The main question I explore in this paper is whether there is a general theory of the phonology-syntax interface that applies to all Chinese dialects. I first review three types of tone sandhi in Chinese, those in Mandarin, Xiamen, and Shanghai. Then I review the interaction between syntax and the domain of tone sandhi in these dialects. Next I discuss the analysis of the tone-syntax interface. In particular, I compare several current approaches to foot formation in Chinese and argue for the view that foot formation depends on stress assignment. The stress-based analysis applies to all the Chinese dialects in question, as well as better-known stress languages such as English.

1. Introduction

Chinese is a tone language, in which tones are lexically specified. In general, all full syllables (or heavy/bimoraic syllables) carry a lexical tone, whereas weak syllables (or light/monomoraic syllables) do not. In addition, a full syllable can usually carry at most two tones (LH or HL); more complex tone patterns (LHL or HLH) can occur only in final (or pre-pausal) position, where the syllable is extra long. This suggests that each mora can carry at most one tone (H or L).

When two (or more) syllables occur next to each other, their tones may undergo change. This is referred to as tone sandhi, which occurs in many Chinese dialects (and languages). In addition, whether tone sandhi takes place between two syllables or not often depends on the syntactic structure. It is fairly well understood what kinds of tonal changes will take place within a given sandhi domain. However, there is less agreement on how sandhi domains are formed, even though most scholars believe that they are sensitive to syntax. The focus of this paper will be on the relation between syntax and sandhi domains.

In section 2 I review three types of tone sandhi, those in Mandarin, Xiamen (also called Southern Min or Amoy), and Shanghai. In section 3 I discuss how sandhi domains interact with syntax. In section 4 I review current analyses of tone sandhi in the three dialects. In section 5 I compare several approaches to foot formation in Chinese, a crucial process in understanding domains for tone sandhi, and I will argue for a stress-based approach. In section 6 I offer some concluding remarks.

2. Three types of tone sandhi

It is common in Chinese dialects that when syllables occur next to each other, their tones undergo change, or tone sandhi. In this section I review three types of tone sandhi. The first occurs in Mandarin, known as Tone 3 Sandhi (T3S, also called Third Tone Sandhi). T3S is discussed at length first in Cheng (1968) and subsequently in many other works. In
T3S a T3 becomes a T2 when it precedes another T3. The rule is given in (1) and exemplified in (2). Unless otherwise noted, transcription is given in Pinyin, where 0 indicates lack of lexical tone and 1-4 indicate the four full tones in Mandarin (or Standard Chinese).

(1) T3 T3 → T2 T3

(2) mai3 ma3 → mai2 ma3
buy horse       bury horse
‘to buy a horse’  ‘to bury a horse’

As far as the listener is concerned, T3S neutralizes the difference between T3 and T2 and can create ambiguities, because T2-T3 and T3-T3 are both realized as T2-T3 on the surface. For example, although mai3 ‘buy’ and mai2 ‘bury’ have different tones underlyingly, ‘to buy a horse’ and ‘to bury a horse’ sound the same to the listener.

Because T3S neutralizes the distinction between T2 and T3, it is tempting to assume that T3S is a reduction process in which the first syllable loses its original tone because it has less stress than the second (Iwata 1999, 2001, Dell 2004). However, there are some problems with this assumption. First, T3S applies regardless of which syllable has more stress. For example, in trisyllabic structures such as [zhan3-lan3 guan3] ‘exhibition hall’, there is good evidence that the first syllable has more stress than the second (Chao 1968, Hoa 1983, Yan and Lin 1988, Yang 1992), yet T3S still applies. Similarly, in a disyllabic expression, T3S applies even if the first syllable has contrastive or emphatic stress. For example, in bu shi [hao3 jiu3], shi [mi3 jiu3] ‘not GOOD wine, but RICE wine’, the underlined syllables have contrastive stress, yet they still must undergo T3S. Second, there is a difference between merger and reduction. A reduction is a change from a contrastive form to a non-contrastive form, such as from a full vowel to a schwa. On the other hand, a merger can be a change from one contrastive form to another contrastive form. Because T2 is not a neutral tone but a full tone, which contrasts with other tones, T3S is not a true reduction rule, but a conditioned merger from one contrastive form to another, similar to [sÆ∫/__i] in English (e.g. this year) or Japanese, where both /s/ and /∫/ are contrastive sounds.

It is also tempting to consider T3S as a dissimilation rule. In particular, because T3 is a low tone (in nonfinal positions), T3S seems to be a dissimilation between two low tones (L-L). On the other hand, there is no dissimilation between the H’s in T1-T1 (H-H), or between the L’s in T4-T2 (HL-LH), or between L’s in T4-T3 (HL-L). So the nature of T3S remains unclear. In any case, why T3S occurs is independent from the question of how the domains for T3S are formed. Since our main interest is in the latter, I do not pursue the nature of T3S any further.

The second type of tone sandhi occurs in Shanghai (and many other dialects in the Wu dialect family), whereby the tones of non-initial syllables are lost and the tone of the initial syllable determines the pitch contour of the entire domain. In particular, if we ignore the effect of onset voicing on the vowel, there are two syllable tones in Shanghai, LH and HL (Duanmu 1999 argues that the underlying syllable tones in Shanghai are L and H and
they surface as LH and HL respectively by a polarity rule). Their influence on the tonal pattern of the domain is abbreviated in (3), where S is a stressed syllable and s an unstressed one, and where a hyphen separates tones on different syllables.

(3) S Ss Sss+
    HL H-L H-L-L…
    LH L-H L-H-L…

The fact that the initial syllable has more stress is not always obvious to the native speaker (Selkirk and Shen 1990), but phonetic studies show that the rime of the initial syllable is nearly twice as long as that of non-initial syllables (Zhu 1995), in support of the view that the initial syllable has more stress. Some examples are shown in (4), transcribed in phonetic symbols.

(4) HL-HL $\Rightarrow$ H-L
    se + pe
    'three cups'
    HL-LH $\Rightarrow$ H-L
    se + bø
    'three plates'
    LH-HL $\Rightarrow$ L-H
    sz + pe
    'four cups'
    LH-LH $\Rightarrow$ L-H
    sz + bø
    'four plates'

It can be seen that the output tone pattern depends entirely on the first syllable, regardless of the underlying tones on the second syllable.

The third type of tone sandhi occurs in Xiamen and many other Min dialects. In this case, every full syllable has two lexically specified tone patterns: a final pattern, which is used when the syllable occurs in isolation or in the domain final position, and a nonfinal pattern, which is used when the syllable occurs in a nonfinal position. The final and nonfinal tones in Xiamen are shown in (5), according to the transcription of Chen (2000) (in Chao digits, where 1 is the lowest pitch and 5 the highest).

