ZHIMING BAO

SYLLABLE STRUCTURE AND PARTIAL REDUPLICATION IN CLASSICAL CHINESE*

Bisyllabic, partially reduplicated words in classical Chinese fall under two major categories: i-words and r-words. An i-word consists of syllables which share the same initial (*bi-ba*), whereas an r-word consists of syllables which share the same rhyme (*buk-suk*). Unlike partial reduplication of the more familiar type, partially reduplicated words in classical Chinese show no fixed segmental shape. An important property of such words is the tone pattern: if the two component syllables differ by only one segment, then they bear the same tone. The proper analysis of such data supports 1. the hierarchical structure of the syllable; 2. total copying in partial reduplication; and 3. the rime as the tone-bearing unit. In addition, the data show that the syllabic position of the medial glide in classical Chinese is indeterminate.

0. INTRODUCTION

Classical Chinese contains a large number of disyllabic words composed of syllables which share some common phonological material. Two specimens appear in (1).

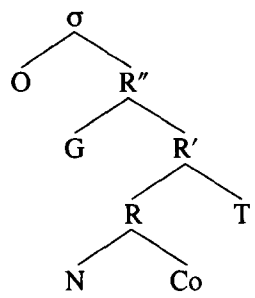
(1) a. ʔu T1 – du T1 'tiger'
b. p'ɑŋ T1 – p'at T4 '(of rain) heavy'

In (1a), the two syllables have the same rhyme, *u T1*, but different initials, *ʔ* and *d*. In (1b), they share the same initial, *p'*, but different rhymes, *ɑŋ T1* and *at T4*.[1] Typically, such dissyllabic, partially reduplicated words cannot be further decomposed, and many of them have variant written forms. They are not restricted to any morphological category. Among the nominal expressions, many are names of animals (such as those in (1a)) or plants. Often, their etymological origin is obscure, and they come into the Chinese language throughout its history.

The proper analysis of the partial reduplication data in classical Chinese has implications for theoretical issues of reduplication, syllable structure, and the representation of tone and the tone bearing unit. In this paper I investigate the formal properties of the kind of data exemplified in (1), and show that (i) partial reduplication must involve total copying; (ii) syllable structure is hierarchical; and (iii) tone is adjoined directly to the tone bearing unit without the skeleton (x-slots or CV templates) playing a mediating role. The syllable structure I argue for is shown in (2) (σ: syllable; O: onset; R/R′/R″; rime projections; G: glide; N: nucleus; Co: coda; T: tone):

Journal of East Asian Linguistics **4**, 175–196, 1995.

(2)



I will show that a flat syllable structure, most forcefully argued for in Clements and Keyser (1983), is inadequate to handle the partial reduplication data in classical Chinese. Perhaps the most startling property of the syllabic hierarchy shown is the recursion of the rime node (Harris (1983), Levin (1985)). The adjunction of tone (T) to rime (R), the tone bearing unit, creates an autosegmental *tier* on the syllabic *plane* (for the distinction between tiers and planes, see Archangeli (1985), Steriade (1986), Prince (1987), McCarthy (1989)). I will show that a structure in which tone is represented on its own autosegmental plane fails to explain the tone patterns that emerge from the partial reduplication data in classical Chinese.

The paper is organized as follows. In section 1 I discuss the patterns of partial reduplication, establishing the generalizations which govern the relationship between segments and tones of partially reduplicated words; in section 2 I give the analysis; in section 3 I discuss the syllabic position of medial glides and show that they are structurally indeterminate; in section 4 I discuss cases problematic for the proposed analysis; and the conclusion is in section 5.

1. Types of P-Words

A note on the collection and treatment of the data is in order. The main part of the data is collected from Wang (1968), which contains some 1,000 partially reduplicated words collected from classical writings; 45% of them are i-words, the rest are r-words. Additional data are collected from Fu (1960), Ju (1960) and Chou (1962).[2] The latter two are exhaustive studies of all reduplicated words from the *Book of Odes*, believed to have been edited by Confucius. From Fu (1960) I collected words which denote birds (characters with the 'bird' radical), animals ('animal' radical), fish ('fish' radical), insects ('insect' radical), plants ('tree' and 'bamboo' radicals), and weeds ('grass' radical). For obvious reasons the data from the various sources overlap. Where appropriate, the data are transcribed in the sound system of the *Guangyun*, a dictionary compiled around 1000AD, as reconstructed by Wang (1958). Words which are transcribed to reflect a period

earlier than that of the *Guangyun* will be marked by a star *.[3] In such cases, the tones will remain those of the *Guangyun*. This practice does not imply that tone does not change over time. But historical reconstruction of tone is hindered by lack of textual evidence; the *Guangyun* is the earliest complete document which classifies the characters in accordance with the four tones, *ping* 'even', *shang* 'rising', *qu* 'departing' and *ru* 'entering'.

In most reconstruction work, the maximal syllable in the *Guangyun* consists of four segments: C_1GVC_2, where $C_{1,2}$ is a consonant, G a glide, V a vowel (Wang (1959), Chen (1976); but see Ting (1979), Pulleyblank (1991)). C_2 may be a post-vocalic glide, a stop (*p/t/k/ʔ*) or a nasal (*m/n/ŋ*).[4] Tone is superimposed on the entire string (Wang (1956), Cheng (1973), Ting (1979)). For ease of exposition I will use these symbols as slots in the syllabic template of classical Chinese. In traditional terms C_1 is the *initial*; G, V, C_2 and tone (T), form the *rhyme*. I will examine the relationship between the segmental make-up of a p-word and its tone pattern and motivate a hierarchical structure of the syllable as we go along.

There is a direct relationship between the segmental properties of syllables making up a p-word and the tones that the syllables bear. Without exception, syllables of an r-word differ by C_1, and they bear the same tone. We have seen this in (1a). More specimens follow:

(3)	a. buk T4 – suk T4	'small trees'
	b. mɑk T4 – ɣɑk T4	'mantis'
	c. kəŋ T3 – məŋ T3	'snouted sturgeon'
	d. kəu T2 – ləu T2	'arched top of a cart'
	e. kɑ T1 – ŋɑ T1	'wild goose'
	f. ɣa T1 – ma T1	'toad'
	g. ʔu T1 – du T1	'tiger'
	h. ɣɒm T2 – dɒm T2	'lotus'

The data in (3) all show different C_1, identical VC_2 and tone. Thus, the first syllable in (3c) begins with *k*, and the second syllable with *m*. Their VC_2 and tones are identical: both syllables have *əŋ T3*.

