Two Stress Patterns of Shanghai Compounds

Xiaonong Sean Zhu

Linguistics Department, Faculties
Australian National University

ABSTRACT - Stress in Shanghai is not uniformly left-headed, as suggested in the literature. It is determined by *tone* categories and is related to duration and F0 profile. This paper suggests that both left- and right-headed stress can exist at the same lexical level in a language.

INTRODUCTION

Stress is one of the most ignored areas in Chinese phonetics and phonology, especially in dialects other than Mandarin. Based on a formal analysis on Shanghai stress, Duanmu (1995) claims that there exists a metrical system in addition to a tonal system in the language, that the stress of Shanghai compounds is uniformly left-headed, and that it is the syllable number of the words in a compound that determines the stress pattern.

This paper will provide with acoustic data and show that stress in Shanghai is closely related to duration and F0 profiles. Specifically, the stronger syllable, which is supposed to be the head of a metrical domain, in a disyllabic compound is manifested to be longer, in a relative sense, and retains (part of) its underlying tones. I will show that stress in Shanghai is more complicated than would be expected and that Duanmu's last two claims are incorrect.

TONES AND AUDITORY IMPRESSION OF STRESS IN SHANGHAI

Stress patterns in Shanghai are predictable from *tone* categories, but not the reverse. (Italic *tone* refers to traditional citation contour such as [14], and capitalised Tone to Contour elements such as H and L.) If we describe Shanghai *tones* using the five-point scale, stress pattern is accompanying and redundant property of the *tones*. If we use Tones to represent Shanghai *tones*, stress is then necessary in describing the different tone sandhi results. There are five *tones* in Shanghai, referred to as T1 to T5 and characterised as in the left part of table 1. There are twenty-five (=5 *tones* on S1 x 5 *tones* on S2) *concatenations* for disyllabic lexical *tones*, but twenty (=5x4) *types*, with T1 and T2 neutralised (referred to as T0) on S2 (Zhu 1995: 194, 211), and five *classes* defined by the S1 *tones*, which are characterised in the right part.

Table 1	T1 [51]	T2 [34]	T3 [14]	T4 [55]	T5 [13]		T1+X	T2+X	T3+X	T4+X	T5+X
Tone	HL	LH	LH	LH	LH		H+L	L+H	L+H	L+H	L+LH
Register	Upper	Upper	Lower	Upper	Lower		Upper	Upper	Lower	Upper	Lower
Truncation	Long	Long	Long	Short	Short	Stress	s w	s w	s w	s w	w s

Stress is likely to be cued primarily in the dimension of F0, with duration also being important, and intensity playing a negligible role (Fry 1958). In Shanghai, F0 height (not shapes), the main acoustic correlate to stress in non-tonal languages, is used essentially for tonal contrast, and thus can do little to cue stress. Auditorily, Shanghai disyllabic lexical compounds are stressed on S1, except those with T5 on S1 which are stressed on S2. In the first group, S1, when carrying T1, is much more prominent than S2, moderately more prominent than S2 when bearing T2 or T3, and only marginally more prominent than S2 when bearing T4.

DURATION AND STRESS

Stress in Shanghai is constantly positively correlated to duration in relative terms due mainly to Truncation. Table 2 gives mean duration in ms of Shanghai disyllabic words and monosyllables, averaged from four speakers. '[]0' means the input tone on S2 is the pooled T1 and T2. 'T1+X' indicates a disyllabic compound with T1 on S1. For example, the first syllable with T1 is 158 ms

before T0, 187 ms before T3. Table 3 gives duration ratios of S1 and S2 to monosyllable with same tone, and S2 to S1. For example, T1 is 172 ms on S1 and 212 ms in citation; the ratio between the two is 0.81. T3 is 129 ms after another T3 and 190 ms on S1, and the ratios between the two is 0.68. The S1 duration is affected by 1) C2 voicing, 2) in part by pitch height of S1, and 3) stress, and the S2 duration is affected by 1) S2 pitch shape, 2) C2 voicing, 3) S1 Truncation, and 4) stress (Zhu 1995). In the ratios given in Table 3 all these factors have been averaged except the stress.