(5) | Categories | A | B | C | D | E | F | G |
    | Final tone (T) | 44 | 24 | 53 | 21 | 22 | 32 | 4 |
    | Nonfinal tone (T’) | 22 | 22 | 44 | 53 | 21 | 4/53 | 21 |

The nature of tonal alternations in Xiamen is not very clear. Chen (2000) suggests that the tonal alternation in (5) forms a music-chair like ‘cycle’ or ‘clock’. However, any clock one might see is probably an accident and cannot be extended to other Min dialects. In any case, such issues need not concern us. In (6) are some examples of tonal alternations, from Chen (2000: 433), transcribed in Roman letters, where A-C are final tones and A’-C’ are non-final tones. Tonal alternations are seen in the underlined words.
Tonal alternation is not limited to disyllabic expressions. For example, the compound in (7) has four syllables; three have non-final tones (T') and one has the final tone (T).

\[
\begin{array}{ll}
\text{(7)} & \\
T' & T' \quad T' \quad T \\
\text{bi-kok sio-tsia} & \\
\text{'America Miss (Miss America)'} & \\
\end{array}
\]

It is worth noting that the final tone is used for the last stressed syllable. Consider the examples in (8), where 0 refers to lack of tone on an unstressed syllable.

\[
\begin{array}{ll}
\text{(8)} & \\
T' & T' \quad T \\
\text{bue nng kun} & \text{bue nng kun} \\
\text{buy two pounds} & \text{buy two pounds} \\
\text{‘buy two pounds’} & \text{‘buy some’} \\
\end{array}
\]

When \textit{nng kun} means ‘two pounds’, the syllables are stressed, and the final tone occurs on \textit{kun}. When \textit{nng kun} means ‘some’, the syllables are unstressed, and the final tone occurs on the preceding syllable.

Strictly speaking, since each syllable is choosing between its own two tones, the alternation in Xiamen should not be called tone sandhi (i.e. the choice is determined by the position of a syllable and not by the tone pattern of a neighboring syllable). However, I follow the tradition and use tone sandhi loosely to include the Xiamen case.

I have not discussed tone sandhi in other dialects, such as Tianjin, Danyang, and Wenzhou, which are still different cases. However, those in Mandarin, Shanghai, and Xiamen offer a rich body of data on the tone-syntax interface.

3. Interaction between syntax and sandhi domains

The interaction between syntax and domains of tone sandhi is well known in the three dialects we are looking at. First, consider the tone sandhi (T3S) in Mandarin (see Duanmu 2000, chapter 11 for more details). In (9) we see that T3S is sensitive to syntactic branching. A left-branching structure usually has just one pattern, but a right-branching structure can have two or more.
(9) Left-branching

\[
[ [3 \ 3 \ 3] ] \rightarrow \ 2 \ 2 \ 3 \ only
\]

\[
[ [\text{mai} \ \text{hao} \ \text{jiu}] ]
\]

buy good wine

\‘finished buying wine\’

Right-branching

\[
[3 \ [3 \ 3]] \rightarrow \ 3 \ 2 \ 3 \ or \ 2 \ 2 \ 3
\]

\[
[\text{mai} \ [\text{hao} \ \text{jiu}] ]
\]

buy good wine

\‘to buy good wine\’

\[
[[[3 \ 3] \ 3] ] \rightarrow \ 2 \ 2 \ 2 \ 3 \ only
\]

\[
[[[\text{zhan-lan} \ \text{guan} \ \text{li}] ]
\]

exhibition hall

\‘inside of exhibition hall\’

\[
[[3 \ 3]] \rightarrow \ 3 \ \underline{2} \ 3 \ or \ 2 \ \underline{2} \ 3
\]

\[
[\text{xiao} \ [\text{zhi} \ \text{lao-hu}] ]
\]

small paper old-tiger

\‘small paper tiger\’

The difference is due to the cyclic application of T3S, moving from inner brackets to outer brackets. The analysis of left-branching structures is shown in (10), where underline shows the syllables to which T3S applies at each step.

(10) Left-branching: one output pattern

\[
[[3 \ \underline{3}]] \rightarrow [[2 \ 3]] \rightarrow 2-2-3
\]

\[
[[3 \ \underline{3} \ 3]] \rightarrow [[[2 \ 3]]] \rightarrow 2-2-3
\]

For right-branching structures, there is an additional factor, namely, a T3 before a T3 that has already changed into T2 can still change into T2 optionally (Shen 1994). This is stated in (11), where the double-underlined 2 is changed from an underlying T3.

(11) \[3-2 \rightarrow \underline{2}-2 \ (optional)\]

Two right-branching structures are analyzed in (12).

(12) Right-branching: many output pattern

\[
[3 \ [3 \ 3]] \rightarrow [3 \ [2 \ 3]] \rightarrow 3-2-3 \ or \ 2-2-3
\]

\[
[3 \ [3 \ [3 \ 3]]] \rightarrow [3 \ [3 \ [2 \ 3]]] \rightarrow 3-2-2-3 \ or \ 2-2-2-3
\]

It is also known that tree structure alone is insufficient for predicting the outcome of T3S. For example, the two expressions in (13a) have the same tree structure and input tones, but the first can become \[2 \ 3 \ 3 \ 1\] and the second cannot (* indicates a bad form). Similarly, each other pair in (13) have the same tree structure, but they do not have the same surface tone patterns (only relevant patterns are shown).
The above examples show that T3S is sensitive to not only the branching structure of syntactic trees but also additional information, either syntactic labels for the trees or information on a compound-phrase distinction, as well as information on word categories, to be discussed below.

Next consider the case in Shanghai. The example in (14) shows that tone sandhi is sensitive to syntax. The domains of tone sandhi are indicated by parenthesis.

```
(14)     LH    LH →    L-H
         (tsʰo    ve)     (tsʰo)    (ve)
      fry    rice     fry    rice
  ‘fried rice’ (noun)     ‘to fry rice’ (verb phrase)
```

A disyllabic NN (noun-noun) or AN (adjective-noun) can only form one tonal domain, where the output tonal pattern is determined by the first syllable. In contrast, a disyllabic VO (verb-object) phrase forms two tonal domains, where each syllable keeps its own tones. A non-final monosyllabic domain, such as ‘fry’ in the VO phrase, usually takes just L or H in Shanghai, instead of LH or HL. According to Duanmu (1999), the full metrical analysis of the VO phrase in (14) is (mm),[(mm).m], where ‘fry’ forms a bimoraic foot (mm) and ‘rice’ is lengthened to three moras, which are equivalent to a disyllabic foot [(mm).m]. A bimoraic foot (mm) usually takes one tone, whereas a disyllabic foot [(mm).m] can take two tones.

Finally, consider the effect of syntax in Xiamen. Some examples are shown in (15), from Chen (2000: 433-434). The domains of tone sandhi are indicated by parenthesis.
As shown, a domain may contain one, two, or three (or more) syllables. In addition, the same adjacent words ‘wind’ and ‘blow’ may belong to the same or different domains depending on the syntactic structure, to be explained below.

It is interesting to note that the sandhi domains in the three dialects are often different even for the same syntactic structure. Some examples are shown in (16)-(19), where domain boundaries are indicated by vertical lines. The words given are from Mandarin. Shanghai and Xiamen may use different words in certain cases (for example, Shanghai uses ‘eat’ instead of ‘drink’ for wine), but as long as the words have the same number of syllables, the domains boundaries remain the same. Also, not all Mandarin syllables in these examples are T3, so the domain boundaries cannot always be tested in terms of T3S. However, current researchers generally agree on the domains for T3S, and so I did not try to make up sentences that consist of T3 syllables only. Finally, tonal domains can be affected by speech rate, in that faster speech has fewer domains. The examples below show the domains for normal speech rate only.