The segmental makeup of i-words shows a great deal of variation, and so do their tone patterns. If the two syllables of an i-word differ only by G, then the tones are identical:

(4)	a. tĭan T2 – tyăn T2	'toss and turn'
	b.*k'ĭan T2 – *k'yăn T2	'emotionally attached'
	c. ɣɑ T1 – ɣwɑ T1	'lotus'
	d.*mĭan T1 – man T1	'uncivilized(?)'
	e. k'ĭăt T4 – k'uăt T4	'miserable'

In (4a,b), the first syllable has the pre-vocalic glide *ĭ*, the second syllable has *ŭ*. In (4c,d), one syllable contains a glide, and the other lacks one. And in (4e), *ĭ* contrasts with *u*. The syllables share the same C_1, VC_2, and tone.

The next set of data show that if the component syllables differ by V, the tones are identical:

(5) a.	siəu T1 – sĭau T1	'red spider'
b.	bi T1 – ba T1	'(a kind of fruit)'
c.	nʐĭěm T2 – nʐĭɛm T2	'soft'
d.	mĭěm T2 – mĭɛm T2	'encouraging'
e.	lĭět T4 – lĭăt T4	'cold'
f.	ĭap T4 – ĭəp T4	'wet'

I have been unable to find i-words showing C_2 to be the only segment which differs between their component syllables. So I have nothing to report on the tone pattern of this type of i-word. i-words of this type do occur when C_2 is a stop. I will discuss this type later.

If the syllables differ by more than one segment, the tone patterns are not predictable. Consider the data in (6):

(6) a.	dɑu T1 – du T1	'wild horse'
b.	ĭəu T1 – ĭɛn T1	'centipede'
c.	kai T2 – kəu T1	'water-chestnut'
d.	liɛu T1 – lĭaŋ T3	'loud'

All words show different VC_2 in their respective syllables: the first syllable in (6a) contains *au*, while the second syllable lacks C_2. In (6b), it is *əu* vs. *ɛn*, and so on. The tones are identical in (6a,b), different in (6c,d).

The data in (7) show different GV in their respective syllables. The tones are the same in (7a), different in (7b).

(7) a.	tieŋ T2 – tuŋ T2	'rush'
b.	tien T1 – tɑu T2	'turn upside down'

The data in (8) show different GVC_2 between the component syllables. Again the tones are not predictable:

(8) a.	kiem T1 – ka T1	'reed'
b.	ma T2 – mŭəi T2	'Chinese pokeweed'
c.	kəu T2 – kiei T3	'Lycium chinense'
d.	tiei T3 – tuŋ T1	'rainbow'

In (8a), the first syllable contains *iem*, the second syllable, *a*. In (8c), the

first syllable has *əu*, the second syllable, *iei*. Such i-words have variable tone patterns: the tones are identical in (8a,b), different in (8c,d).

The data in (9) show no segmental difference; only tonal difference is observed:

(9) a.	gĭə T3 – gĭə T1	'owl'
b.	ʔўe T1 – ʔўe T3	'(a kind medicinal herb)'
c.	ʔɑn T3 – ʔɑn T1	'peaceful'

Due to the tonal difference, the p-words are considered i-words rather than r-words. They are not cases of total reduplication.

When the tones are the same, we have total reduplication:

(10) a.	pɐŋ T3 – pɐŋ T3	'worrisome'
b.	mɑŋ T1 – mɑŋ T1	'big'
c.	t'ɑu T1 – t'ɑu T1	'(water) flowing'
d.	tsɑn T3 – tsɑn T3	'bright'

In addition to the forms discussed above, we do not find the forms enumerated in (11):

(11) a. cgvc t – c'gvc t'
b. cgvc t – cg'vc t'
c. cgvc t – cgv'c t'
d. cgvc – c'gv'c
e. cgvc – c'gvc'

In (11), lower-case letters represent segments, and primed segments in the second syllable are different from their counterparts in the first syllable. (11d,e) do not occur, regardless of tone.[5] *Prima facie*, these types may conceivably be partially reduplicated since the component syllables share some common segmental material: *gvc* in (11a) and *c,vc* in (11b), and so on. Their non-occurrence needs to be explained.

This completes the survey of the data. The descriptive generalizations that emerge from the discussion are given in (12).

(12) a. If the syllables differ in C_1, they have the same tone.
b. If the syllables differ in G, they have the same tone.
c. If the syllables differ in V, they have the same tone.
d. If the syllables differ in more than one segment, they may, but need not, have the same tone.

2. The Analysis

Before we proceed with the analysis, three methodological issues need to be addressed. First, in any theory of reduplication (Marantz (1982), Yip (1982), Broselow and McCarthy (1983), Clements (1985), McCarthy and Prince (1986), Steriade (1988), Bao (1990a)), some form of base copying is involved. The theory must therefore know what the base of copying is. In the case at hand, the etymological origin of the data is obscure. We do not know which syllable (i.e., character) of a p-word is the base from which the p-word is derived through reduplication. Second, the partial reduplication data do not show an invariant phonetic form. The two syllables of an r-word may have any segment in the C_1 position; and those of an i-word may have any number of segments making up their GVC_2. In addition, modern reconstruction of classical Chinese differs from one scholar to another (Karlgren (1948), Dong (1954), Wang (1959), Ju (1960), Shen (1960), Chou (1962), Chen (1976), Pulleyblank (1984, 1991), Yu (1985), and references cited there). For these reasons our analysis cannot commit us to etymological claims which cannot be empirically verified or claims about the phonetic shape of partial reduplication or fixed reduplicative templates with or without pre-specified material.

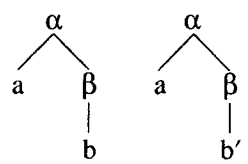
Third, the p-words we are studying were not created at the same time (Wang (1958)), and there is no clear philological evidence that the mechanism serves any productive morphological function with identifiable semantic content. This poses a serious problem for our analysis, which for practical purposes approaches the data from a synchronic perspective. The choice of the *Guangyun* is linguistically arbitrary but sinologically sound. The *Guangyun* is a milestone in Chinese linguistics on which reconstruction of earlier phonological systems is based and from which later systems are derived. Although our approach does not shed light on the exact nature of p-word formation, it nevertheless allows us to examine the syllable structure of the p-words in the *Guangyun* phonology.[6]

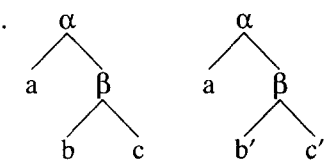
The analysis to be developed here focuses on the relationship between the two syllables of a p-word in terms of the syllabic constituent by which they differ. The intuition is that the mechanism of partial reduplication in classical Chinese, whatever it may be, produces properties which can be extracted out of p-words. Such properties, which I will refer to as P_σ, in turn shed light on the nature of partial reduplication, syllable structure, and the position of tone within the syllable. The analysis therefore takes the form of reverse engineering: P_σ expresses the properties of p-words as created by rules of partial reduplication, which remain obscure due to the historical nature of the data in question.

