the same registered as having had the meaning 'extinguish' in Hann texts, *e.g.*, by implication in

8 *Shyy jih* 25 ("Liuh shu" 3): 戌者萬物盡滅.
"The *shiu* period [of the ramal cycle] is when all creatures are exhaustively extinguished."

More explicitly, the meaning 'extinguish' is expressed paronomastically, in

9 *Bor huu tong*, "Wuu shyng 1": 戌滅也. (***smet** > **swit** = *[**sə-**]**met** > **met**)
"*shiu* is 'to extinguish'."

What is of still more interest in connection with the disappearance of prefixes is how the iambic form seems to have come to be reflected in a great variety of semanticized forms in texts. Note the following:

10 *How Hann shu* 39: 降及戰國奢僭益熾削滅禮籍.
"Reaching down to the time of the Warring States, wastefulness and profligacy increased and flourished. They abolished proper forms and records."

11 *How Hann shu* 87B: 冰霜見日必至消滅.
"When ice and frost are exposed to the sun, they will inevitably melt."

---

and was looking for a way to rationalize their respective scriptions accordingly. In fact the *Shuo wen* entry says 从哭从亡會意亡亦聲, one of the infrequent cases where Sheu Shenn explicitly identifies a character as a *huey yih* type. This analysis, with its 从亡 […] 亡亦聲 wording, generally is taken to mean that he was attributing both a semantic and a phonetic role to the 亡 component, and this in turn supports the speculation that 喪 *sang* and 亡 *wang* are cognate with each other. We can still say that 亡 ***mang** is phonophoric in 喪 ***smang**, we simply have to acknowledge that graphically it has arisen fairly late as a secondary phonophoric element.

12 *Hann shu* 60: 咎異何足消滅.
"As for anomalies and aberrations, what is sufficient for them to be eliminated?"

13 *Hann shu* 36: 百異消滅而眾祥並至.
"When the 'hundred' aberrations are eliminated, a cascade of good omens will come forth."

14 *Hann shu* 36: 凶災銷滅.
"Misfortune and disaster have been wiped out."

15 *How Hann shu* 25: 則凶妖銷滅害除福湊矣.
"[…] then misfortune and perversions would be wiped out, harm would be eliminated and good fortune would come in abundance."

16 *How Hann shu* 95: 先零東羌歷載為患. 潁前陳狀欲必掃滅.
"The Eastern Chiang have been a nuisance at Shian-ling for years on end. Jeong formerly presented a scheme for their inevitable extermination.

17 *How Hann shu* 99: 左右掃滅我曹.
"Left and right, they will wipe us out."

18 *Tzuoo juann*, Cherng 2: 余姑翦滅此而朝食.
"Let me deal with exterminating these, then I shall take my breakfast."

19 *How Hann shu* 59: 永因數為諫陳興復漢室翦滅纂逆之策.
"Yeong on this basis several times admonished us, laying out a plan for restoring the House of Hann and exterminating the usurpers and rebels."

20 *Shyyjih* 8: 諸所過無不殘蔑.
"No place by which he passed failed to be wiped out." (Said in reference to the depredations of Shianq Yeu 項羽 during the civil war.)

21 *Hwainan-tzyy* 1: 與高辛爭為帝. 遂潛于淵. 宗族殘蔑.
"(Gonq Gong) contended with Gau Shin to become ruler. He subsequently was consumed in the vortex, and his clan and lineage were exterminated."

22 *Tzuoo juann*, Jau 4: 執齊慶封而盡殪其族.
"Having seized Chinq-feng of Chyi, they exterminated his lineage."

23 *Tzuoo juann*, Jau 27: 盡殪卻氏之族黨.
"(The Chief Minister) extinguished all branches of the Chiueh lineage."