(16) [[Old Li] [often [drink [rice wine]]]]
    [[老 李] [经常 [喝 [米 酒 ]]]]
Shanghai | Xiamen |
Mandarin |          |

(17) [[Old Li] [like [Shaoxin’s [rice wine]]]]
    [[老 李] [喜欢 [绍兴 的 [米 酒]]]]
Shanghai | Xiamen |
Mandarin |          |

(18) [Sarajevo [is-in Yugoslavia]]
    [萨拉 热窝 [在 南斯 拉夫]]
Shanghai | Xiamen |
Mandarin |          |

(19) a. fry rice (phrase) b. fry rice (compound)
炒 饭 [phrase] 炒 饭 [compound]
Shanghai | Xiamen |
Mandarin |          |
The example in (18) shows that, unlike what is suggested by Selkirk and Shen (1990), a tonal domain cannot be a prosodic word, because a polysyllabic word such as ‘Sarajevo’ or ‘Yugoslavia’ is one prosodic word, yet it can form two domains in Shanghai and Mandarin.

4. Current analyses of sandhi domains
I first discuss how each dialect has been analyzed. Then I discuss whether a general approach to the three dialects can be found.

4.1. Xiamen
The most detailed discussion of Xiamen is probably Chen (1987, 2000), according to which the sandhi domain in Xiamen is the phonological phrase (p-phrase). Adopting a suggestion from Lin (1994), Chen (2000: 459) defines the phonological phrase boundary as in (20).

(20) Xiamen p-phrase
   {Right, X_{max}^\prime}, X_{max}^\prime not lexically governed

The right-hand boundary of a phonological phrase occurs at each right-hand boundary of a syntactic phrase (the maximal projection of an element X, or X_{max}^\prime), unless the X_{max}^\prime is governed by a lexical word. The term ‘govern’ is a formal syntactic notion; it can be loosely understood as ‘being next to’. For example, in an SVO (subject-verb-object) sentence, there are four right-hand X_{max}^\prime boundaries, labeled as 1-4 in (21), where I is the Inflection node (INFL), which is null in this sentence. The X_{max}^\prime boundaries that trigger phonological phrase boundaries are underlined.

(21) [[NP]_1 I [V [NP]_2]_3]_4
    (Ting sio-tsia) (pang hong-ts\'e)
    (Ting miss) (fly kite)
    ‘Miss Ting flies a kite.’

The first right-hand X_{max}^\prime boundary is that of the subject NP; since this NP is governed by I, which is not a lexical word, the right-hand boundary counts as a phonological phrase boundary. The second right-hand X_{max}^\prime boundary is that of the object NP; however, since the object is governed by the verb (a lexical word), its boundary does not count. The third right-hand X_{max}^\prime boundary is that of the verb phrase; since verb phrase is governed by I (not a lexical word), its right-hand boundary counts as a phonological phrase boundary. The fourth right-hand X_{max}^\prime boundary is that of the sentence, which is governed by either I or C (complementizer); since neither I or C is a lexical word, the right-hand boundary of the sentence counts as a phonological phrase boundary, which happens to coincide with that of the verb phrase.

To see more examples of domain formation, consider the difference between (22) and (23), adapted from Chen (2000: 459-460). The point of interest is the adverb, which forms a domain of its own in (23) but not in (22)
(22) \[[\text{NP}] \text{I} [\text{AP} \text{V}]\]

(Ting sio-tsia) (kʰun-lat tʰak-tsʰeq)
(Ting miss) (diligent study)
‘Miss Ting studies hard.’

(23) \[[\text{NP}] \text{AP} \text{I} [\text{V} [\text{NP}]\]

(Ting sio-tsia) (kai-tsai) (tse tsi t pan ki)
(Ting miss) (fortunately) (take this one flight)
‘Fortunately, Miss Ting took this flight.’

In (22) the adverb is governed by the verb (a lexical word), so it does not form an independent domain. In (23) the adverb is supposedly governed by I (not a lexical word), so it forms an independent domain.

Chen’s analysis of Xiamen is clearly superior to that of Duanmu (1995), which claims that the sandhi domain is a right-headed metrical foot. Criticisms of Duanmu (1995) are given in Chen (2000: 436-437) and not repeated here.

There is, however, a possible alternative to Chen’s analysis. First, what seems to be at work is probably not a domain for tone sandhi, but rather the positions that trigger the final tone. Second, these positions seem to need two conditions, stated in (24).

(24) Conditions for the use of the final tone

\begin{itemize}
  \item a. The syllable should be stressed
  \item b. The syllable should be followed by a pause or an unstressed syllable
\end{itemize}

The conditions can be seen in the examples in (25), from Chen (2000: 280), where T is a final tone, T’ is a non-final tone, and 0 indicates lack of tone on an unstressed syllable.

(25) \[
\begin{array}{cccc}
\text{T’} & \text{T’} & \text{T} & \text{T} & \text{0} & \text{0} \\
\text{bue nng kun} & \text{bue nng kun} & \\
\text{buy two pounds} & \text{buy two pounds} & \\
\text{‘buy two pounds’} & \text{‘buy some’} & \\
\end{array}
\]

In the first case, all syllables are stressed, but only the final one is before a pause, hence it alone takes the final tone. In the second case, when nng kun means ‘some’, the two syllables have no stress; the initial syllable is the only one that has stress, and it is before an unstressed syllable, and thus, it takes the final tone.

The conditions in (24) may also explain why the $X_{\text{max}}$ in Chen’s definition cannot be lexically governed. In most cases, being ‘not lexically governed’ means that the $X_{\text{max}}$ is before a functional element (which is either unstressed if it is present or a pause if it is absent). Consider the difference between [Adverbial I] and [Adverbial V]. In [Adverbial I], the last syllable of the adverbial takes the final tone, because it is before a pause (if I is empty) or an unstressed syllable (if I is present). In [Adverbial V], the last syllable of the
adverbial takes the non-final tone, because it is before a lexical word (not a pause or an unstressed syllable).

Finally, let us consider two additional facts. First, in a compound made of two disyllabic nouns, familiar expressions tend to use one domain, whereas unfamiliar expressions tend to use two domains. This is consistent with the alternative analysis: a familiar expression tends to be spoken faster, and so there is usually no pause between the two nouns, and hence the final tone does not occur in the middle. In contrast, in an unfamiliar expression there tends to be a pause between the two nouns, and so the final tone can occur in the middle. The second fact is what Chen (2000: 471) calls the ‘rhythmic effect’. An example is shown in (26).

(26) ‘woman dressed up as man; nobody suspected’

\[
\begin{array}{cccccccc}
\text{woman} & \text{dress-up} & \text{male} & \text{attire} & \text{no} & \text{person} & \text{know} \\
\text{lu} & \text{ban} & \text{lam} & \text{tsong} & \text{bo} & \text{lang} & \text{gi} \\
( ) & ( ) & ( ) & ( ) & ( ) & & \\
\end{array}
\]

Regular

\[
\begin{array}{cccccccc}
\text{woman} & \text{dress-up} & \text{male} & \text{attire} & \text{no} & \text{person} & \text{know} \\
( ) & ( ) & ( ) & ( ) & ( ) & & \\
\end{array}
\]

Rhythmic

The rhythmic reading can alter the regular tonal domains probably because it can reduce pauses in the regular reading (such as the one after ‘woman’) as well as adding new pauses (such as the one after ‘dress up’).

4.2. Shanghai

There are two approaches to tonal domains in Shanghai. According to the first, a tonal domain is a phonological word (Selkirk and Shen 1990). According to the second, a tonal domain is a left-headed metrical foot (Duanmu 1997, 1999).

In the analysis of Selkirk and Shen, a phonological word starts from the left edge of a lexical word and extends until the left edge of the next lexical word. Also, for Selkirk and Shen a compound is a single lexical word. Consider the examples in (27).

