ICCL-5 poster
June 1996
Tsing Hua University

PROSODIC STRUCTURE IN CHINESE CHARACTERS

James Myers
University of Michigan

ABSTRACT

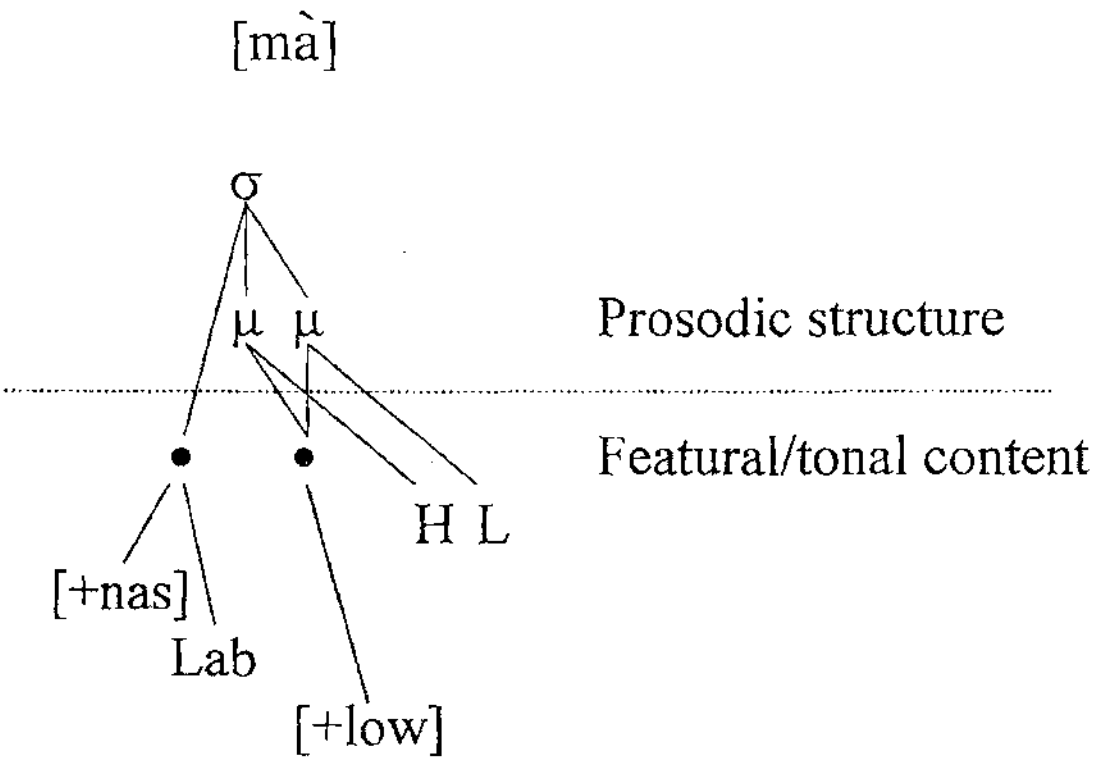
Prosodic phonology represents phonological structure as independent of phonological content. This concept has been very useful in describing a variety of phonological patterns in both spoken and natively acquired signed languages. In this poster I argue that prosodic structure is also found in the shape of Chinese characters. Specifically, a single shape template can be used to describe four otherwise unrelated generalizations about Chinese characters:

(1) the location of long strokes and enlarged components;
(2) the legal patterns for duplicated elements;
(3) the location of curved vertical strokes;
(4) tendencies in the location and form of radicals.

The existence of a single explanation for these disparate generalizations suggests that prosodic structure was a real force in the evolution of modern Chinese characters. It is also possible that this structure is used by experienced readers and writers of Chinese in the processing and mental storage of characters.

WHAT IS PROSODIC STRUCTURE?

Prosodic phonology represents phonological *structure* (e.g. moras μ and syllables σ) as independent of phonological *content* (e.g. features, tones):



The fundamental claim made by the theory of prosodic phonology is that prosody is represented by ***constituents***. For example, English uses a binary, left-headed metrical foot, which can be represented with the following template:

[x .]