All of the cases listed here, including the binomial phrase *jinn mieh* < ***tsins-met** seen in examples (8), (22) and (23), are, I suspect, cases of semanticization of the same original iambic form of *mieh* < ***sə-met**. The binomial forms for 'extinguish' in examples ten through twenty-three can be summarized as follows:

削滅	*shiue-mieh*	< ***sewk-met**
消滅	*shiau-mieh*	< ***sew-met**
銷滅	*shiau-mieh*	< ***sew-met**
掃滅	*sao-mieh*	< ***ssuw-met**
翦滅	*jean-mieh*	< ***tsanʔ-met**
殘殪	*tsarn-mieh*	< ***ddzan-met**
盡殪	*jinn-mieh*	< ***tsins-met**[10]

Each one can be translated in a way that gives full lexical credit to the first member, *e.g.*, 'scrape away and extinguish', 'wipe out and extinguish', etc. This is a legitimate way of reading these binomes when viewed synchronically from the perspective of Classical Chinese literary style, with its strong predisposition toward parison and metrical regularity. All the same, there is I think sometimes a sense in which this seems a bit like translation overkill. It is natural enough, of course, to understand the phrase *jinn-mieh* 盡殪 as *adjunct-head*, 'completely exterminate', rather than 'wipe out and exterminate' ('wipe out' would in fact be an etymologically correct rendering of the verb *jinn* 盡), and there is little doubt that that is how it is conventionally taken. But it seems unlikely that the phrase was really meant to draw a contrast with an *incomplete* extermination, and in origin, the 'completely' part is likely to be the result of a semanticization of the ***sə-** protom, just as are the other first members of the rest of these two-word phrases.

10 The characters 滅 and 殪 are regularly seen in early texts as variants of each other, standing for the same word *mieh* < ***met** 'extinguish', though *sensu stricto* they are distinguished from each other by the Middle Chinese rime tables in that 滅 is given as a third division word and 殪 as fourth.

To recapitulate, my suggestion is that each of these two-part forms is in origin a semanticized reflection of the original iambic form ***sə-met** where the stress became trochaic and the first, now stressed, syllable was lexicalized as a consequence of losing its prefixal identity and what was perhaps a morphological function. Notice that the hypothesis that I am proposing applies equally well irrespective of whether we can identify the prefix as morphemic or not. If OC had monosyllables with "true" initial consonant clusters, not morphemically analyzable, the hypothesis would still apply. We could describe the process (somewhat artificially) as consisting basically of these steps; (i) epenthesis, (ii) full bisyllabification, (iii) an iambic stress pattern yielding to trochaic, (iv) semanticization. Using the *mo lih* 摩厲 data as an example, we could suppose the following: (o) ***mlat-s** > (i) ***mᵊlat-s** > (ii) ***mə-lat-s** > (iii) ***mməʼ-lat-s** > (iv) ***mmaj-lat-s**, which now at stage (iv) is OC for *mo lih* 摩厲. I am not suggesting that these steps actually must have proceeded in this way one-by-one, each on an equal footing with the others, but only that the whole semanticizing and lexicalizing process can be seen heuristically to consist in these points. How the process might have transpired in the spoken language we can only surmise, but in the written language only points (o) and (iv) are viable.

Notice also that in the same way that the order of the two components of the "compound" expression *mo lih* 摩厲 appears to be fixed, as we observed above, so also for all of the forms with *mieh* 蔑 as the second component. To the extent that they are coordinate compounds, the order of the two members ought in theory to be reversible, but in practice I am unaware of any cases in early classical texts of *mieh* 蔑 coming first followed by a second member from among the words listed above. This empirical observation, it seems to me, lends further support to the hypothesis that these two-word phrases reflect in origin the semanticization and lexicalization of an earlier iambic prefix. What I am calling here semanticization and lexicalization is simply a kind of folk-etymology.[11] The protom came to be identified with a semantically congruent, independently existing word with which it was phonetically similar, especially in the initial. As