(27) (tsh o\textsuperscript{h} o\textsuperscript{v}e) (ts h o\textsuperscript{v}e) (tsh i\textsuperscript{h} i\textsuperscript{v} i\textsuperscript{v} pe)(zo)

\[
\begin{array}{cccccccc}
\text{fry} & \text{rice} & \text{fry} & \text{rice} & \text{drink a cup tea} \\
\text{fry} & \text{rice} & \text{to fry rice} & \text{to drink a cup of tea} \\
\end{array}
\]

The compound ‘fried rice’ is one lexical word, and so it forms one domain. The verb phrase ‘to fry rice’ has two lexical words, and so it forms two domains. In ‘to drink a cup of tea’, ‘a’ and ‘cup’ are both thought to be non-lexical words, which do not start their own domains. Therefore there are two domains.

There are four problems with the above approach. First, it cannot explain why only lexical words (nouns, verbs, and adjectives/adverbs) can start a tonal domain but grammatical words (pronouns, classifiers, determiners, prepositions, and tense/aspect markers) cannot. Second, it cannot explain why a polysyllabic word can form two or more domains. Some examples are shown in (28).
In fast reading, all the words in (28) can form one domain. However, in careful reading a word can form two or three domains. If a tonal domain is a phonological word, all the words should form one domain only. The third problem is that a compound can also form more than one domain. Consider the examples in (29).

(28) (ko-r-fu) (ŋi-kʰa'2-sō) 'Golf'
(ʔo-ta)-(li-ja) (ʔo-ta-li-ja) 'Nixon'
(nø-sz)-(la-fu) (nø-sz-la-fu) 'Yugoslavia'
(ku-r-)(pa-tɕo-fu) (ku-r-pa-tɕo-fu) 'Gorbachev'
(ʔe-se²)-(tɕu-pi-ja) (ʔe-se²-tɕu-pi-ja) 'Ethiopia'
(dʑe²-kʰa'2)-(sz-lu)-(va²-ka2) (dʑe²-kʰa'2-sz-lu-va²-ka2) 'Czechoslovakia'

In fast reading, all the words in (28) can form one domain. However, in careful reading a word can form two or three domains. If a tonal domain is a phonological word, all the words should form one domain only. The third problem is that a compound can also form more than one domain. Consider the examples in (29).

(29) (sâ ɦo²-y) *(sâ)(ɦo²-y)
business school 'business school'
(ŋy-ji ɦo²-y) or (ŋy-ji)(ɦo²-y)
language school 'language school'

If the expressions are compounds, they should both form one domain only, yet the second can form two domains. If the expressions are not compounds, they should both form two domains, yet the first cannot form two domains. The fourth problem is that contrastive stress can create a new domain. Consider the examples in (30).

(30) Normal Contrastive
(lo-fu tɕ) (lo-fu)(tɕ)
'tiger head' 'tiger HEAD'

In normal reading, the compound ‘tiger head’ forms one domain. When there is contrastive stress on ‘head’ (as in ‘not tiger TAIL, but tiger HEAD’), it starts a new domain. If tonal domains are unrelated to stress, as Selkirk and Shen assume, additional assumptions are needed to explain why contrastive stress can start a new tonal domain.

Let us now consider the analysis that a tonal domain in Shanghai is a left-headed foot, based on Duanmu (1999). This analysis explains a number of things right away. First, it explains why it is the first syllable whose tone is kept: there is a general rule that in all Chinese dialects only stressed syllables can keep their underlying tones. Second, it agrees with the phonetic fact that the domain-initial syllable is phonetically much longer than non-initial syllables (Zhu 1995). Third, it explains why polysyllabic words can form disyllabic domains from left to right: a common process of foot formation is to count syllables, rather than words. Fourth, it explains why contrastive stress can start a new domain: since a domain is a left-headed foot, every stress starts a new domain. Finally, it
explains why grammatical words do not start a tonal domain. The reason is that grammatical words are not assigned stress, owing to independently motivated stress rules. In particular, in Duanmu’s analysis, Shanghai has the same rule as English for compounds and phrases, which is called Nonhead Stress, stated in (31).

(31) Nonhead Stress: A syntactic nonhead is stressed
Examples: In [X N] compounds, X is stressed
In [V O] phrases, O is stressed

Grammatical words are unstressed because they are mostly syntactic heads; for example, a classifier is the head of a classifier phrase. In addition, there is a rule for assigning stress to polysyllabic words and other free syllables, given in (32).

(32) Left-to-right foot formation (for polysyllabic words and free syllables):
Form syllabic trochees from left to right.

Now consider the examples in (33), where Ø is an empty syllable (Burling 1966, Liberman 1975, Giegerich 1985, Burzio 1994, and Hayes 1995).

(33) x x (x) x
(ts^h_0 ve) [ts^h_0] (ve Ø) (tci^2 i^2 pe)(zo Ø)
  fry  rice fry  rice drink a cup tea
‘fried rice’ ‘to fry rice’ ‘to drink a cup of tea’

In the compound ‘fried rice’, Nonhead Stress is assigned to the first word, so there is one domain. In the phrase [V O] ‘to fry rice’, Nonhead Stress is assigned to ‘rice’, which can form a binary foot with an empty syllable (a pause). The first word ‘fry’ will form another kind of domain by itself—a bimoraic foot as a heavy syllable (instead of a disyllabic foot); its tone will also be somewhat shorter than that on ‘rice’. In ‘to drink a cup of tea’, ‘a’ is the syntactic head of the determiner phrase and ‘cup’ is the syntactic head of the classifier phrase, so neither gets stress. In contrast, ‘tea’ is the syntactic nonhead of the classifier phrase, so it has stress. The verb ‘drink’ is not assigned stress either; however, the first three syllables can form a foot by the left-to-right foot formation rule.

Finally, compare (34) and (35). In (34), ‘a cup’ does not form its own domain when preceded by a verb, but does when it is in initial position. In contrast, ‘one cup’ always forms its own domain.

(34) (x) x (x) x
(tci^2 i^2 pe)(zo Ø) (i^2 pe)(zo Ø)
[drink [[a cup] tea]] [[a cup] tea]
‘to drink a cup of tea’ ‘a cup of tea’

(35) x x x x
(tci^2 i^2 pe)(zo Ø) (i^2 pe)(zo Ø)
In the present analysis, ‘a cup’ forms a domain via the left-to-right foot formation process. In contrast, ‘one’ is the nonhead of the classifier ‘cup’, so ‘one’ has stress and ‘one cup’ always stars a domain.

I am not aware of a better alternative to the stress-based analysis of Shanghai. This analysis is also adopted by Chen (2000) and is consistent with Yue-Hashimoto’s (1987) classification of Shanghai as an initial-dominance language.

### 4.3. Mandarin

There seems to be little disagreement that some kind of domains must be formed before the T3S rule applies in Mandarin. For example, most analyses (e.g. Shih 1986, Chen 2000, Duanmu 2000, Feng 2004, Wang 2004) seem to agree with the domain boundaries in the following examples, where domain boundaries are indicated with a vertical line ‘|’.