Table 2	T1+ X	T2+ X	T3+ X	T4+ X	T5+ X	Citation
[]0	158 73	174 121	172 122	55 139	57 169	T1: 212
[]3	187 81	191 135	202 129	77 161	65 193	T2: 240
[]4	162 39	174 51	177 55	59 54	55 63	T3: 238
[]5	180 46	202 62	211 60	78 66	68 85	T4: 71
Ave	172	185	190	67	61	T5: 92

Table 3		T1	T2	T3	T4	T5
1.	$X_{[]}/X_i$	0.81	0.77	0.80	0.94	0.66
2.	$[]X_i/X_i$	0.43	0.63	0.63	0.70	0.86
3.	$[X_{[]}X_i/X_{[]}]$	0.42	0.65	0.68	0.81	1.39

The ratios in line 3 indicate that the stress of T5+X compound is [w s] since S2 is much longer than S1 (1.39:1), and that the stress of the other four combinations is [s w] since S1 is longer. The other two lines also show T5 makes a group while the other four another.

Next consider the four [s w] combinations (T1 to T4 + X). Line 1 shows, on S1, T4 (0.94) is relatively longer than the other three (about four fifths). Line 2 shows, on S2, T1 (0.43) is relatively shorter than the other three (about two thirds). Line 3 shows both. I interpret this as there are three prominence degrees in the [s w] pattern. S1 is much stronger than S2 when carrying T1, moderately when carrying T2/T3, and only marginally when carrying T4.

The most comparable pair is T3 and T5 which differ in Truncation. Compare [T3 T3] and [T5 T5] first. The first T3 in [T3 T3] is only reduced a little with respect to citation T3 (0.85=202/238), but the final T3 is greatly reduced (0.54=129/238). Had the S1 in [T5 T5] has stress, similar proportions as in the case of T3 would have shown up. But the reverse is what we see. The first T5 is greatly reduced with respect to citation T5 (0.74=68/92), but the final T5 is only slightly reduced (0.92=85/92). Next examine [T3 T5], of which T5 is only 60 ms, comparable with the first T5 of [T5 T5] (68); both significantly shorter than the final T5 (85) of [T5 T5]. Now look at [T3 T3] and [T5 T3]. The second T3 of [T3 T3] is only almost half as long as citation T3 (0.54=129/238), but the T3 after T5 is reduced not much (0.81=193/238). On average, T3 on S1 is 80% (=190/238) as long as citation T3, but T5 on S1 is only 65% (=60/92) of citation T5. Not coincidentally, the average duration of initial T3 (190) is effectively the same as T3 after T5 (193), while the average duration of initial T5 (60) is exactly the same as T5 after T3 (60). All these indicate either Truncation or stress or both makes the difference. Below I'll show that Truncation is irrelevant to duration pattern through indirect ways.

First, the final T5 is similar after T2 (66), T3 (60), and T4 (66). They are comparable to the initial T5 (68), but differ from the final T5 (85) in [T5 T5]. Here Truncation does not matter because T2 and T3 are Long and T4 is Short, but all have the same duration patterns. Second, as with the T3~T5 pair, there is another, Upper Register, pair of T2 and T4. If Truncation matters, similar duration patterns should arise from the T2~T4 pair. After a series of examination in the same steps as done to the T3~T5 pair, we found a different picture: T2 and T4 create similar duration patterns, allowing some factors which do not exist in the T3~T5 pair. Third, there are two pairs in terms of Register: T2~T3 and T4~T5. Both T2 and T3 are Long and create similar duration patterns. Both T4 and T5 are Short and similar duration patterns are expected. However what we see is different duration patterns for T4 and T5. First, T4 on S1 (67) almost remains unchanged in duration (0.94) with respect to citation (71), while T5 is remarkably reduced on S1 (0.66=61/92). T4 (71) is significantly shorter than T5 (92) in citation ($p<0.001$), but it (67) is longer than T5 (61) when both are on S1 ($p<0.05$). In other words, S1 is strong when carrying T4, but weak when carrying T5. Second, T4 after another T4 (54) is 81% of

T4 on S2 (67). On the other hand, T5 after another T5 (85) is 1.39 times longer than T5 on S1 (61). This indicates that T4 is reduced on S2 compared with on S1, while T5 is prolonged on S2 compared with on S1. Third, compare T4 after T5 (63) with after T4 (54), and T5 after T4 (66) and after T5 (85). After T5, S2 is strong, so it is longer than that after T4 which is weak.