11 Brinton and Traugott (2005: 18) in their recent study of lexicalization and language change distinguish synchronic from diachronic lexicalization, the former having to do with the coding of conceptual categories and the latter dealing with "adoption into the lexicon" of linguistic entities that had origins as components other than lexical items. I am concerned here only with the second of these two, the diachronic type of lexicalization, and as can be seen, I am using the term lexicalization in a very straightforward way. Brinton and Traugott tend to focus their study on lexicalization especially as it relates to grammaticalization. This aspect does not figure in the OC material analyzed here.

the phonological structure of the language changed such that initial clusters were lost throughout, and as whatever morphological processes that these clusters may have represented became obsolescent, speakers were left with few ways to understand these excrescent protomic syllables. Folk etymology was one common and natural recourse.

It may have happened in some cases that rather than being identified with a pre-existing word the first syllable was instead lexicalized as a neologism. Consider, for example, the following:

24 *Liijih* , "Shao yi" : 劍則啟櫝蓋襲之. 加夫襓與劍焉.
"(When presenting) a sword, open the cover of the wooden case and turn back its lining; lay the cloth sheathing together with the sword on it."

The curious word in this line is 夫襓 *fu-rau* < ***pa-nrew**, which looks very much like an iambic form of a **p**-prefixed word meaning 'wrapping'; here the specific sense is 'cloth sheathing for a sword'. The fused form could be 表 *beau* < ***pjew?** 'outside garment, wrapping'; the apparent phonophoric in 表 *beau* is 毛 *mau* [表 < 衷, see *SWGL* 3689, Duann juh among others] which would suggest an OC ***pjaw?**, but this may reflect a *shye sheng* structure that arose after the merger of the **ew** and **aw** distinction in the traditional 宵 *shiau* rime class. The protom 夫 *fu-* < ***pa-** comes to be written with classifier 145, 衣, thus: 衭 and is entered in the *Shuo wen* as 襲衭也, a succinct and somewhat imprecise explanation that seems to suggest a part of the lapel lining of a garment. The same character is entered in the various Middle Chinese rime dictionaries with a pronunciation corresponding to modern *fu* and a meaning 'lapel', *e.g.*, *Goang yunn*, *pyng* 10 : 衭衣前襟也. Graphic variants are given in the *Jyi yunn* with classifier 050, 巾, 'napkin', 帎 and classifier 120, 糸, 'silk', 紨. For all three characters, 衭, 帎 and 紨, the dictionary gives both the *Shuo wen* entry and the same entry seen in the *Goang yunn* and cites further the *Bor yaa*, which identifies the character as the protom of the binomial 衭襓 *fu-rau*, glossed as 劍衣 'sword sheath'. All of this suggests that the protom came at least by Eastern Hann times to be seen as an independent character standing for a distinct word in its own right, with a meaning having something to do with 'lapels' or 'linings', or 'lapel-linings', but certainly not any longer as simply a meaningless first syllable in a bisyllabic form *fu-rau*. The semantic relation between the original sense of 'cloth sheathing for a knife', described in the *Liijih* context as 'turned back' or 'opened out', and 'lapel' is likely to be the 'turning back' or 'folding out' of a covering, whether lapel or knife sheath.

Jenq Shyuan 鄭玄 (127–200) was not fooled by this. He says the following in his commentary to the *Liijih* passage given above:

25 Jenq Shyuan *juh* 注: 夫褖劍衣也 […] 夫或為煩. 皆發聲.
"*fu-rau* is a covering for a sword […] 夫 is sometimes written as 煩; in either case it expresses the pronunciation."

Clearly the explanation for the alternate scription with 煩 *farn* < ***ban** is that the protom ***pa**-, which according to the *Goang yunn* had a reading with a voiced initial **b**-, thus OC ***ba**-, acquired a final -**n** through assimilation to the initial of the second syllable *rau* < ***nrew**. Equally clearly, Jenq Shyuan was not puzzled by the use of characters to represent sounds alone in polysyllabic words, an appreciation of a kind of orthographic subtlety that seems not to have prevailed much beyond the end of the Hann in connection with writing native Chinese texts, but that became widespread in producing Chinese translations of Indic Buddhist texts.