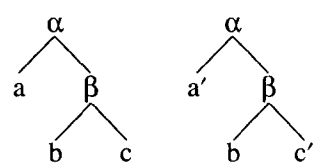
2.1. *The Property* P_σ

The main analytical tool is the notion P_σ. It picks out the smallest constituent by which the two syllables of a p-word differ. For example, if one syllable of a p-word has *a* as the nuclear vowel and the other syllable has *e*, other things being equal, the nucleus is the smallest constituent which differentiates the two syllables of the p-word – the two syllables are 'related' by their differing nuclei. In this case the p-word is said to have property $P_\sigma(N)$, where N is the nucleus. P_σ expresses relations between segments or sequences of segments in the same syllabic position. Thus, the r-word such as *buk T4 – suk T4* (3a) has the property $P_\sigma(C_1)$ because *b* in the first syllable differs from *s* in the second syllable, and both segments occupy syllable-initial position. P_σ says nothing about *b* in the first syllable and *k* in the second, since they occupy different positions in their respective syllables. All r-words have the property $P_\sigma(C_1)$.

To illustrate the notion P_σ further, consider the hypothetical p-words in (13), where lower-case letters are segments (terminal nodes) and Greek letters represent non-terminal constituents:

(13) a. 

b. 

c. 

In (13a), the p-word has the property $P_\sigma(\beta)$ since that is the constituent by which the two component syllables differ. In (13b), the p-word also has the property $P_\sigma(\beta)$, β being the node which dominates *b*/*b′* and *c*/*c′*. In (13c), α is the minimum node that dominates the segments *a*/*a′* and *c*/*c′*, the p-word therefore has the property $P_\sigma(\alpha)$, even though both syllables share the same segment *b*.

I stipulate the following well-formedness conditions on p-words:

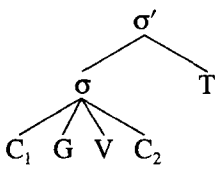
(14) Let W be a p-word of the form σ_1–σ_2
   a. there is at least one subsyllabic constituent α such that W has the property $P_\sigma(\alpha)$.
   b. there is at most one subsyllabic constituent α such that W has the property $P_\sigma(\alpha)$.

A p-word must therefore have one and only one constituent which differentiates the two components. Totally reduplicated words are not p-words because there is no constituent α such that they have the property $P_\sigma(\alpha)$, violating (14a).

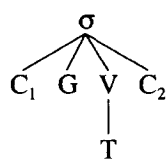
2.2. *Syllable Structure*

Two representations may be dismissed as inadequate in explaining partial reduplication in classical Chinese. Neither the 'flat' structure (Clements and Keyser 1983) in (15a), with tone adjoined to the syllable, nor the structure in (15b), with tone on its own autosegmental plane, is capable of explaining the tone patterns generalized in (12).

(15) a.



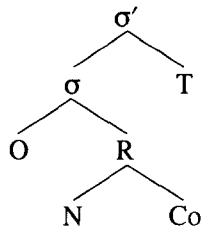
b.



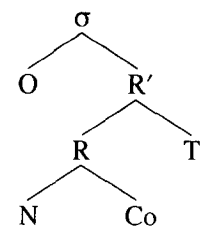
Independent of how tone is represented, the structure in (15a) makes the generalizations in (12) mysterious: there is no explanation why tones of a p-word may differ if its component syllables differ by more than one segment, as in i-words like *kai T2 – kəu T1* (6c). Whatever the rules of partial reduplication, they must be able to refer to contiguous segments, VC_2, indicating a closer relationship between V and C_2 than between V and C_1, or between C_1 and T. The same criticism can be directed at the structure in (15b), which is the standard representation in current non-linear phonology (for a recent exposition of the theory, see Goldsmith (1990)). In this representation tone forms an autosegmental plane and is therefore expected to exhibit stability from operations taking place on other autosegmental planes. To derive the generalizations, one needs to stipulate that tone may be affected only when two or more segments are being targeted by rules of partial reduplication. The stipulation is needed whether the syllable itself is hierarchical or flat. Such stipulation is *ad hoc*. Tone interacts with segments in partial reduplication, and such interaction needs to be captured directly in the representation.

Consider now how the elements of a syllable C_1GVC_2T may be grouped into hierarchical constituency. Putting aside the constituency of G for the moment, we can have the following two hierarchical structures:

(16) a.



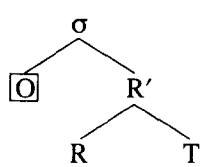
b.



where O is the onset (= C_1), N the nucleus (= V) and Co the coda (= C_2). The only difference between the two structures is the location of tone: it is adjoined to the syllable σ in (16a) (Cheng (1973), Ting (1979), Lin (1989)), to the rime R in (16b) (Bao (1990a, 1990b)), in accordance with the traditional analysis which groups tones in the rhyme. The two structures make the same prediction concerning the tone patterns of r-words. Both predict, correctly, that r-words have the same tone since they have the property P_σ(O), and O does not dominate tone (T) in either structure. But (16a) makes the wrong prediction concerning the tone patterns of i-words: it predicts that all i-words have the same tone since T is sister to σ. This prediction is wrong; i-words may have different tones: *kai T2 – kəu T1* 'water-chestnut' (6c). See the generalization stated in (12d).

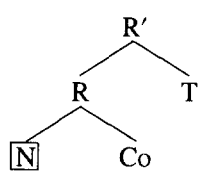
Given the structure in (16b), we may deduce the following types of tone patterns of partial reduplication:

(17) a. P_σ(O):



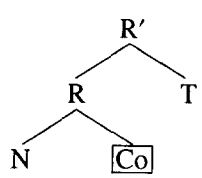
tones are identical

b. P_σ(N):



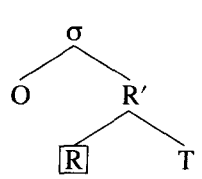
tones are identical

c. P_σ(Co):

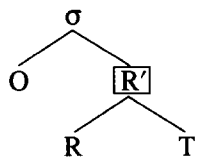


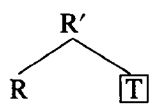
tones are identical

d. P_σ(R):



tones are identical

e. $P_\sigma(R')$:  tones may be different

f. $P_\sigma(T)$:  tones are different

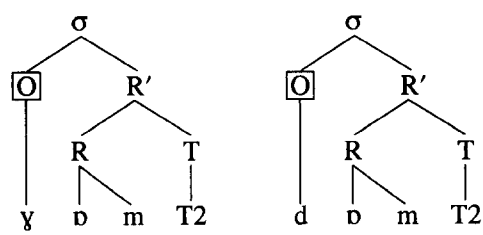
p-words with the property P_σ(Co) will be dealt with in section 2.5; those with the property P_σ(G) are discussed in section 3. The types in (17a,b,d,e,f) are all attested among the partial reduplication data: P_σ(O) data are shown in (3); P_σ(N) data in (5); P_σ(R) data in (6a,b), P_σ(R′) data in (6c,d), and P_σ(T) data in (9).