("x" represents the ***head***, "." represents the ***nonhead***)

Prosodic constituents demonstrate their existence by providing *simple, uniform explanations* for a wide variety of different phenomena:

(1) <u>Prominence (stress)</u>.

In many languages, the location of prominence is predictable by the use of metrical feet. Syllables dominated by the heads of metrical feet are stressed. Thus, in English, feet are built from right to left:

[x]	[x .]	[x] [x .]	[x .][x .]
big	happy	stupendous	obligatory

[x .] [x .][x .][x .] [x .][x .][x .]
supercalifragilisticexpialidocious

(2) <u>Prosodic morphemes</u>.

In some languages, prosodic constituents can be used to build new words. For example, in Arabic, the plural affix is the right-headed foot [σ_{μ} $\sigma_{\mu\mu}$]:

(from McCarthy and Prince 1990)

<u>singular</u>	<u>plural</u>	
nafs	nufuus	"soul"
qidH	qidaaH	"arrow"

(3) Rule domains.

Phonological rules can refer directly to prosodic constituents. For example, aspiration in American English occurs only at foot boundaries, and vowels that are not linked to heads are reduced:

[x .]	[x .]
[ə t^hamɪk]	[a ɾəm]
"atomic"	"atom"

(4) Morphological domains.

Morphology can also refer directly to prosodic constituents. In English, the comparative *-er* can only suffix to words with a single foot:

[x .]	
happy + er	ACCEPTABLE
[x] [x .]	
stupendous + er	UNACCEPTABLE
[x] [x]	
abstract + er	UNACCEPTABLE

"PHONOLOGY" BEYOND SPEECH

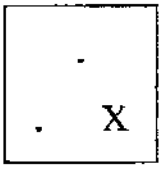
Scholars have argued that patterns similar to phonology can be found outside of spoken language. Evidence is particularly strong in ***natively acquired sign languages***. Sign languages are linguistically and neurologically very similar to spoken languages in many ways (Poizner, Bellugi and Klima 1989). It is thus not surprising that researchers have found evidence for prosody in sign language. For example, Padden and Perlmutter (1987) argue that hand movement patterns in American Sign Language are best described with prosodic constituents like moras and syllables.

Prosody has been claimed to be important in nonlinguistic domains as well. For example, Lerdahl and Jackendoff (1983) show how the ***rhythmic structure of music*** parallels the metrical structure of spoken language.

What about ***orthography***? Many scholars have independently claimed that writing systems also have an internal structure analogous to phonology (e.g. Albrow 1972, Eden 1961, Herrick 1966, McCawley 1989, Wang 1983, Watt 1979).

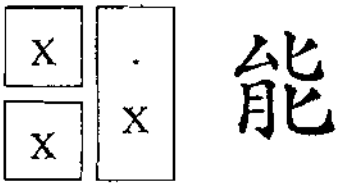
THE CHINESE CHARACTER TEMPLATE

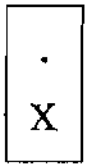
I propose that the basic prosodic constituent in Chinese characters is the following two-dimensional template:

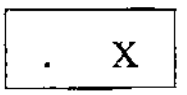


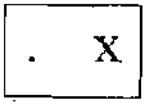
("x" represents the ***head***, "." represents the ***nonheads***)

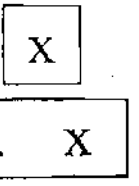
Characters are built out of subsets of this template that include a head, and out of combinations of such templates (sometimes there is unanalyzable material left over):











器

This single concept can be used to explain *four distinct patterns* in the shape of Chinese characters.