Having excluded Truncation, we have only stress left which is responsible for the different duration patterns for T3 and T5. I thus conclude that, from the point of view that duration is positively correlated with stress, [T5 'X] compounds have a right-headed stress while all the other compounds have left-headed stress.

F0 PROFILE AND STRESS

There are two F0 shapes for the five sandhi classes: falling T1+X and rising for the other four combinations. The falling shape is out of question, so I will consider the rising shape only. Figure 1 plots the F0 contours of T0 (a), T3 (b), T4 (c), and T5 (d) after T2 to T5: different S1 input *tones* + same S2 input *tone*. Italic and boldface respectively stand for Upper and Lower Register; u and v for unvoiced and voiced C2. Each panel is divided into two parts by the second '0 ms' line-up which designates the F0 onset of S2 *tones*. In the right part are the second-syllable F0 contours with which we are concerned, and in the left part are the first-syllable F0 contours which serve as reference for the F0 on S2.

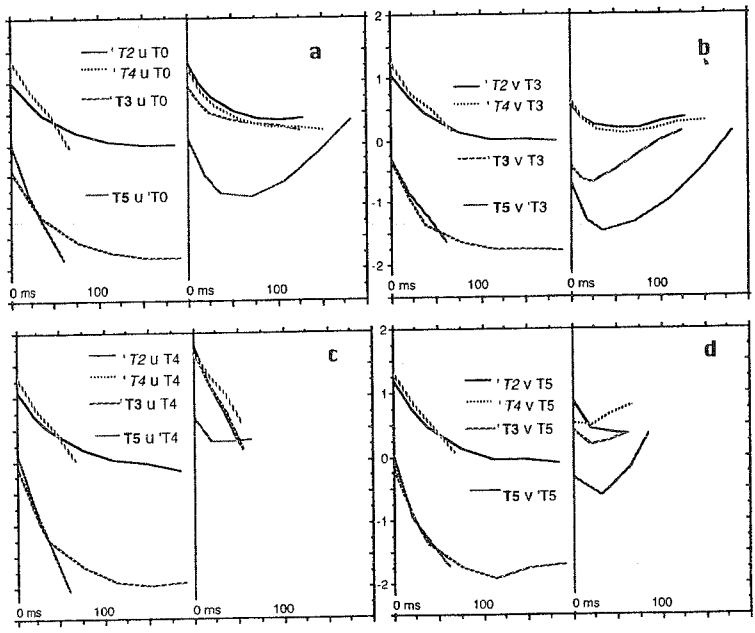


Figure 1. Second-syllable F0 contours of T0 (a), T3 (b), T4 (c), and T5 (d) after Hi-target *tones*. These F0 contours have been normalised using the Logarithmic Z-score algorithm (Zhu 1995), so the vertical axes in the figure indicate the standard deviations away from the mean in logarithmic terms.

There are two groups with respect to the F0 shapes of S2: T2/3/4+X and T5+X. The strong S2 contours after T5 are either fully rising or level while the weak S2 contours after other *tones* are basically level in panels (a), (b), and (d), and falling in panel (c).

Six possible factors might make these differences: 1) the underlying Tones of S1, 2) Register of S1, 3) Truncation of S1, 4) the offset F0 height of S1, 5) voicing of C2, and 6) stress placement. The first four can be excluded immediately. Let us look at panel a. First, the underlying Tones of S1 with T2 to T5 are all LH, so Tone does not matter. Second, the Register of S1 with T2 or T4 is Upper, and that with T3 and T5 is Lower. If S1 Register were a factor, the S2 contours would be the same after T3 and T5, and different from those after T2 and T4. But T0 after T3 is effectively the same as those after T2 and T4, so S1 Register is not a factor. Third, similarly, if S1 Truncation were a factor, T0 contours would be the same after T4 and T5, and different from those after T2 and T3. But they are not. T0 after T4 is effectively the same as those after T2 and T3, and different from that after T5. Fourth, the offset F0 height of S1 is not a factor. If it were, T0 contours would be the same after T3 and T5. But the S2 contour after T3 is the same as those after T2 and T4, and different from that after T5. Fifth, check the role of C2 voicing. Compare 'T2/4 u T0' (a) with 'T2/4 v T3' (b). The input *tones* on S1 are the same T2 or T4. The two input *tone*s on S2 lose their tonal characteristics and receive an H from S1, so T0 and T3 on S2 are irrelevant here. So the only difference is the C2 voicing: voiceless in panel (a) and voiced in (b). Since the S2 contours after T2/T4 are effectively the same in the two panels, I conclude that C2 voicing does not matter.