2.3. *Derivations*

The analysis we have developed has three components, given in (18):

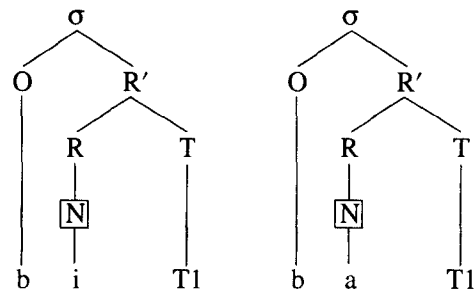
(18) a. Property P_σ
b. Well-formedness conditions (14)
c. Hierarchical syllable structure (16b)

In addition, we need to assume that a p-word is derived from a single syllable, which is copied to produce a disyllabic string. The derivations below illustrate how the analysis works (unnecessary nodes are omitted):

(19) a. Derivation of (3i): *ɣɒm T2 – dɒm T2* 'lotus'
Base syllable: ɣ.ɒm T2
Copying: ɣ.ɒm T2 – ɣ.ɒm T2
P_σ(O): 

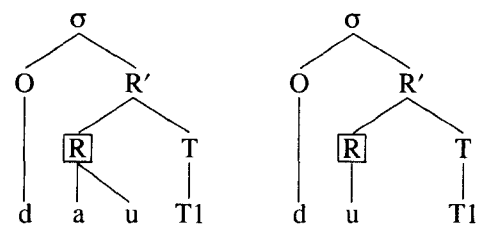
b. Derivation of (5b): *bi T1 – ba T1 a kind of fruit*
Base syllable: b.a T1
Copying: b.a T1 – b.a T1

P_σ(N):



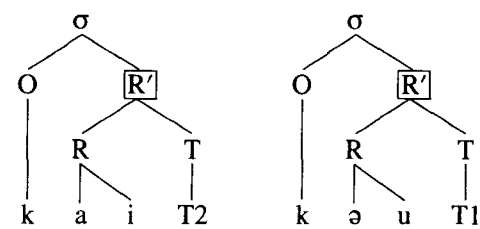
c. Derivation of (6a): *dau T1 – du T1* 'wild horse'
Base syllable: d.u T1
Copying: d.u T1 – d.u T1

P_σ(R):



d. Derivation of (6c): *kai T2 – kəu T1* 'water-chestnut'
Base syllable: k.ai T2
Copying: k.ai T2 – k.ai T2

P_σ(R′):



All derivations start with a base syllable, which is copied to yield a string of two exact copies. The rules of partial reduplication, whatever they may be, apply to derive partially reduplicated words with specific properties: P_σ(O) in (19a), P_σ(N) in (19b), P_σ(R) in (19c), and P_σ(R′) in (19d). The choice of base syllables is random.

2.4. *The Generalizations Explained*

We now proceed to derive the generalizations in (12) and explain the non-occurrence of the forms in (11). We need one more assumption, namely, (20) (Steriade (1988), Bao (1990a)):

(20) Partial reduplication involves total copying of the base

This guarantees that the two syllables of a p-word have exactly the same phonological content prior to the rule(s) of partial reduplication, whatever they might be. The generalizations in (12) follow from our analysis without further stipulation. (17a,b) provide an explanation of the generalizations in (12a,c), while (17d,e) lead to the generalization in (12d). Note that R′ and R dominate the same string of segments. Therefore, if a p-word differs by VC_2, it may have the property $P_\sigma(R)$ or $P_\sigma(R')$. This means that such p-words may have identical tones when they have the property $P_\sigma(R)$ or different tones when they have the property $P_\sigma(R')$: the content of the generalization (12d). The generalization (12b) will be explained shortly.

The non-occurrence of the forms in (11) also follows from our analysis. This is partially due to the well-formedness conditions (14), which require that a p-word have one and only one property expressible by $P_\sigma(\alpha)$ and partially due to the constituency requirement on α. Thus, the form *cgvc-c'gv'c* (cf. (11d)) cannot have the properties $P_\sigma(O)$ and $P_\sigma(N)$ at the same time, this possibility being ruled out by (14b). Since *c* and *v* do not form a constituent, the form, lacking a single syllabic constituent expressible by P_σ, violates (14a). The same argument applies to the other forms in (11).

2.5. *The Gap in the Data:* P_σ(Co)

As we noted before, one gap in our data is the absence of i-words of the form $C_1GVS\text{-}C_1GVS'$, i.e., i-words with the property $P_\sigma(Co)$, where Co is a sonorant (nasal or glide, denoted by S). The prediction is that they have the same tones. *Prima facie* this prediction is not borne out in i-words with the property $P_\sigma(Co)$, where Co is an obstruent, as the following data show:

(21) a.	diei T1 – diet T4	'weed'
b.	luk T4 – lu T1	'well pulley'
c.	lĭuk T4 – lĭu T1	'wild goose'
d.	*ʈĭɑ T1 – *ʈĭɑk T4	'sugar cane'

In classical Chinese, as in many modern Chinese dialects, there is an interdependency between syllables ending in a stop and T4. Based on this observation we can speculate that i-words of this type have identical tones at some level of derivation. Thus, the data in (21) may be derived from forms in (22) by a rule such as the one given in (23):

(22) a. diei T1 – diet T1
b. luk T1 – lu T1
c. lĭuk T1 – lĭu T1
d. *ʈĭɑ T1 – *ʈĭɑk T1

(23) T → T4 /VC]$_\sigma$, where C is *p, t, k* or *ʔ*

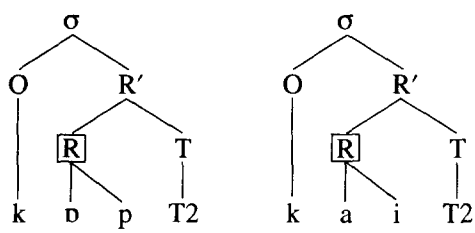
It must be acknowledged that i-words with the property P_σ(Co) are compatible with the predictions of our analysis, not compelling. p-words of the form C_1GVST – C_1GVS'T', if found, constitute counterexamples to the proposed analysis.[7]

Given the rule in (23), we may consider the i-words in (24) as having the property P_σ(R), rather than P_σ(R′):

(24) a. fĭuk T4 (T2) – fĭəu T2 'turtle-dove'
b. xĭɐt T4 (T1) – xĭɛu T1 'short-nosed dog'
c. kɒp T4 (T2) – kai T2 'gecko'

T4 is derived from T2 and T1, respectively. As an illustration, consider the structure of the i-word (24c), shown in (25):

(25)



Since T4 is derived from T2 by the rule in (23), the minimum constituent that differentiates the two syllables in (25) is R, rather than R′. The same is true of (24a,b), *modulo* the glides.