The examples surveyed so far can be categorized into four types: (i) binomes easily understood as consisting in two independent words, *e.g.*, 死亡 *syy wang* and 消滅 *shiau-mieh*, (ii) binomes not so easily understood as consisting in two independent words, except through lexicalization of the protom as a neologism, *e.g.*, 袚褖 *fu-rau*, (iii) binomes that cannot be understood on the basis of the separate meanings of their written components at all, but only through an *ad hoc* explanation, *e.g.*, 不來 *buh lai* 'wildcat' and 不驚 *buh jing* 'attentive', (iv) riming binomes, *e.g.*, 漂搖 *piau-yau*. Categories (i) and (ii) exhibit the kind of semanticization that I think underlies such binomes in general and reflects what were once iambic prefixes. Category (iii) is, as we have said, rare, in that such binomes have not been semanticized systematically.[12] Category (iv) is, I suspect, a special case of categories (i) or (ii), depending on whether or not the protomic member of the binome existed as an independent word when the binome arose. Riming binomes of this type are called by the descriptive category name *dye yunn* 疊韻 'doubled rimes' in the Chinese lexicographical tradition, and this is usually translated into English as 'partial reduplication', meaning simply that the

12 It is relatively easy to see how the name 不來 *buh lai* for 'wildcat' could be folk-etymologized as "the wild cat that never comes out," *i.e.*, that is typically not seen in the open, but this is an *ad hoc* interpretation of the phrase itself rather than an example of semanticization of the protom alone. All the same, it still shows that what was once likely to have been an iambic prefix ***pə**- has been re-analyzed as a meaningful, independent word.

riming part of the syllable is reduplicated but the initial is not.[13] Generally, no explanation is offered as to where the initial of the protom comes from.

Sagart discusses reduplication briefly in chapter 14 (*op. cit.*: 137–38) observing generally that such *dye yunn* binomes likely had "a variety of origins." He includes here the possibility that partial reduplicatives may have arisen through a process of reduplication of prefixed words and that in the case of fusing prefixes the reduplication applied only to the word-base (what I called 'word-root' above), exclusive of the prefix. Thus, once the fusing prefix had supplanted the original word-base initial in the gonotom, the resulting binome, with the word-base initial preserved in the protom but replaced by the prefix in the gonotom, looked like a partial reduplication involving the final alone. He cites several "East Asian parallels for this kind of construction," but does not give any Chinese examples.[14] He goes on to say that "[p]refixed reduplications can be reconstructed when the evidence indicates that the initial of the first element […] is an earlier prefix that has replaced the root initial" (*loc. cit.*). As an example of this kind of *dye yunn* reduplication he gives the word 果蠃 *guoo luoo* < MC kwaX-lwaX 'kind of gourd', which he derives from a putative reduplicated form ***^{A}k-loj?-loj?**. My suspicion is that the binome 果蠃 *guoo luoo* reflects instead the semanticization as 果 *guoo* < MC kwaX 'fruit' of a protomic ***kə-** of an earlier iambic form ***kə-loj?**, perhaps from a still earlier monosyllabic form with an initial cluster, such as ***kloj?**. This analysis gives both a simpler explanation for the origin of the binome in question than that which Sagart has proposed and is at the same time consistent with the very general and widespread pattern of semanticization that we have been able to discern elsewhere.

The real problem with the word 果蠃 *guoo luoo* is knowing whether this is really a 'kind of gourd', as Sagart glosses it, or is in fact the generic word for 'fruit', for which we find the written variant 果蓏. Sagart (*op. cit.*: 107) reconstructs the simple word 果 *guoo* < MC kwaX 'fruit' as OC ***^{A}k-loj?**,

13 See Chou 1962: 97–201 for one of the most extensive, classic studies of reduplication in early literary texts, chiefly the *Shy jing*.