(1) PROMINENCE (SIZE)

In general, the largest of a *vertically* arranged set of strokes or components is located at the *bottom*:

三 土 未 多 官

Likewise, the largest of a *horizontally* arranged set of strokes or components is located at the *right*:

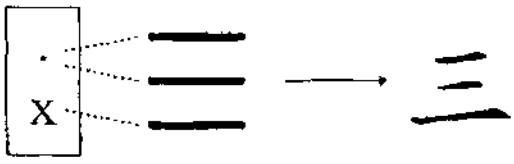
介 川 州 林 朋

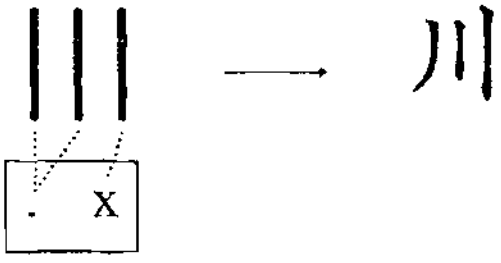
Both generalizations can hold in the same character, as in 井.

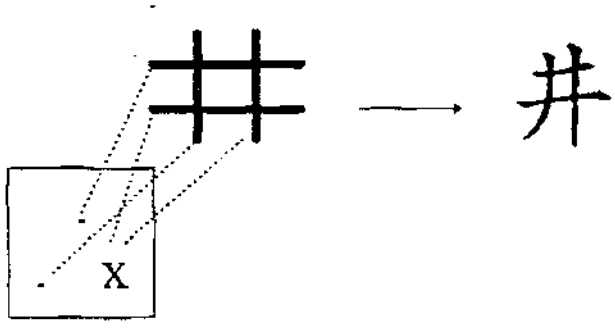
Moreover, the generalizations hold both within entire characters and also within each of the component parts:

圭

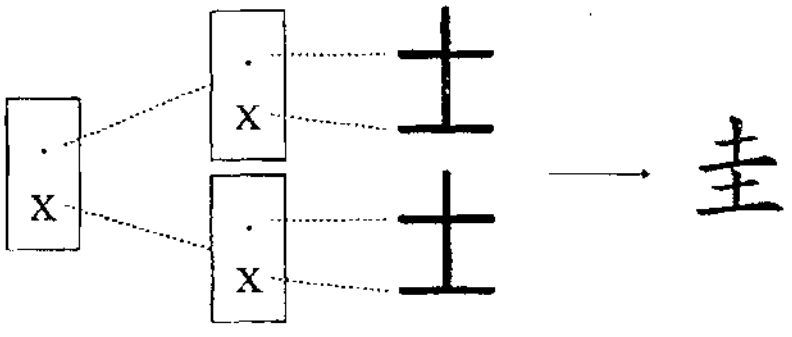
These patterns can be described by *linking schematic forms to templates* (see Wang 1983 for arguments for such schematic forms). Strokes or components that are linked to heads receive prominence.







Templates may be arranged ***hierarchically***, resulting in different degrees of prominence:

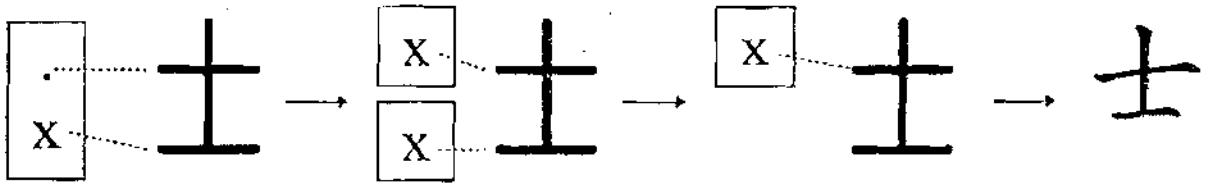


Virtually all exceptions to these generalizations are characters that are lexically marked to undergo ***Topmost Prominence***:

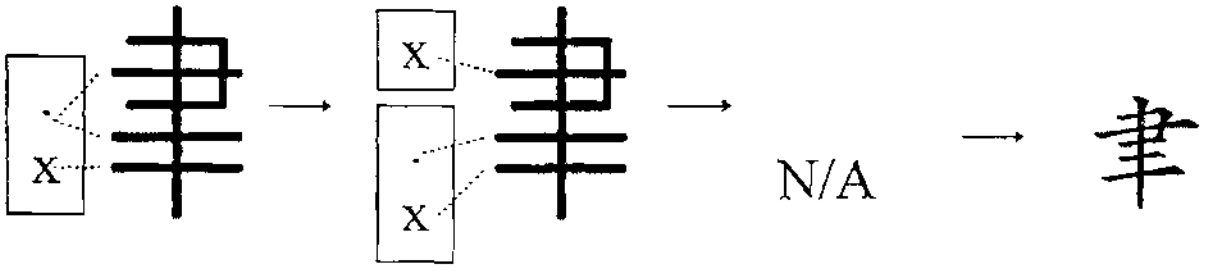
士 末 聿 畢 重 毳

These can be analyzed in a three-step process:

(a) Link to a template
(b) Topmost Prominence (if lexically marked)
(c) ***Clash***: Prominence added in (a) is removed if *directly adjacent* to prominence added in (b)



(a) (b) (c)



(a) (b) (c)

(Clash thus parallels stress clash in spoken language.)

(2) DUPLICATION TEMPLATES

Characters formed with duplicated components only appear in the *shapes that are allowed by the template*: [. .], [:], [∴].

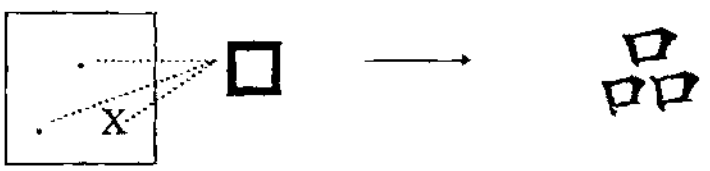
The following shapes are never found:

(Note that the duplicated elements in characters like 三 川 黑 巡 龠 靈 either consist of a single stroke or do not create independent components.)

There are some cases where duplicated components appear in a 2x2 shape, but with only one exception all are *split into two legal template shapes* (2x1 or 1x2):

爽 繼 齒 爾 綴

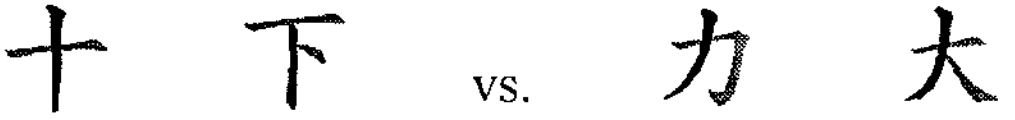
These generalizations can be described by *multiply linking a single schematic form* to template slots:



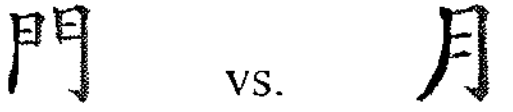
(3) THE CURVED VERTICAL STROKE

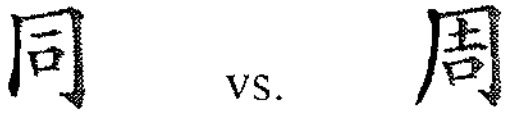
Roughly speaking, *the leftmost vertical stroke* in a character or a component is curved:

More precisely, it is curved if it is the leftmost stroke *and there is significant material to the right*:



Still more precisely, the curved vertical stroke only appears in a component that is *tall and narrow*:



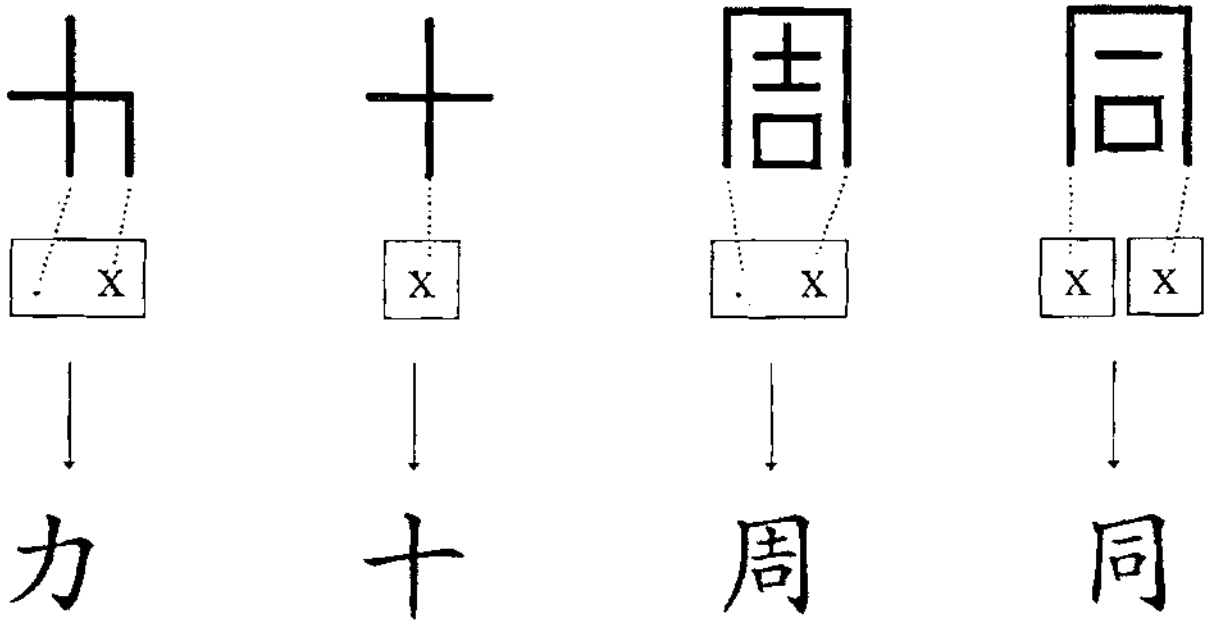


(There are other factors as well which will not be discussed here, e.g. curving does not occur with vertical strokes attached at the bottom or topped by a slanting stroke: 廿 乍)

These observations can be captured with a small set of *natural assumptions* and a *simple rule*:

(a) Only the rightmost element is linked to a head
(b) All the remaining elements are linked to the nonhead
(c) Wide characters are linked to two templates

RULE: **The leftmost vertical stroke becomes curved only if it is linked to a *nonhead*.**



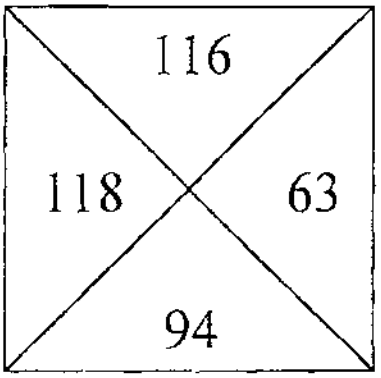
Without prosodic structure, Wang (1983) had to use a *more complex* and *less explanatory* description:

A vertical stroke on the extreme left of a character or component becomes curved if the matrix containing it has no fewer horizontal partitions than vertical partitions.

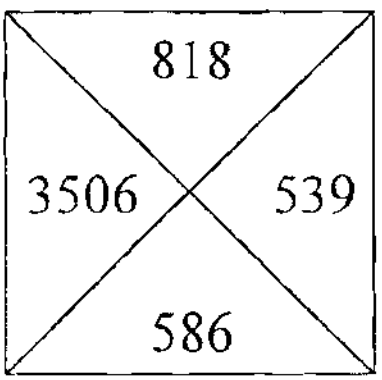
(4) RADICALS

Although radicals can wrap around or intersect with other components, most appear at the top, bottom, left and/or right. *The left is more common than the right, and the top is more common than the bottom.*

This is true whether one counts the *number of radicals* that can appear in at least one of these positions (Rose-Innes 1942)...