3. The Medial Glides

Let's now consider the position of the medial glides within the syllable structure. Typically r-words have identical glides:

(26) a. xuɒi T1 – duɒi T1 'sick'
b. xўəi T2 – jўəi T2 'abundant'
c. byən T1 – ʔyən T1 'intertwining'
d. kiep T4 – siep T4 'sulphur butterfly'
e. vĭuk T4 – fĭuk T4 'hoopoe'
f. k'ĭuŋ T1 – gĭŋ T1 '(a kind of fragrant herb)'

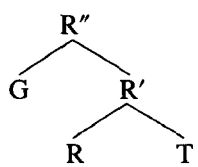
g.	gǒoŋ T1 – sǒoŋ T1	'despicable'
h.	biei T2 – liei T2	'ficus pumila'
i.	ʔieu T2 – nieu T2	'(a kind of fast horse)'
j.	tsǐo T1 – gǐo T1	'galleyworm'

Since r-words have the property P_σ(O), these data suggest that medial glides are part of the rime R. The i-words below also support syllabifying the medial glides into the rime:

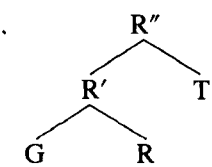
(27)	a.	dɑŋ T1 – dieu T1	'cicada'
	b.	dɑŋ T1 – diei T3	'aspen'
	c.	kəu T2 – kiei T3	'licium chinense' (= (8c))
	d.	tieŋ T2 – tuŋ T2	'rush' (= (7a))
	e.	kiem T1 – ka T1	'reed' (= (8a))
	f.	tiei T3 – tuŋ T1	'rainbow' (= (8d))
	g.	mǐɛn T1 – ma T2	'cicada'
	h.	myən T1 – mǐu T1	'night owl'
	i.	kyət T4 – kǐəu T1	'falcon'
	j.	duok T4 – *dǐo T1	'spider'
	k.	k'ǐět T4 – k'uət T4	'grub'
	l.	bǐu T1 – bỹɐt T4	'antelope'

In all cases the medial glides in the two component syllables are different. In (27a–c), for example, the first syllable has no medial glide; the second syllable has the glide *i*. Other glide pairs are: *i*/*ǐ* versus *ø* in (27d–g), *y* versus *ǐ* in (27h,i), *u* versus *ǐ* in (27j), *ǐ* versus *u* in (27k), and *ǐ* versus *ỹ* in (27l). The tone patterns are expected: some have the same tone (27a,d,e,h,k); others have different tones. These i-words all seem to show that the medial glides form a constituent with the following segment(s). We have shown that R and R′ must be constituents as defined on the tree (16b) since there are p-words with the properties P_σ(R) and P_σ(R′). Therefore the medial glides may be analyzed as sister to R′ or R, as shown in the structures below:

(28) a.



b.



Immediately, we derive the generalization in (12b): if a p-word has the property P_σ(G), then the p-word must have the same tone, under either of the two structures in (28).

These structures, however, make different predictions. Consider a p-word with the segmental string *cgvc – cgv'c'*. The structure (28a) predicts two

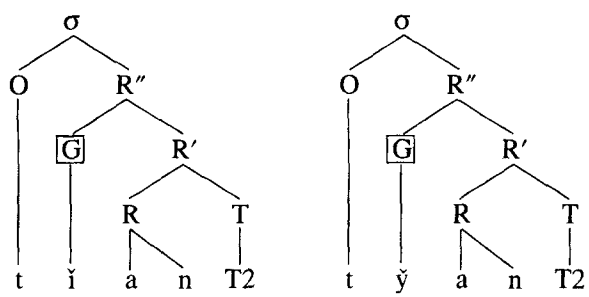
possible analyses: the p-word may have the property $P_\sigma(R)$ or the property $P_\sigma(R')$. The structure (28b) predicts only one possible analysis: such p-words must have the same tone, with the property $P_\sigma(R)$. Empirical evidence supports (28a):

(29) a.	ĭəu T1 – ĭɛn T1	'centipede' (= (6b))
b.	dzĭəu T1 – dziei T1	'beetle larvae'
c.	lĭəu T1 – lĭe T1	'eared owl'
d.	liɛu T1 – lĭaŋ T3	'loud' (= (6d))
e.	fĭu T1 – fĭəu T2	'(a kind of bird)'

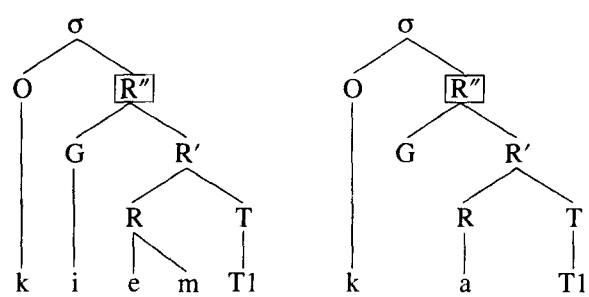
The words in (29a,b,c) have identical tones, while those in (29d,e) have different tones, contrary to the prediction of the structure in (28b). The tone patterns exhibited in (29) are the expected result; see the generalization stated in (12d).

The derivations involving the medial glides are shown in (30). G is postulated for ease of reference only.

(30) a. Derivation of (4a): *tĭan T2 – tỹan T2* 'toss and turn'
Base syllable: t.ĭan T2
Copying: t.ĭan T2 – t.ĭan T2
$P_\sigma(G)$:



b. Derivations of (8a): *kiem T1 – ka T1* 'reed'
Base syllable: k.iem T1
Copying: k.iem T1 – k.iem T1
$P_\sigma(R'')$:



As we remarked earlier, tones need not be different in i-words with the property $P_\sigma(R')$. The i-word, *kiem T1 – ka T1*, shown in (30b), has the same tone, whereas the i-word in (8d), *tiei T3 – tuŋ T1* 'rainbow', has different tones. The two i-words are structurally identical.

From the structure in (28a) we predict the following:

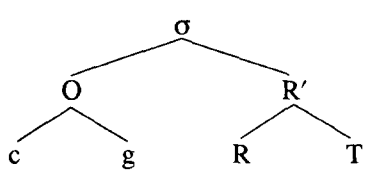
(31) r-words of the form *cgvc t – c'g'vc t* do not occur.