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ma3yi3</td>
<td>2 3  ‘ant’</td>
</tr>
<tr>
<td></td>
<td>mi3-jiu3</td>
<td>2 3  ‘rice-wine’</td>
</tr>
<tr>
<td></td>
<td>ni3 hao3</td>
<td>2 3  ‘you good (How are you?)’</td>
</tr>
</tbody>
</table>

#### (36) Left-branching vs. Right-branching

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>[[ [ 3 3 ] 3 ] 3 ] → 2 2 2 3 only</td>
<td>[3 [3 [3 3 ]]] → 2 3 2 3, 3 2 2 3, or 2 2 2 3</td>
<td></td>
</tr>
<tr>
<td>exhibition hall inside</td>
<td>small paper old-tiger</td>
<td></td>
</tr>
<tr>
<td>‘inside of exhibition hall’</td>
<td>‘small paper tiger’</td>
<td></td>
</tr>
</tbody>
</table>

#### (38) Two Domains vs. One Domain

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I want buy book</td>
<td>small paper old-eagle</td>
<td></td>
</tr>
<tr>
<td>‘I want to buy a book’</td>
<td>‘small paper eagle’</td>
<td></td>
</tr>
<tr>
<td>[gou [ [bi</td>
<td>ma] xiao ] ]</td>
<td>[gou [ [hen hao] yang] ]</td>
</tr>
<tr>
<td>dog than horse small</td>
<td>dog very good raise</td>
<td></td>
</tr>
<tr>
<td>‘Dogs are smaller than horses’</td>
<td>‘Dogs are very easy to raise’</td>
<td></td>
</tr>
</tbody>
</table>

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>wu-wu-</td>
<td>wu-wu</td>
<td>qi-wu-</td>
</tr>
<tr>
<td>‘five-five-five-five’</td>
<td>‘seven-five-five-seven’</td>
<td></td>
</tr>
</tbody>
</table>

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>xiang mai</td>
<td>gu-piao</td>
<td>xiang MAI gu-piao</td>
</tr>
<tr>
<td>want buy stock</td>
<td>want buy stock</td>
<td></td>
</tr>
<tr>
<td>‘want to buy stocks’</td>
<td>‘want to BUY stocks’</td>
<td></td>
</tr>
</tbody>
</table>
In (36) we see that any disyllabic structure can form a domain, since T3S applies to all of them. In (37) we see that even though all the structures form one domain, alternative patterns are possible in some but not in others; this indicates that cyclic application of T3S is needed within a domain (Shen 1994, Chen 2000, Duanmu 2000). In (38) we see that expressions that have the same branching structures can form one or two domains, and therefore additional information (such as syntactic labels or word categories) is needed to account for the difference. In (39) we see that flat structures can also form multiple domains. Finally in (40) we see that emphasis or contrastive stress can affect domain boundaries.

The domain boundary is primarily determined on the basis of whether T3S is obligatory. In particular, T3S must apply to adjacent 3-3 within a domain, but is optional for adjacent 3-3 between two domains. For example, consider the examples in (39), repeated in (41).

(41) a. | wu3-wu3- | wu3-wu3 | \( \rightarrow \) \( 2 3 2 3 \) (*3 2 2 3)
   ‘five-five-five-five’

b. | qi1-wu3- | wu3-qi1 | \( \rightarrow \) \( 1 3 3 1 \) (or 1 2 3 1)
   ‘seven-five-five-seven’

In (41b), the medial 3-3 need not change, which indicates that they are in separate domains. This is confirmed in (41a), where the first and third T3 must change, which indicates that the initial 3-3 are in one domain, so are the final 3-3. Similarly, consider (40), repeated in (42).

(42) a. No emphasis
   | xiang3 mai3 | gu3-piao4 |
   want buy stock
   \( \rightarrow \) \( 2 3 3 4 \) (or 2 2 3 4)
   ‘want to buy stocks’

b. Emphasis on ‘buy’
   | xiang3 MAI3 gu3-piao4 |
   want buy stock
   \( \rightarrow \) \( 3 2 3 4 \) (*2 3 3 4)
   ‘want to BUY stocks’

In (42a), the medial 3-3 need not change, which indicates that they are in separate domains. In (42b), the medial 3-3 must change (i.e. the tone on ‘buy’ must change), which indicates that they are in the same domain. In addition, a single syllable does not form its own domain (unless it has phrasal stress and is in final position); therefore all the syllables in (42b) are in the same domain. (42b) also shows that, unlike what is suggested in some studies (e.g. Iwata 1999, 2001, Dell 2004), the syllable that changes its tone under T3S cannot be thought of as having less stress; in (42b) ‘buy’ has emphatic stress, yet it still must change its tone.
While most analyses agree on where domain boundaries should be, there are different views on how the boundaries are determined. I discuss this in the next section.

5. Foot formation in Chinese
Most analyses agree that T3S in Mandarin is sensitive to foot formation. However, opinions are divided on how feet are formed in Chinese. In addition, there is the question of whether T3S domains are identical to foot domains. In this section I argue for an approach in which foot formation is based on stress assignment and T3S domains are not always identical to foot domains. I also compare the stress-based analysis with three other approaches, those of Chen (2000), Wang (2004), and Feng (2004).

5.1. The stress-based analysis
The analysis I offer is a revised version of what is proposed in Duanmu (2000). The main claim is that, while languages can differ in word stress, all languages have the same rule for compound and phrasal stress. I present a general theory of stress first. Then I discuss foot formation.

5.1.1. A general theory of stress
The idea that word stress can be variable is not controversial. For example, English has many lexically idiosyncratic stress patterns, such as Canada-banana, Alabama-America, and alpine-sardine. However, the claim that all languages have the same rule for compound and phrasal stress may seem counter-intuitive; for one thing, many languages are often thought to have no stress. For example, in Hyman’s (1977) survey of 444 languages, 113 are thought to have no stress, including Chinese. So what is the reason to think that all languages have stress, and that they all have the same rules for compound and phrasal stress? Let us begin with contrastive stress. I assume that all languages have contrastive stress, even in those that do not seem to have lexical stress. An example in Chinese is shown in (43), where contrastive stress is shown in uppercase.

(43) Wo xing HUANG, bu xing WANG
I name HUANG, not name WANG
‘I (am) named HUANG, not named WANG.’

I assume further that in all languages the word under contrast has extra stress. But why is it so? Apparently, the word under contrast carries more information of interest. In other words, we give more stress to words with more information and less stress to words with less information. This may seem common sense, but I will show that it explains many well-known stress effects. So let us state the principle explicitly in (44).

(44) The Information-Stress Principle:
A word or phrase that carries more information than its neighbor(s) should be stressed.
The Information-Stress Principle explains two well-know facts right away. First, pronouns usually do not carry stress, because pronouns are used for entities that are either obvious or have already been mentioned, and so they usually do not carry much information. Second, as pointed out by Hayes (1995: 373) and many others, phrasal stress is often quite flexible. In the present analysis, it is because the information load of a word is not fixed but is dependent on the context, including the speaker and the listener.

Next consider how information load is determined. Following Information Theory (Shannon 1948), the information load of a form is defined in terms of its probability of occurrence, given in (45).

(45) Information load:
    The more predictable a form is, the less information it carries.

Consider a simple phrase [Art N] (article-noun), such as the car. In English, there are only two choices for an article before a singular noun, the or a. Each has a probability of $\frac{1}{2}$ in this environment. In contrast, English has thousands of nouns, and so each has a very small probability of occurring after an article. Therefore, in this structure the article carries far less information than the noun. Not surprisingly, if we omit the article, there is not much loss of information, but if we omit the noun, the information loss is more serious. Similarly, consider the difference between prepositions and verbs. Suppose that a language has six prepositions and 1000 verbs (and suppose that we can determine the category of a word through syntactic context). Their information load is analyzed in (46).