14 While he does not give any Chinese examples, he does mention examples from Written Burmese, Gyarong, and Paiwan to illustrate the process concretely (*loc. cit.*). Sagart is certainly correct to point out that whatever their diachronic explanations, these kinds of reduplications are typologically characteristic of a wide range of East Asian languages. For a splendidly precise and detailed synchronic analysis of such forms in Vietnamese see Emeneau 1951: 159–200. See Baxter and Sagart 1997: 64–66 for a slightly earlier, equally brief, statement of the same approach to reduplication.

inferring the cluster, he says, on the basis of (a) the binome 果蠃 *guoo luoo* < ***[A]k-loj?-loj?** 'kind of gourd' and of (b) the word 蓏 *luoo* < MC lwaX < ***[a]Cə-loj?** 'plant fruit'.[15] When the first part of the word 果蠃 *guoo luoo* is understood as reflecting the semanticization of a syllabic ***kə-** in a bisyllabified word for a 'kind of gourd', as I suggested above, rather than the partial reduplication that Sagart suggests, it does not support a hypothesis of a cluster ***k-l-** in the OC for 果 *guoo* 'fruit'; it supports instead such a reconstruction for the gonotom, *viz*., in the word that is said to mean a 'kind of gourd'. If we are willing to equate 蓏 *luoo* lexically with the generic word 果 *guoo* 'fruit', taking the two as a pair of lexical "doublets," and to understand 果蠃 *guoo luoo* 'kind of gourd' as the same word as 果蓏 *guoo luoo* 'fruit', we then would have a good basis for positing ***k-l-** in the OC word for *guoo* 'fruit'.[16] As it happens, the evidence of transmitted texts supports precisely such a conclusion. The two-character riming binome 果蓏 *guoo luoo*, perfectly homophonous with 果蠃 *guoo luoo*, is well attested in literature from the Warring States period, *e.g*.:

26 *Juang-tzyy*, "Ren jian shyh": 夫楂梨橘柚, 果蓏之屬.
"in any case of a cherry-apple, pear, orange, or pomelo,—all belonging to the fruit-or-berry producing type of tree—."

Whether the *Juang-tzyy* line was originally intended to say 'fruit producing kind of tree', taking 果蓏 as a single lexical item, or 'fruit and berry producing kind of tree', taking it as a two-word *dvandva* compound, we cannot know for sure. But we can know how the commentators explained it, and in every instance they opted for the *dvandva* explanation, striving to distinguish a 果 *guoo* from a 蓏 *luoo*. Cherng Shyuan-ing 成玄英 (*fl*. mid-seventh cen.), for example, in his *shuh* 疏 to this line says:

27 *Juang-tzyy shuh* 疏: 在樹曰果, 在地曰蓏.
"If it's on a tree, it's a 果 *guoo*; if it's on the ground, it's a 蓏 *luoo*."

The same distinction was made more than five centuries earlier in the *Shuo wen*, which said in its entry for 蓏 *luoo*:

15 He implies without any explicit discussion that the ***C-** in 蓏 *luoo* < ***[a]Cə-loj?** 'plant fruit' is to be understood phonetically as ***k-**.

16 Alternatively, we might prefer to write ***kl-**; the distinction between ***k-l-** and ***kl-** being whether the **k-** segment is seen as a separable prefix or as an inherent element of a morphologically unanalyzable cluster.

28 *SWGL* 0235: 在木曰果, 在地曰蓏. (*cf.* Lat. FRUCTUS *vs.* FRUX.)

Alternatives to the "tree *vs.* ground" distinction included the following:

29 *Leu shyh chuen chiou*, "Jonq Shiah," Gau Yow 高誘 (*ca.* 168–212) note:
有核曰果, 無核曰蓏.
"If it has a kernel, it is a 果 *guoo*; if not, it is a 蓏 *luoo*."