Having excluded the first five candidates, we have only stress left. T2, T3, and T4 make one group, and T5 another, according to the F0 contours on S2. The only difference between the two groups is that the stress is placed on S1 in the first group, but on S2 in the second.

FORMAL ANALYSIS

In this section I will show, from formal analysis, that a right-headed stress exists in Shanghai compounds. Consider the two examples (1 and 3 below) discussed in Duanmu (1995) who suggests that the contrast between the two 'provides evidence for left-headed stress'.

1 a. HL LH LH
H L L
(*tçì tshï-pong*)
x
'chicken wing'

b.
HL L H
*(*tçì*) (*tshï-pong*)
x x

2 a. LH LH LH
L H L
(*keu biq-deu*)
x
'dog's nose'

or

b.
LH L LH
(*keu*) (*biq-deu*)
x x
'dog's NOSE'

Duanmu (1995:231) states: 'if W1 is monosyllabic and W2 disyllabic, there is just one domain' as in 1. This is applicable only to the compounds in which the first syllable of W2 carries any of T1 to T4. In that case, there is a stress clash if the compound is split into two domains, see 1b. If the first syllable of W2 in the compound bears T5, there can either be one (2a) or two domains (2b). In 2b, the two domains do not create a stress clash because the stress in the T5+X combination is right-headed.

3 a. LH LH HL
L H L
(*lu- sung thòng*)
x
'Russian soup'

or

b.
L H HL
(*lu-sung*) (*thòng*)
x x
'Russian SOUP'

4 a. LH LH LH
L L LH
(*luq xo meng*)
x
six number gate
'gate number six'

but

b.
L LH LH
*(*luq xo*) (*meng*)
x x

Duanmu (p.231) states, 'If W1 is disyllabic and W2 monosyllabic, there can be either one or two domains' as in 3. This is only applicable to the compounds whose first syllable carries any of T1 to

T4. If the first syllable is in T5, there can be only one domain. If there are two domains, as in 4b, a stress clash is created because the stress of combination T5+X is right-headed.

Several conclusions can be drawn from the above discussion. First, the stress in Shanghai is not 'uniformly left-headed' as Duanmu claims. Second, it is not the syllable number of the words in a compound that determines the stress pattern, as Duanmu suggests. Rather the stress pattern of a compound with more than two syllables is determined by the *tone* of S1 in W1 or W2 in the compound. Third, although Duanmu's above two claims are wrong, a metrical analysis for the Shanghai stress is still tenable. In other words, the stress clash plays the role in determining which pattern is correct and which is wrong.

Now consider units larger than compounds in which both left-and right-headed stress arise. The following example is discussed in Zhang (1992:266).

5 a.
```
LH  LH  LH LH
LH  LH  L  H
(keu) (ngo) (lo-pe)
 x    x    x
dog  bit  boss
'the dog bit the boss'
```
or b.
```
L   H    L H
(keu ngo) (lo-pe)
 x        x
```

Both three domains (5a) and two domains (5b) are well-formed. Zhang suggests that the domain merger in 5b due to the rhythmic effect is preferable to the three domains in 5a. However, as Duanmu (1995:255) argues, the rhythmic effect cannot explain the ill-formed merger in 6b.

6 a.
```
LH LH LH  LH
L  H  LH  LH
(lo-pe) (ngo) (keu)
 x      x     x
boss   bit   dog
'the boss bit the dog'
```
but b.
```
L  H  L   H
*(lo-pe) (ngo keu)
 x       x   x
```

According to Duanmu's metrical explanation, if the verb and object in 6b merge into one domain, there is a stress clash. The stress on the object is greater than that on the verb, so it cannot be deleted. Neither can the stress on the verb since 'word stress is left-headed...Thus, the object and verb cannot merge' (p.256). In 7 he gives a metrical analysis for Ex. 6.