(46) | P    | V    |
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>1,000</td>
</tr>
<tr>
<td>$\frac{1}{6}$</td>
<td>$\frac{1}{1,000}$</td>
</tr>
<tr>
<td>low</td>
<td>high</td>
</tr>
</tbody>
</table>

In this case, the information load of an average verb far exceeds that of an average preposition. The example explains a well-known stress difference between functional (or closed-class) words and lexical (or open-class) words. A functional word is usually unstressed, because its information load is low. In contrast, the information load of a lexical word is high, and so it is usually stressed.

The Information-Stress Principle can also explain a well-known correlation between frequency and reduction, namely, frequent word tend to undergo stress loss and vowel reduction more often than infrequent words (Bybee 2001). For example, the second syllable is unstressed and reduced in information but not in importation, because information is a frequent word while importation is not. Similarly, the word trombone is an infrequent word for non-musicians, and so the first vowel is stressed and unreduced; for trombonists, on the other hand, the word is frequent and so the first vowel is unstressed and reduced. In the present analysis, a frequent word has less information, and so it tends to have less stress and is more likely to undergo reduction.

The Information-Stress Principle also points to a default rule for compound and phrasal stress. Consider phrasal stress first. A typical phrase has the structure [X YP] (or
In standard X-bar syntax, the head is an element at the word (or affix) level, and the nonhead is an element at the phrase level. Since there are more possible phrases than possible words (or affixes), the occurrence of a nonhead (phrase) is less predictable than the occurrence of a head (word or affix). For example, consider a hypothetical language that has 1000 verbs and 1000 nouns. In a verb phrase [V NP], supposing the NP is made of up to two words (say an [N N] compound), the probabilities of an average V and an average NP can be calculated and are shown in (47).

\[
\begin{array}{ccc}
\text{V} & \text{NP (up to [N N] compounds)} \\
\hline
\text{Number of forms} & 1,000 & 1,000 \times 1,000 \\
\text{Probability of each form} & 1/1,000 & 1/1,000,000 \\
\end{array}
\]

Clearly, the probability of an average NP is far smaller than that of an average V, and so the information load of an average NP is far greater than an average V (in a real language the difference is even greater, because a real language typically has more nouns than verbs, and an NP can contain more than two words). By the Information-Stress Principle, therefore, the default phrasal stress should go to NP, which is the case in English (and in Chinese, to be exemplified below). By the same argument, in any phrase [X YP], phrasal stress should go to YP.

The analysis of compound stress is similar. In English, an [N1 N2] compound is parallel to a [YP X] phrase, where N1 is the counter-part of YP and N2 the counter-part of X. Therefore, N1 should receive compound stress, which is true in typical [N N] compounds in English.

The phrasal stress rule and compound stress rule are originally given as two separate rules in Chomsky and Halle (1968), which are also assumed in Halle and Vergnaud (1987). In the present analysis, they are related, namely, in both cases the syntactic nonhead is assigned stress. The unified rule is called Nonhead Stress, first proposed in Duanmu (1990) and repeated in (48), which ultimately derives from the Information-Stress Principle.

\[
\text{(48) Nonhead Stress:} \\
\text{In the syntactic structure [X YP (or [YP X])], where X is the syntactic head and YP the syntactic nonhead, YP should be stressed.}
\]

In (49) I compare Chomsky and Halle’s analysis with the Nonhead Stress analysis. For clarity syntactic heads are underlined and relative stress among stressed elements is ignored.

\[
\begin{array}{ccc}
\text{(49) Syntax} & \text{Nonhead Stress} & \text{Chomsky and Halle} \\
[V \ N] & \text{buy CARS} & \text{(same)} \\
[P \ N] & \text{in SCHOOL} & \text{(same)} \\
[N \ N] & \text{WRIST-watch} & \text{(same)} \\
[N [N N]] & \text{GOLD WRIST-watch} & \text{(same)} \\
[[N N] N] & \text{WRIST-watch store} & \text{WRIST-watch STORE}
\end{array}
\]
In some cases, such as [N [V N]] and [N [V N]], Chomsky and Halle assign stress to more than one word. In the first seven structures, the syntactic heads are as traditionally understood. In [D N] the syntactic head is D, following Abney (1987). In [N’s N] the syntactic head is the possessive {s}. The noun phrase [A N] is [A (F) N], where the syntactic head is the inflectional element {F}. This follows from the view that noun phrases are headed by a functional element, not always present in English but required in some other languages (Ritter 1991, Cinque 1993). For example, the Chinese counter-part to an English [A N] phrase is [A de N], where the head de is a functional element (Dai 1992, Duanmu 2000).

In most cases the analyses predict similar results. Where they differ, the Nonhead Stress analysis seems to be more accurate. For example, in [N’s N] and [A N], both words have stress in the Nonhead Stress analysis, in agreement with the traditional judgment (Kenyon and Knott 1944, Jones 1950). Similarly, as Hayes (1995: 373-382) argues, in structures like [[N N] N] and [[[N N] N] N], no stress should be assigned after the main stress (here the first N). Finally, the present analysis agrees with the observation that verbs are less likely to be stressed than nouns (Ladd 1980: 90-92, Hayes 1995: 376). This is because (a) verbs often occur as syntactic heads, which do not receive Nonhead Stress, and (b) there are usually more nouns than verbs in a language, and so an average noun carries more information than an average verb.

The idea that there might be a universal rule for compound and phrasal stress has been proposed in different forms by Cinque (1993) and Zubizarreta (1998). According to them, the degree of stress is related to the depth of a syntactic tree: the deeper a branch, the more stress it has. Since a syntactic head generally does not branch, and the syntactic nonhead is a phrase, which can always branches, the syntactic nonhead is generally deeper and has more stress than the syntactic head. Thus, their theories predict similar stress patterns as Nonhead Stress.

5.1.2. Stress in Chinese

As in English, word stress in Chinese is unpredictable. For example, Mandarin has stressed monosyllabic words, such as ma3 ‘horse’, and unstressed monosyllabic words, such as the progressive marker zhe0 ‘be doing’. Stress in disyllabic words is also unpredictable. For example, some Mandarin words are heavy-heavy (both stressed), such as fan1-qie2 ‘tomato’ and ku3-gua1 ‘bitter melon’, whereas others are heavy-light (stressed-unstressed), such as luo2-bo0 ‘turnip’ and huang2-gua0 ‘yellow melon (cucumber)’. Indeed, there are minimal pairs that are distinguished by stress alone, such as those in (50). The vowel difference in the monosyllabic pair is due to reduction in the unstressed word.
Minimal pairs contrasted by stress

Monosyllabic

<table>
<thead>
<tr>
<th>Heavy</th>
<th>Light</th>
</tr>
</thead>
<tbody>
<tr>
<td>zhao2 着</td>
<td>zhe0 着</td>
</tr>
<tr>
<td>‘contact’</td>
<td>‘be doing’</td>
</tr>
</tbody>
</table>

Disyllabic

<table>
<thead>
<tr>
<th>Heavy-Heavy</th>
<th>Heavy-Light</th>
</tr>
</thead>
<tbody>
<tr>
<td>da4-yi4 大意</td>
<td>da4-yi0 大意</td>
</tr>
<tr>
<td>‘main idea’</td>
<td>‘careless’</td>
</tr>
<tr>
<td>bao1-han2 包含</td>
<td>bao1-han0 包涵</td>
</tr>
<tr>
<td>‘include’</td>
<td>‘forgiving’</td>
</tr>
</tbody>
</table>

Moreover, word stress can vary from speaker to speaker, or from dialect to dialect. For example, As Li (1981: 37) points out, ‘cotton’ can be heard as mian2-hua1 (heavy-heavy) or mian2-hua0 (heavy-light) from Mandarin newsreaders. Similarly, xil-gua0 ‘watermelon’ is heavy-light in Beijing Mandarin, but xil-gua1 or heavy-heavy in Taiwanese Mandarin. (Some studies, such as Hoa 1983, further distinguish two kinds of heavy-heavy patterns: medium-heavy vs. heavy-medium. However, I agree with Chao 1968 that such a distinction is hard to verify.)