30 *Yih jing*, "Shuo guah", Koong Yiing-dar 孔穎達 (574–648) note:
草實為蓏.
"The fruit of a grass is called 蓏 *luoo*." (Whereas, by implication the fruit of a tree or shrub is 果 *guoo*.) (*cf.* Lat. FRUMENTUM.)

Commentators to texts throughout the mediaeval period repeatedly invoke one or another of these various distinctions in order to explain the difference between a 果 *guoo* and a 蓏 *luoo*. In all cases the distinctions, always expressed as complementary pairs, are efforts to account in semantic terms for what in origin was likely no more than the bisyllabification, *i.e.*, the *dimidiation*, of an earlier monosyllable with an initial ***kl-** (or ***k-l-**) cluster. The proposed distinctions are invariably contrived contrasts imposed on the two parts of the binomial expression, giving each part in the end an ostensible independent lexical status, however artificial. This illustrates still another type of semanticization that seems to have characterized some of the binomes arising through the process we have sketched above, and it suggests that Sagart is right to suspect an OC ***^{A}k-loj?** for 'fruit'.

Sagart also claims that the homophonous and graphically related binome for 'wasp', 蜾蠃 [*sic*, conventionally 蜾蠃] *guoo luoo* < ***^{A}k-loj?-loj?**, "allows us to posit *guo*$_3$ 蜾 ***^{A}k-loj?** > kwaX" (*ibid.*) In this case, since the protom does not show any evidence of having been semanticized and the two-syllable expression has every appearance of constituting a single, unanalyzable word, to say that the OC *word* for 'wasp' was ***^{A}k-loj?** is unobjectionable, though the binome 蜾蠃 itself does not provide any basis for presuming that the character used to write that monosyllable ever was 蜾 alone (but see *infra*, footnote 17.)[17]

17 Schuessler (2007: 269) enters the word 蜾蠃 *guoo luoo* as a 'species of small wasp'. He gives a number of possible cognate forms in Tibeto-Burman languages, none of which shows a medial *r* or *l*. On this basis he suggests that Chinese 蜾蠃 *guoo luoo* is not a dimidiation (from an earlier ***^{A}k-loj?**), but is rather "simply a reduplicative compound of the common type CV-lV or *CV-rV," and he refers to his discussion of reduplication in the

In setting out his methodological assumptions for reconstructing Middle and Old Chinese Sagart reminds us that using the *Chieh yunn* as a guide to the sound system of Middle Chinese is not the same as using the comparative method, and this traditional approach to the Middle Chinese sound system "confers an excessive weight on the literary tradition at the expense of the oral tradition" (1999: 9). Sagart suggests that the kind of bisyllabic words with their loosely attached prefixes that we have been talking about are characteristic chiefly of the colloquial language, and that this explains why we cannot, according to him, find clear traces of such forms in the *Chieh yunn*. If my suspicion about the extent of semanticization of protomic syllables originating in iambic prefixes is right, then the situation is not as bleak as Sagart implies. The data of the *Chieh yunn* and of other similar Middle Chinese riming dictionaries do preserve traces of these prefixes, as does the huge corpus of transmitted literary texts, but they do so largely by giving what were for the compilers of the dictionaries and authors of the texts simply two-word phrases, in the case of the riming dictionaries often entered as glosses to one of the two words as a single character entry.[18] That those two-word phrases have arisen as semanticized derivatives of the originally iambic forms is something that the dictionary compilers were unlikely to have recognized. This is an entirely understandable development given the nature of the Chinese writing system, where every character, no matter what its origin, inevitably carries a lexical identity. And the writing system, at bottom, can be nothing more than a reflection of the language. The possibility of writing prefixes *qua* prefixes simply does not exist as a normal practice in the Chinese script, and has not existed since at least the second century B.C. By virtue of its prestige and cultural weight the script forces a lexical identity on any attempt to

---

"Morphology and Word Derivation" chapter of his introductory material. That discussion (*op. cit.*: 24–25) identifies reduplication as a process that produces what Schuessler calls "expressives," but there is no explanation of the sense in which an insect name might be an "expressive" that would account for the word 蜾蠃 *guoo luoo* 'wasp' falling into this derivational category. Sagart's suggestion of an earlier ***^{A}k-loj?** for 蜾蠃 *guoo luoo* 'wasp' seems, all things considered, a likelier surmise.