The empirical validity of this prediction depends on one's intuitive notion of r-words. In Wang (1956), r-words are defined as p-words whose syllables share the same tone, medial glide (if any), vowel, and coda segment (if any). Under this view (31) is empirically valid. In Ju (1960) and Chou (1962), r-words are classified into two classes: strict versus loose. The so-called strict r-words meet the definition given by Wang; loose r-words are similar to their strict counterparts except for the medial glides: the syllables making up a loose r-word have different medial glides, or one syllable contains a medial glide, and the other does not. In no one's pre-theoretic intuition are p-words with different tones classified as loose r-words. The proposition in (31) is true of strict r-words, by definition. The loose r-words, however, are anomalous. A few specimens follow:[8]

(32) a.	ɣuɐn T1 – lɐn T1	'(a kind of fragrant herb)'
b.	*məu T1 – *diəu T1	'(a kind of cicada)'
c.	tɕĭək T4 – mək T4	'bat (mammal)'
d.	bɑŋ T1 – ɣwɑŋ T1	'(a kind of grain)'
e.	bwæt T4 – k'æt T4	'(a kind of weed)'
f.	t'ieŋ 2 – ɣyeŋ T2	'(a kind of herb)'
g.	*mўan T3 – *ian T3	'(big animal like the fox)'

These data suggest that the medial glides form a constituent with the preceding consonant, so they have the property $P_\sigma(O)$, as all r-words do. This analysis leads to the structure below:

(33)



Here we are faced with a theoretical dilemma: either we exclude loose r-words from the set of r-words and arrive at the structure in (28a); or we include loose r-words in the set of r-words, and thus our pre-theoretic notion of r-words would include both strict and loose r-words. In this case the medial glide is structurally ambiguous, analyzable as part of the rhyme (28a)

or part of the onset (33). This is a dilemma out of which there is no safe escape. To exclude loose r-words from consideration we must resort to definition; to include them we must admit the structural indeterminacy of the medial glide.

4. Some Problematic Cases

The proposed analysis accounts for the structural properties of strict p-words. It has no exception in all my sources: Chou (1962), which comprises p-words from the *Book of Odes*; Wang (1968), which contains p-words from classical writings (including, but not limited to, the *Book of Odes*); and Chen (1992), which contains p-words from the *Guangyun*. It is not surprising since the generalizations stated in (12) are true not only of p-words in classical Chinese, but also of partially reduplicated game languages (the so-called *fanqie* languages, see Chao (1931), Yip (1982), Li (1985), Lin (1989), Bao (1990a)), and other partial reduplication data in modern Chinese (Bao (1993)).

There are a few dissyllabic words which resist analysis. These are the so-called loose p-words (or sub-p-words) in the terminology of traditional Chinese phonological literature (see footnote 2). Some of the data are discussed as examples of *dui-zhuan* 'interchange,' whereby one rhyme with a coda parallels one without one (*uk* vs. *u*) (Chen (1992)). Strictly speaking, the loose p-words are not p-words at all since they do not satisfy the definition of p-words. They can be broken down into two types. Of the first type are words which look like r-words but whose component syllables show an additional difference. The data in (34) are of this type:

(34) a.	ʔĭək T4 – ĭə T2	'plantago'
b.	biek T4 – *t'iei T1	'wild ducklings'
c.	dzĭĕt T4 – liei T1	'crickets'
d.	lu T1 – bĭuk T4	'daikon radish'
e.	diei T1 – kyet T4	'goatsucker'
f.	*k'əi T2 – *diei T2	'harmonious'

(34a–c) data show two differences, in the onset and in the coda (the first syllable ends in a stop, hence acquires the tone T4). In addition to the onset and coda difference, (34d,e) show different medial glides. According to Wang (1958, section 9), *əi* and *(i)ei* in (34f) historically belong to two different rhymes, Rhyme 8 and Rhyme 6, respectively (see footnote 4). They are considered 'close enough' so that the word counts as a r-word. The properties exhibited by these data cannot be expressed with a single P_{σ}.

The other type of problematic cases is exemplified in (35):

(35) a. sĭět T4 – ʃyět T4 'cricket'
b. sĭuk T4 – ʃĭaŋ T1 '(a kind of fine horse)'

(35a) shows different medial glides (P_σ(G)), and (35b) shows different rimes (P_σ(R)). Both show an additional alternation, *s* versus ʃ. They look like i-words but are more complicated than the ones we have seen. According to Wang (1958) the two segments (*s* and ʃ) are considered 'similar' enough to make the words in (35) legitimate i-words. This notion of similarity, however, is not expressible within the proposed analysis.

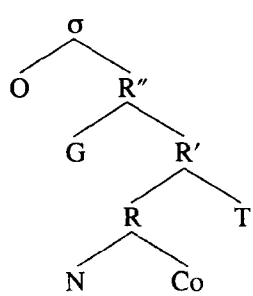
5. Conclusion

In the preceding sections we have presented an analysis of partial reduplication in classical Chinese. We have shown that partially reduplicated data, though diverse in their phonetic shapes, can nevertheless be analysed in a simple manner. The various forms of partial reduplication can be described and explained by means of a few formal devices. The main tool in our analysis is the notion P_σ, which expresses the relation between the two syllables of a p-word. A p-word must meet the two conditions stated in (14), repeated below:

(14) Let W be a p-word of the form $\sigma_1 - \sigma_2$,
a. there is at least one subsyllabic constituent α such that W has the property $P_\sigma(\alpha)$
b. there is at most one subsyllabic constituent α such that W has the property $P_\sigma(\alpha)$

We have shown that the two syllables of a partially reduplicated word share certain properties that indicate total copying of a common, monosyllabic base. Although we do not know the exact mechanism which forms the p-words, the tone patterns we noted in (12) will be inexplicable if it does not involve total copying. In addition, we have shown that syllable structure is hierarchical, as in (36):

(36)



supporting a large body of research on the internal structure of the syllable (Steriade (1982), Harris (1983), Levin (1985), Itô (1986), Fudge (1987), Kenstowicz and Rubach (1987), Lin (1989), Duanmu (1990), among others). A flat syllabic structure is not compatible with the data.

An important theoretical consequence of our analysis is that tone does not form a separate autosegmental plane. If tone were represented as forming an autosegmental plane, the generalizations on tone patterns of partial reduplication stated in (12) cannot be explained. One would have to resort to *ad hoc* stipulations to derive the structural dependency of tone on tone bearing units (here, the rime). Such dependency is captured in terms of the adjunction structure in (36). Tone mapping rules (Williams (1971/76), Goldsmith (1976, 1990), Pulleyblank (1986)) adjoin tone to tone bearing units directly, creating a projection of the latter in the sense of X-bar theory.