One might argue that the difference between the word pairs in (50) is not due to stress but due to the presence or absence of tone. If so, Chinese does not have lexical stress. This proposal has two shortcomings. First, it must explain why toneless syllables are phonetically unstressed (see below). Second, stress must be assumed for Chinese anyway, if not for word stress, at least for phrasal stress. So there is little theoretical gain in not recognizing word stress.

The phonetic difference between stressed and unstressed syllables in Chinese is very clear to native speakers. Indeed, the phonetic cues are similar to those in English (Fry 1955, 1958). Consider the examples in (51) and (52).

(51) English heavy-light

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Peter</td>
<td></td>
</tr>
<tr>
<td>Length</td>
<td>long</td>
</tr>
<tr>
<td>Reduction</td>
<td>no</td>
</tr>
<tr>
<td>Tone</td>
<td>yes</td>
</tr>
</tbody>
</table>

(52) Mandarin heavy-light (Pinyin: ba4-ba0, via reduplication of ba4)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>paa- po</td>
<td>‘dad’</td>
</tr>
<tr>
<td>Length</td>
<td>long</td>
</tr>
<tr>
<td>Reduction</td>
<td>no</td>
</tr>
<tr>
<td>Tone</td>
<td>yes</td>
</tr>
</tbody>
</table>

In English, the heavy syllable is longer, has an unreduced vowel, and a tone (what is called ‘pitch accent’ in intonational phonology; Pierrehumbert 1980). In contrast, the weak syllable is short, has a reduced vowel, and no pitch accent. Similarly, in Chinese, the heavy
syllable is longer, has an unreduced vowel, and a lexical tone, whereas the weak syllable is short, has a reduced vowel, and no lexical tone (Lin and Yan 1988). Thus, in both languages the difference between the two syllables is clear.

If the manifestation of stress is similar in English and Chinese, why then is there a popular view that Chinese has no stress? There seems to be three reasons. First, many words in Chinese are monosyllabic, and stress is usually more obvious in a polysyllabic word (i.e. when stressed and unstressed syllables occur next to each other). Second, minimal pairs in which stress is contrastive are not many. According to Li (1981: 37), there are some 30,000 disyllabic words (including compounds) in a large Mandarin dictionary, and about 2,000 are clearly heavy-light. Of those about 100 pairs are contrasted by stress, and only half of those show ‘large semantic differences’ between each pair (for example, ‘wife/woman’ probably shows a small semantic difference and ‘include/forgiving’ probably shows a large semantic difference). In other words, stress contrast plays only a minor role in the Chinese lexicon. The third reason is that the stress difference between the two syllables in a heavy-heavy word is often not as obvious as that in English. Consider the examples in (53) and (54).

(53) English heavy-heavy: stress is clear

<table>
<thead>
<tr>
<th></th>
<th>meat</th>
<th>ball</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length</td>
<td>long</td>
<td>long</td>
</tr>
<tr>
<td>Reduction</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Tone</td>
<td>yes</td>
<td>no</td>
</tr>
</tbody>
</table>

(54) Mandarin heavy-heavy: stress is less clear

<table>
<thead>
<tr>
<th></th>
<th>rou4</th>
<th>wan2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length</td>
<td>long</td>
<td>long</td>
</tr>
<tr>
<td>Reduction</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Tone</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>

In English, both syllables are long and have unreduced vowels. However, the first syllable also has a pitch accent, whereas the second does not. Thus, the stress difference is still quite clear. In Mandarin, both syllables are long and have unreduced vowels. In addition, both have lexical tones. Thus, the stress difference is much less obvious than that in English. It is interesting to compare the Chinese example in (54) with the English example in (55).

(55) English heavy-heavy: stress is less clear

<table>
<thead>
<tr>
<th></th>
<th>Red</th>
<th>Cross</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length</td>
<td>long</td>
<td>long</td>
</tr>
<tr>
<td>Reduction</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Tone</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>
In this case both words have pitch accents, and so their stress difference is less obvious than that in meat ball. According to Kenyon and Knott (1944) and Jones (1950), in expressions like Red Cross and real deal, the two words have equal stress, which is similar to the case of heavy-heavy words in Mandarin.

There is a further similarity between English and Chinese that is worth pointing out. Some scholars, such as Chomsky and Halle (1968), consider the second syllable in Red Cross to have more stress when the expression occurs in final position, but not when it occurs in non-final position. This is shown in (56).

$$\begin{array}{cccccc}
\text{x} & \text{x} & \text{x} & \text{x} \\
\text{x} & \text{x} & \text{x} & \text{x} \\
\text{Red Cross} & \text{Red Cross office}
\end{array}$$

Interestingly, the same proposal has been made for Mandarin (Chao 1968, Xu 1982, and Hoa 1983), which is illustrated in (57).

$$\begin{array}{cccccc}
\text{x} & \text{x} & \text{x} & \text{x} \\
\text{x} & \text{x} & \text{x} & \text{x} \\
\text{rou4} & \text{wan2} & \text{rou4} & \text{wan2} & \text{tang1} \\
\text{‘meat ball’} & \text{‘meat ball soup’}
\end{array}$$

According to Hoa (1983), examples like (57) involve a rule that reverses the stress in a syntactic unit that is not final. In particular, Hoa assumes that compound and phrasal stress in Chinese is right-headed, or MS (medium-strong), following Chao (1968). The underlying stress for ‘meat ball soup’ is [[MS] S]. However, a nonfinal syntactic unit must undergo stress reversal, so that [[MS] S] becomes [[SM] S], where the final S is stronger than the initial S. Similarly, in Beijing Daxue ‘Peking University’, [[MS][MS] becomes [[SM][MS]]. One might suggest that perhaps stress reversal is unnecessary; all we need is a template MS or MWS (medium-weak-strong), regardless of inner syntax. Thus, ‘meat ball’ is MS, ‘meat ball soup’ is MWS, and ‘Peking University’ is MWWS. However, as Hoa (1983) argues, inner syntax is clearly relevant. For example, Beijing Dongwuyuan ‘Peking Zoo’ is [MW MWS], not [MW WWS] (unless in fast speech, perhaps), and dongwuyuan damen ‘zoo entrance’ is [MWW MS], not [MWW WS].