18 The *Goang yunn*, for example, enters 蜾 as 蜾蠃蟲也, but for 蠃 it says simply 同上, referring to the preceding entry, which is 裸 *luoo* 'naked' (上聲 24, 果韻). The *Jyi yunn*, by contrast, enters 蜾 along with 蝸 and 蠃 as graphic variants for the same syllable / word *guoo*, glossed simply as 蟲名, never actually giving the binome 蜾蠃 (上聲 34, 果韻). The entry does give 蠃蠃, the binome found in the *Shuo wen* presumably writing the same word *guoo luoo* 'wasp' (*SWGL* 5978). To complicate the picture still more, the *Jyi yunn* lists the character 蜾 a second time with a reading *luoo*, unexpectedly confirming Sagart's suggestion that 蜾 *guoo* itself likely comes from ***^{A}k-loj?** (上聲 34, 果韻).

write a sound value alone, whether prefix or other, whether we like it or not. Jenq Shyuan recognized this phenomenon in the second century A.D.; we live with its consequences eighteen centuries later.

Abbreviations

OC Old Chinese, the language of the Warring States period.
SWGL *Shuowen jieetzyh guulin* 說文解字詁林 (Ding Fwubao 1928).

Bibliography

BAXTER, William H. III
1992 *A Handbook of Old Chinese Phonology*. Berlin: Mouton de Gruyter.
BAXTER, William H. III and Laurent Sagart
1997 "Word Formation in Old Chinese". In Jerome L. Packard, ed., *New Approaches to Chinese Word Formation*. Berlin: Mouton de Gruyter: 35–76.
BEHR, Wolfgang
1994 "'Largo forms' as secondary evidence for the reconstruction of Old Chinese initial consonant clusters." Handout (unpublished): 27ème Congrès International sur les Langues et la Linguistique Sino-Tibétaines, Paris.
BOLTZ, William G.
1968 "An Etymological Note on Old Chinese ***sâng**." In: *Phi Theta Papers* (Publication of the Honor Society in Oriental Languages, University of California, Berkeley), vol. 11: 65–67.
BOODBERG, P.A.
1937 "Some Proleptical Remarks on the Evolution of Archaic Chinese." In: *Harvard Journal of Asiatic Studies*, 2.3–4: 329–72.
BRINTON, Laurel J. and Elizabeth Closs Traugott
2005 *Lexicalization and Language Change*. Cambridge: Cambridge University Pr.
CHOU Fa-kao 周法高
1962 *Jonggwo guuday yeufaa: gow tsyr biann* 中國古代語法: 構詞編. (*A Historical Grammar of Ancient Chinese, part II: Morphology*.) Tair-

beei: Institute of History and Philology, Academia Sinica. Special publications No. 39.

DING Fwubao 丁福保

[1932] *Shuowen jieetzyh guulin* 說文解字詁林. Shanqhae: I shyue. Photo-rpt. Tairbeei: Shangwuh, 1970.

EMENEAU, Murray B.

1951 *Studies in Vietnamese (Annamese) Grammar*. Berkeley & Los Angeles: University of California Pr. University of California Publications in Linguistics, vol. 8.

GASSMANN, Robert H. and Wolfgang BEHR.

2005 *Antikchinesisch – Ein Lehrbuch in drei Teilen*, Teil 1. Bern (etc.): Peter Lang.

KNECHTGES, David R.

1998 "Problems of Translating Descriptive Binomes in the *fu*.". In: *Tamkang Review*, vol. XV: 329–47.

SAGART, Laurent

1999 *The Roots of Old Chinese*. Amsterdam & Philadelphia: John Benjamins.