...or the *number of characters* with radicals in one of these positions (趙 1988):



Moreover, *"left" radicals* tend to be *narrow and vertical*, while *"top" radicals* tend to be *flat and horizontal*:

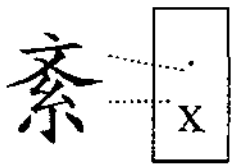
機 眼 綠 娟 茶 寫

Vertical radicals can appear at the bottom, but rarely at the top:

架 省 紊 娶

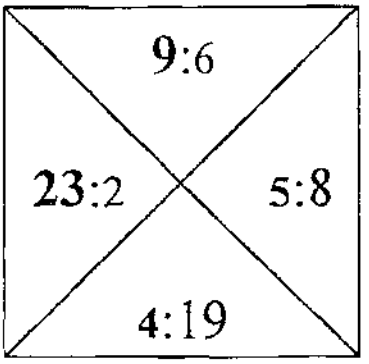
All of these observations can be explained by the hypothesis that ***radicals are linked to nonheads*** if they are *narrow or flat along the relevant dimension*:

Otherwise, they must be linked to *heads*:



This analysis also explains why many radicals appear in *reduced form at the top or left*:

Ratio of ***reduced***:*unreduced* radicals in 趙 (1988)

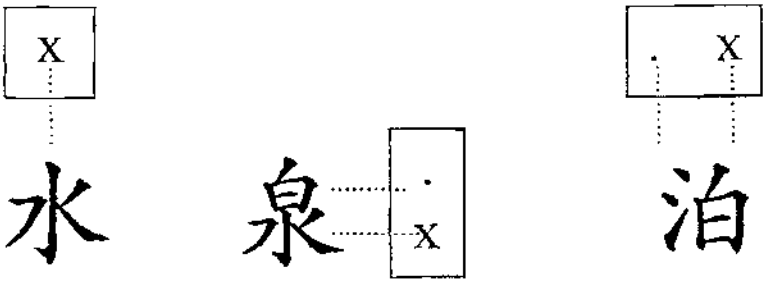


The relation between reduction and position is highlighted by radicals that appear in their *full form at the bottom* but in *reduced form at the left*:

心　忘　忙

水　泉　泊

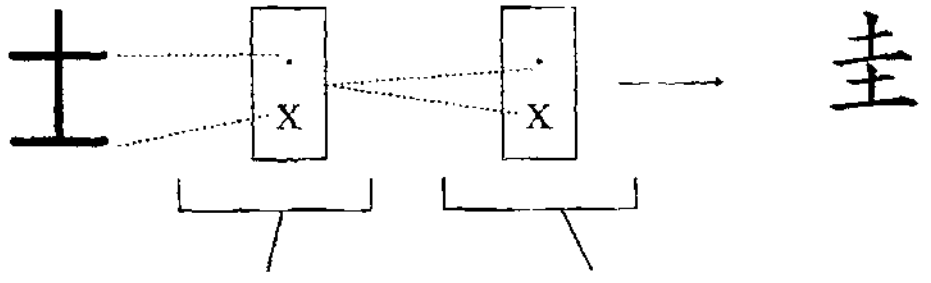
In other words, reduced radicals may only appear when *linked to a nonhead*:



EVIDENCE FROM INTERACTIONS

The claim that *all these patterns involve the same prosodic templates* is supported by interactions like the following.

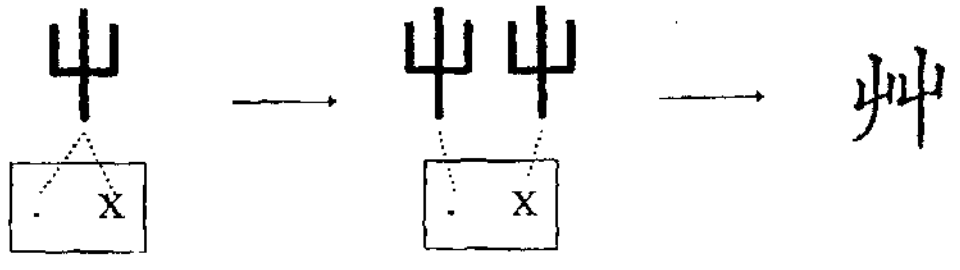
The ***interaction between prominence and duplication*** can be seen in the revised hierarchical analysis of the character 圭:



Describes difference in stroke size within component

Describes duplication and *also* difference in stroke size at character level

The ***interaction between duplication and curving*** suggests that duplication involves copying, although only a single template is used:



EVIDENCE FROM DIACHRONIC CHANGE

Even after the emergence of 楷 書, characters continued to be modified in ways consistent with the prosodic template analysis.