NOTES

* Earlier versions of this paper were presented at the 1991 Annual Meeting of the Linguistic Society of America and at the Linguistics Colloquia at the University of Illinois, Urbana. I benefitted from discussions with J. Cole, O. Fujimura, M. Halle, M. Kenstowicz, C.-L. Shih, R. Sproat, P.-H. Ting, and three anonymous reviewers. Part of the research was conducted while I was a Postdoctoral Fellow at the Ohio State University, Columbus, Ohio, under the faculty sponsorship of Professor Osamu Fujimura. I am grateful to Professor Fujimura for providing a stimulating academic environment and to the Ohio State University for the generous financial support. All errors of fact and interpretation are of my own responsibility.

[1] Following traditional practice, I will refer to words of the type shown in (1a) as double rhyme words, or r-words, and those of the type shown in (1b) as double initial words, or i-words. Collectively, r-words and i-words will be referred to as p-words.

[2] In the course of revising earlier versions of this paper I came across Chen (1992), which is an exhaustive study of the partially reduplicated words contained in the *Guangyun*. Chen traces each p-word to the earliest classical document in which it occurs and divides them into two periods, the Han dynasty (206BC–23AD) being the dividing line. By my count, the number of p-words are as follows:

(i)		i-words	r-words	Sub-i-words	Sub-r-words
	Pre-Han/Han	139	286	49	74
	Post-Han	57	254	34	20

Sub-i-words are like i-words except their component syllables share similar, though not identical, initials; likewise sub-r-words are like r-words, except their component syllables share similar rhymes (see section 4). Chen (1992) has three additional categories: i-words which are also sub-r-words, r-words which are also sub-i-words, and sub-i-words which are also sub-r-words. I include them in i-words, r-words, and sub-i-words, respectively. Chen's corpus contains less data than Wang's (1968). It must be emphasized that all the sources have overlapping data, so the numbers should be treated with caution.

[3] The diachronic effect of p-words can be summarized as follows:

(i)	Time of *Guangyun*	Earlier Time
a.	Yes	Yes
b.	Yes	No
c.	No	Yes

A word may satisfy the definition of a p-word at the time of the *Guangyun* and earlier (a); or at the time of the *Guangyun* but not at an earlier time (b); or at an earlier time but not at the time of the *Guangyun* (c). Only words of the third type will be transcribed with an *. For methodological justification of this practice, see section 2.

[4] The *Guangyun* has the following initials and rhymes (Wang (1958)):

(i) Initials of the *Guangyun*

Labial:	p	p'	b	m	Coronal:	t	t'	d	n	l
Velar:	k	k'	g	ŋ		ʈ	ʈ'	ɖ		
Glottal:	ʔ	j	x	ɣ		ts	ts'	dz	s	z
						tʃ	tʃ'	dʒ	ʃ	
						tɕ	tɕ'	dʑ	ɕ	ʑ nʑ

(ii) Rhymes of the *Guangyun*

1	uŋ, uk	2, 3	oŋ, ok	4	ɔŋ, ɔk
5	e	6	i	7	ə
8	əi	9	o	10-1	u
12	ei	13	ɛi	14	ɑi
15	ai	16	ɐi	17	æi
18, 19	ɒi	20	ɐi	21–2	ěn, ět
23	en, et	26	ɐn, ɐt	24–5, 27–8	ən, ət
29, 30	ɑn, ɑt	31	an, at	32	æn, æt
33	en	34	ɛn, ɛt	35	eu
36	ɛu	37	au	38	ɑu
39, 40	ɑ	41	a	42	aŋ, ak
43	ɑŋ, ɑk	44	ɐŋ, ɐk	45	æŋ, æk
46	ɛŋ, ɛk	47	eŋ, ek	48–9	əŋ, ək
50–2	əu	53	ěm, ěp	54	ɒm, ɒp
55	ɑm, ɑp	56	ɛm, ɛp	57	em, ep
58, 60–1	ɐm, ɐp	59	am, ap		

The glottal stop *ʔ* is reconstructed by Wang as the zero-initial; here I follow Pulleyblank (1991). The voiced obstruents are reconstructed with aspiration; since it is not contrastive, following Guo (1986), I leave out the aspiration marker on the voiced obstruents. The rhymes are shown without pre-vocalic glides, and the numbers are rhyme numbers used in Wang (1958). The pre-vocalic glides are: *i, ǐ, u, w, iw, ǐw, ǐu*. I treat the last three as single glides (Wang (1959), Chen (1976)), and notate them as follows: *y* for *iw*; *ỳ* for *ǐw*, and *ǔ* for *ǐu*. The pre-vocalic glides co-occur with vowels of different 'grades'. Thus, the glide *i* is used with Grade 4 vowels, as in *ien*; *ǐ* with Grade 3 vowels, as in *ǐɛn* (Wang (1959, section 10)). T1 through T4 represent tones, as follows: T1, *ping* 'even'; T2, *shang* 'rising'; T3, *qu* 'departing', and T4, *ru* 'entering'.

[5] These two forms, and (11a), are excluded from the set of p-words by definition, as one reviewer points out. I include them in (11) for formal completeness. By the traditional definition, (11b,c) are possible i-words. The range of i-words actually found in various sources is more restricted than what the traditional definition would allow.

[6] I benefitted from discussions with C.-L. Shih and R. Sproat on this point.

[7] At first glance, the p-word in (i) is anomalous.

(i) t'ɑu T1 – t'ɑi T3 'wash away'

The two syllables differ by their last segments: *u* versus *i*. The tones, however, differ unexpectedly. It turns out that this word is derived from *dəu T1 – dɑi T3*. The rhyme in the first syllable, *əu*, has become *ɑu* by the time of the *Guangyun*. Thus, at an earlier time, the two syllables differ by VC_2; it had the property $P_o(R')$; its tone pattern is expected.

[8] Chou (1962) argues that p-words such as (32a) are *cranberry*-type compounds (see Aronoff (1976)), as such are not genuine p-words. In *yuɐn T1 – len T1*, the second syllable is a general term referring to fragrant herbs; the first syllable, however, has no identifiable meaning, much like *cran-* in *cranberry*. The same can be said of (32b), where **dəu T1* is a general term meaning 'cicada'. However, not all words in (32) may be explained away in this fashion. The word for 'bat' in (32c), for example, consists of two characters which mean 'crab' and 'caterpillar' in isolation. The meanings are not related to the meaning of the dissyllabic word. In the discussion, I will not attempt to make a distinction between genuine p-words and potential compounds. Such a distinction must be motivated by a great deal of detective work in obscure etymology, which is beyond the scope and concern of this article.

REFERENCES

Archangeli, Diana (1985) "Yokuts Harmony: Evidence for Coplanar Representation in Nonlinear Phonology," *Linguistic Inquiry* 16, 335–372.

Aronoff, Mark (1976) *Word Formation in Generative Grammar*, MIT Press, Cambridge, Massachusetts.