7
```
                x
(               x)
 x      (x      x)
(S-S)  (S)     (S)
[lo-pe [ngo    keu]]
 boss   bit    dog
'the boss bit the dog'
```

The *tone* on the verb *ngɔ* is T3 which creates a left-headed stress domain. As discussed above, not all *tones* in Shanghai create left-headed stress. If the verb carries a T5 which creates right-headed stress, and this stress can merge with the primary stress on object, then there is no metrical obstacle to prevent the merger of verb and object. This is shown in 8b an 9b. In them the verbs *luq* and *gaq* carry a T5 which makes possible the merger of verb and object.

8 a.
```
LH  LH  LH  LH
 L   H  LH  LH
(zong-he) (luq) (xü)
  x       x     x
Shanghai drop rain
'it's raining in Shanghai'
```
or b.
```
 L   H   L  LH
(zong-he) (luq xü)
 x         x
```

9 a.
```
LH LH  LH      LH  LH
 L  H  LH       L   H
(lo- pe) (gaq)  (pang-xieu)
  x      x        x
```
or b.
```
 L H   L   L    LH
(lo-pe) (gaq pang-xieu)
 x                x
```

boss squeeze friends
'the boss makes friends'

The metrical analysis of the two domains in 8b and 9b is given in 10.

10 a.

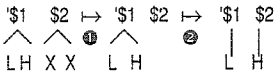
```
        x
    (   x)
 x  (   x)
(S-S) (S) (S)
[zong-he [luq xü ]]
 Shanghai drop rain
'it's raining in Shanghai'
```

b.

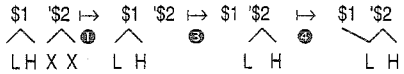
```
        x
    (   x)
 x  (      x)
(S-S) (S)  (S-S  )
[lo-pe [gaq  pang-xieu ]]
 boss squeeze friend
'the boss makes friends'
```

Below is the derivation of Tone patterns of disyllabic compounds.

a) 'T2/3/4 + X



b) T5 + 'X



❶ Delete the Tones on non-first syllable, no matters what stress status it is.
❷ Link the first Tone to S1 if it is strong, and the second Tone to weak S2.
❸ Shift the Tones to the strong S2 if S1 is weak.
❹ Reverse to link the first Tone to S1.

Metrical approach alone cannot explain the tone sandhi of T5+X compounds because it predicts that the Tones on the strong S2 should be preserved. But this is wrong; it is the S1 Tones that have been preserved and shifted onto S2. Then the question is why a metrical system is necessary in Shanghai since a metrical domain can be defined by a *tone*.

I agree to maintain a metrical system in Shanghai for the following reasons. The suggestion that omitting the metrical system and deriving the [w s] pattern from the initial T5 which is defined as [LH, Lower, Short] is at the cost of introducing a complicated system. Moreover, in doing so, we may lose an important generalisation: stress, which may play roles in other respects. For example, the phonological information of *tone* categories is carried not only by Tone, Register, and Truncation, but also by stress.

CONCLUSIONS

This paper gives support to the claim that both tonal and metrical systems can exist in the same language (Hyman & Katamba 1993, Duanmu 1995), and, moreover, that both left- and right-headed stress can exist at the same lexical level in one language. It is assumed that, if both tonal and metrical systems exist in the same language, a **tone** domain is a metrical domain (Kenstowicz 1987, Duanmu 1995). Based on the Shanghai data discussed above, we know the **tone** should be further specified: *tone* or Tone. In Shanghai the metrical domain, the *tone* domain and the Tone domain all cover a phonological word or a compound, though, a metrical domain is a *tone* domain because the *tone* on S1 of a compound predicts the stress of the compound, and it is independent of a Tone domain because neither of them can fully predict the other.

REFERENCES

Duanmu, S. (1995) Metrical and tonal phonology of compounds in two Chinese dialects. *Language* **71** (2), 225-259.

Fry, D.B. (1958) Experiments in the perception of stress. *Language and Speech* **1**, 126-152.

Hyman, L. & F. Katamba (1993) A new approach to tone in Luganda. *Language* **69**, 34-67.

Kenstowicz, M. (1987) Tone and accent in Kizigua - a Banto language. *Centamen Phonologicum*, ed by P-M. Bertinetto & M. Loporcaro, 177-88. Torito: Rosenberg & Sellier.

Zhang, H. (1992) *Topics in Chinese phrase phonology*. San Diego, CA: U of California dissertation.

Zhu, X. (1995) *Shanghai Tonetics*. Australian National University dissertation.