Characters that violated the templates have been lost. For example, Wieger (1965) lists the following obsolete characters with illegal duplication patterns:

ㄌㄧㄥˊ

"noise of voices"

ㄑㄧㄥ

"clamors"

In other components, *the prosodic structure has been changed.* Below, one component appears to have been reinterpreted as linked to a single template rather than two, as shown by the *merging of the horizontal strokes* and *curving of the leftmost vertical stroke*:

三 國 時 代

(趙 1988)

Modern form

CONCLUSIONS

Two distinct claims are argued for in this poster:

(A) Chinese characters have a prosodic structure analogous in interesting ways to that of spoken language;
(B) The proper analysis of this structure requires the specific template discussed here.

Claim (B) may be falsified or modified without affecting the validity of claim (A).

The fact that prosodic structure allows for the simple description of a diverse set of patterns implies that claim (A) is correct. The diachronic evidence suggests that prosodic structure was relevant to scribes in previous centuries.

It is reasonable to suppose that prosodic structure is still used by modern writers, and perhaps also readers, in the processing and mental storage of Chinese characters, although this has not yet been tested.

In any event, the evidence for prosodic structure in orthography shows that phonological principles may apply in domains quite distinct from natural speech.

REFERENCES

Eden, M. (1961) "On the formalization of handwriting." In R. Jakobson (ed.) *Structure of Language and Its Mathematical Aspects: Proceedings of Symposia in Applied Mathematics, Vol. XII*, American Mathematical Society.

Herrick, E. M. (1966) *A Linguistic Description of Roman Alphabets.* Hartford Studies in Linguistics, No. 19. MA Thesis, Hartford Seminary Foundation: Hartford, Connecticut.

Lerdahl, F. and Jackendoff, R. (1983) *A Generative Theory of Tonal Music.* MIT Press: Cambridge, MA.

McCawley, J. D. (1989) "Some graphotactic constraints," *University of Chicago Working Papers in Linguistics* 5:96-103.

McCarthy, J. J. and Prince, A. S. (1990) "Foot and word in prosodic morphology: the Arabic broken plural," *Natural Language and Linguistic Theory* 8:209-283.

Padden, C. A. and Perlmutter, D. M. (1987) "American Sign Language and the architecture of phonological theory," *Natural Language and Linguistic Theory* 5:335-375.

Poizner, H., Bellugi, U. and Klima, E. (1989) *What the Hands Reveal about the Brain.* Cambridge, MA: MIT Press.

Rose-Innes, A. (1942) *Beginner's Dictionary of Chinese-Japanese Characters.* Harvard University Press: Cambridge, MA.

Wang C.-S. (1983) *Toward a Generative Grammar of Chinese Character Structure and Stroke Order.* University of Wisconsin-Madison unpublished Ph.D. thesis.

Watt, W. C. (1979) "Iconic perspectives on linguistic explanation." In G. D. Prideaux (ed.) *Perspectives in Experimental Linguistics.* John Benjamins B. V.: Amsterdam.

Wieger, L. (1965 [1927]) *Chinese Characters: Their Origin, Etymology, History, Classification and Signification.* Reprinted by Paragon Book Reprint Corp. and Dover Publications: New York.

趙侶 (1988) 七體書法字典. 博文書社: 北京.

APPENDIX: EXCEPTIONS

Prominence appears in the middle in only one form:

Some familiar radicals are reduced on the right or the bottom, though this is not the dominant pattern:

刀　到　　火　照

Other exceptions are more general. For example, unlike "bottom" radicals, "right" radicals are not especially "heavy." In particular, they are never horizontal (recall that bottom radicals are usually vertical). I speculate that this is due to biases in visual perception which favor left-right reflective symmetry.

Finally, simplified characters used in the PRC violate some of the generalizations discussed here. For example, the vertical stroke in 书《書》 is not curved as predicted. Either my analysis is not describing psychologically real phenomena after all, or else at least some of the simplified characters are "prosodically unnatural."