Bao, Zhiming (1990a) "*Fanqie* Languages and Reduplication," *Linguistic Inquiry* 21, 317–350.

Bao, Zhiming (1990b) *On the Nature of Tone*, PhD Dissertation, MIT.

Bao, Zhiming (1993) "The Syllable in Chinese," ms., National University of Singapore.

Broselow, Ellen and John J. McCarthy (1983) "A Theory of Internal Reduplication," *Linguistic Review* 3, 25–88.

Chao, Yuen-Ren (1931) "Fanqie yu ba zhong [Eight Varieties of Languages Based on the Principle of *Fanqie*]," *Bulletin of the Institute of History and Philology*, Academic Sinica 2, 320–354.

Chen, Matthew Y. (1976) "From Middle Chinese to Modern Peking," *Journal of Chinese Linguistics* 4, 113–277.

Chen, Yan (1992) "*Guangyun* shuangsheng dieyun lianmian zi de yuyin yanjiu [A Phonetic study of the Partially Reduplicated Words in the *Guangyun*]," *Yuyanxue Renzong* [Linguistics Forum] 17, 60–121, edited by the Chinese Department, Beijing University.

Cheng, Chin-chuan (1973) *A Synchronic Phonology of Mandarin Chinese*, Mouton, De Hague.

Chou, Fa-kao (1962) *A Historical Grammar of Ancient Chinese, Part II: Morphology. (Bulletin of the Institute of History and Philology*, Special Publication No. 39, Academia Sinica, Taiwan.

Clements, George N. (1985) "The Problem of Transfer in Nonlinear Morphology," *Cornell Working Papers in Linguistics* 7.

Clements, George N. and Samuel Jay Keyser (1983) *CV Phonology*, MIT Press, Cambridge, Massachusetts.

Dong, Tong-he (1954) *Zhongguo yuyin shi* [A History of Chinese Phonology], Chinese Culture Press, Taiwan.

Duanmu, San (1990) *A Formal Study of Syllable, Tone, Stress and Domain in Chinese Languages*, PhD Dissertation, MIT.

Fu, Ding-yi (1960) *Lianmian Zidian* [A Dictionary of Dissyllabic Words], Zhonghua Press, Taibei.

Fudge, Erik (1987) "Branching Structure Within the Syllable," *Journal of Linguistics* 23, 359–377.

Goldsmith, John (1976) *Autosegmental Phonology*, PhD Dissertation, MIT.

Goldsmith, John (1990) *Autosegmental and Metrical Phonology*, Basil Blackwell, Oxford.

Guo, Xi-Liang (1986) *Hanzi guyin shouce* [A Manual of Ancient Pronunciation of Chinese Characters], Beijing University Press, Beijing.

Harris, James W. (1983) *Syllable Structure and Stress in Spanish: A Nonlinear Analysis*, MIT Press, Cambridge, Massachusetts.

Itô, Junko (1986) *Syllable Theory in Prosodic Phonology*, PhD Dissertation, University of Massachusetts, Amherst.
Ju, Chi-yung (1960) "An Annotated List of Sound-Correlated Dissyllabic Compounds in the *Book of Odes*," *Bulletin of the College of Arts, National Taiwan University* No. 9, 129–291.
Karlgren, Bernard (1948) *Zhongguo yinyun xue yanjiu* [Etudes in Phonologie Chinoise], translated from the french by Yuen-ren Chao and Fang-kui Li, Commercial Press, Taiwan.
Kenstowicz, Michael and Jerzy Rubach (1987) "The Phonology of Syllabic Nuclei in Slovak," *Language* 63, 463–497.
Levin, Juliette (1985) *A Metrical Theory of Syllabicity*, Doctoral Dissertation, MIT.
Li, Paul (1985) "A Secret Language in Taiwanese," *Journal of Chinese Linguistics* 13, 91–121.
Lin, Yen-hwei (1989) *Autosegmental Treatment of Segmental Processes in Chinese Phonology*, PhD Dissertation, University of Texas, Austin.
McCarthy, John (1989) "Linear Order in Phonological Representation," *Linguistic Inquiry* 20, 71–99.
McCarthy, John and Alan Prince (1986) "Prosodic Morphology," ms., University of Massachusetts, Amherst and Brandeis University.
Marantz, Alec (1982) "Re Reduplication," *Linguistic Inquiry* 13, 435–483.
Prince, Allan (1987) "Planes and Copying," *Linguistic Inquiry* 18, 491–510.
Pulleyblank, Douglas (1986) *Tone in Lexical Phonology*, Reidel, Dordrecht.
Pulleyblank, Edwin G. (1984) *Middle Chinese: A Study in Historical Phonology*, University of British Columbia Press, Vancouver.
Pulleyblank, Edwin G. (1991) *Lexicon of Reconstructed Pronunciation in Early Middle Chinese, Late Middle Chinese and Early Mandarin*, University of British Columbia Press, Vancouver.
Shen, Jian-shi (ed.) (1960) *Guangyun Shengxi* [Sound System of the *Guangyun*], Written Language Reform Press, Beijing.
Steriade, Donca (1982) *Greek Prosodies and the Nature of Syllabification*, PhD Dissertation, MIT.
Steriade, Donca (1986) "Locality Conditions and Feature Geometry," *NELS* 17, 595–617.
Steriade, Donca (1988) "Reduplication and Syllable Structure," *Phonology* 5, 73–155.
Ting, Pang-hsing (1979) "Shang gu yin de yinjie jiegou [Syllable Structure of Ancient Chinese]," *Bulletin of the Institute of History and Philology*, Academia Sinica 50.2, 717–739.
Wang, Guo-wei (1968) "Lian mian li pu [A List of Reduplicated Words]," in *The Complete Works of Wang Guo-wei*, Volume 9, Wenhua Press, Taiwan.
Wang, Li (1956) *Hanyu yinyun xue* [Chinese Phonology], Zhonghua Shuju, Beijing.
Wang, Li (1958) *Hanyu shigao* [A History of Chinese], Science Press, Beijing.
Wang, Li (1959) *Hanyu yuyin shi* [A History of Chinese Phonology], Chinese Social Sciences Press, Beijing.
Williams, Edwin (1971/76) "Underlying Tone in Margi and Igbo," *Linguistic Inquiry* 7, 463–484.
Yip, Moira (1982) "Reduplication and C-V Skeleta in Chinese Secret Languages," *Linguistic Inquiry* 13, 637–661.
Yu, Nai-yong (1985) *Shangku yinxi yanjiu* [A Study of the Sound System of Ancient Chinese], The Chinese University of Hong Kong Press, Hong Kong.

Received 18 March 1993
Revised 13 October 1994

Department of English Language and Literature
National University of Singapore
Singapore 0511
fasbaozm@leonis.nus.sg