It is obvious that most people agree that stress is left-headed except for the final syllable. I suggest that the perceived extra stress on the final syllable is due to the pause that follows (and perhaps also to the boundary tone), which can lengthen the final syllable (Klatt 1975, 1976 for English; Lin et al 1984, Yan and Lin 1988, Yang 1992 for Chinese). It is worth noting that in Chinese, although the final syllable has longer duration, the first syllable has greater amplitude and higher F0 peak (Lin et al 1984, Yan and Lin 1988, Yang 1992). In view of such considerations, I analyze (57) in (58).
In both expressions, compound stress goes to the initial syllable (the syntactic nonhead), and so there is one syllabic foot, shown in parentheses. In addition, since each syllable is heavy, it is a moraic trochee and has its own stress (this is the ‘dual-trochee’ analysis introduced in Duanmu 1999). The final syllable is longer (indicated as x+) because of the following pause Ø. Now x+ may be perceived as having more stress than the initial syllable, even though the initial syllable has greater amplitude and F0 peak, but the difference is small. On the other hand, both the initial syllable and the final syllable should be perceived as being stronger than the medial syllable.

There is a popular assumption that stress (or accent) is an abstract notion that can be realized differently in different languages (e.g. Hayes 1995:5, Kaji 1999: 217). The above discussion shows that, as far as English and Chinese are concerned, there is no evidence for this assumption, because there is no fundamental difference in the phonetic realization of stress.

5.1.3. Foot formation and T3S
Let us now consider the formation of syllabic feet in Chinese (ignoring moraic feet for now) and the application of T3S. I assume the analysis in (59) and (60).

(59) Syllabic foot formation in Chinese:
In compounds and phrases, assign Nonhead Stress.
A foot has at least two syllables, with stress on the first syllable.
For free syllables and flat structures, build binary feet from left to right.

(60) Application of T3S:
Apply T3S cyclically from within each foot and upward.
T3S is optional between two cyclic domains.

The analysis predicts that a disyllabic word or compound forms a foot, and a polysyllabic word or a sequence of digits form binary feet from left to right. Next consider a [V O] phrase, shown in (61), where parentheses indicate foot boundaries and vertical lines indicate the cyclic domain of T3S.

(61)  x  x  x  
( x  x+ ) ( x  x  x+ )  
| mai3 (hao3 shu1) | 2 3 1  
[ buy [good book]]  
‘to buy good books’

In the object, Nonhead Stress goes to ‘good’. In [V O], Nonhead Stress goes to the object, which already has stress. Now consider T3S, which starts from within the foot; since
‘book’ is not T3, T3S has no effect. Next, T3S applies to the next node up, which is between ‘buy’ and ‘good’, giving [2 3]. The analysis shows that for T3S to apply, the two syllables need not be in the same foot. Instead, all three syllables are in the domain of T3S, even though the first is outside the foot. Next consider the example in (62), where | is a cyclic domain boundary.

(62)  x        x
(wo3  xiang3) | mai3  (hao3  jiu3)  \rightarrow  2 3 3 2 3 (or 2 2 3 2 3)
[I  [want    [buy  [good wine]]]]
‘I want to buy good wine’

In ‘good wine’, Nonhead Stress goes to ‘good’. In ‘buy good wine’, Nonhead Stress goes to ‘good wine’, which already has stress. In ‘want buy good wine’, Nonhead Stress goes to ‘buy good wine’, which is already stressed. The subject ‘I’ is a pronoun, which may or may not receive Nonhead Stress. If it does not, it can still form a foot with the next syllable, via the left-to-right foot formation rule. Next consider T3S. The first foot forms its own domain. For the second foot, T3S produces 2-3 in the first step; the next step up is between ‘buy’ and ‘good book’, where T3S need not apply, because ‘hao’ has already changed to T2. This ends the second domain. Now both ‘want’ and ‘buy’ are still T3, but they are in different domains. Therefore, T3S occurs optionally between them. Next consider the compounds in (63) and (64).

(63)  x
| (zhan-lan  guan  li) |
[[exhibition  hall] inside]
‘inside of exhibition hall’

(64)  x  x  x  x
xiao  zhi  lao-hu  \rightarrow  | (xiao  zhi  lao-hu) |
[small  [paper  [old-tiger]]]
‘small paper tiger’

In (63) the word ‘exhibition’ has initial stress. On the next two cycles, the syntactic nonhead is on the left, so stress should be on the left, which is already true. The result is one foot, which coincides with a cyclic T3S domain. In (64), the compound ‘old tiger’ first gets initial stress. On the next two cycles, the syntactic nonhead is on the left, so stress should be assigned to the left. This gives three adjacent stresses. However, since a foot needs at least two syllables, the three stresses cannot all remain. I assume that in this case only the initial stress is kept, because it is the ultimate syntactic nonhead. It is worth noting that in (63) and (64) there are no free syllables. Next consider the phrases in (65) and (66).
In (65), the innermost structure is ‘than horse’, where Nonhead Stress goes to ‘horse’. In the VP, V is the syntactic head, so Nonhead Stress goes to the adverbial, which is already stressed. Finally, the subject is not the syntactic head of the sentence, so it has stress, too. The two stresses create two feet. In (66), the innermost structure is ‘very good’, where Nonhead Stress goes to ‘very’. In the VP, V is the syntactic head, so Nonhead Stress goes to the adverbial, which is already stressed. Finally, the subject is not the syntactic head of the sentence, so it has stress, too. Now ‘dog’ is a monosyllable and cannot form a foot by itself. It can either lose its stress, or it can keep its stress and form a foot with a pause (which is available between the subject and the object, I assume). Either way, the domains are different from those in (65). Finally, consider the effect of emphatic (or contrastive) stress, shown in (67) and (68).

In (67), ‘stock’ forms a foot, with stress on the first syllable. In ‘buy stock’, Nonhead Stress goes to ‘stock’, which is already stressed. Finally, ‘want’ gets stress from left-to-right foot formation. So there are two feet. In (68), ‘want’ has emphatic stress, which can override that on ‘stock’. Thus, there is one foot. The verb ‘ant’ is outside the foot, yet it is in the same cyclic domain, because T3S can apply beyond the foot.

5.1.4. Comparison of domains
I have shown earlier that the domains of tone sandhi are often different in Shanghai, Mandarin, and Xiamen. Now if compound stress and phrasal stress are the same for all languages, as I have suggested, foot formation should be, too. How then do we explain the differences in the tone sandhi domains? I suggest that the same set of information is
present for all the dialects, but Shanghai, Mandarin, and Xiamen make use of different aspects of the information for their tone sandhi. In particular, I suggest (69).

(69)  Boundaries for tone sandhi:
Shanghai: Foot boundaries
Mandarin: Cyclic domain (a foot plus adjacent free syllables)
Xiamen: Pre-pause heavy syllable

The examples below show the analysis of sandhi domains for the three dialects, where parentheses indicate foot boundaries, Ø indicates a functional element or a pause, and | indicates boundaries in Mandarin and Xiamen.

(70)  [[Old Li] [often [drink [rice wine]]]]

| Shanghai | ( ) ( ) ( ) ( ) | | |
| Mandarin | | | | | |
| Xiamen | | | | | |

(71)  [[Old Li] [like [Shaoxin’s [rice wine]]]]

| Shanghai | ( ) ( ) ( ) ( ) | | |
| Mandarin | | | | | |
| Xiamen | | | | | |

(72)  [[Sarajevo is-in Yugoslavia]]

| Shanghai | ( ) ( ) ( ) ( ) | | |
| Mandarin | | | | | |
| Xiamen | | | | | |

(73)  a. fry rice (phrase)  b. fry rice (compound)

| Shanghai | ( ) | ( ) |
| Mandarin | | | |
| Xiamen | | | |

The above analysis differs from the view that dialects can have different rules for foot formation. Instead, all dialects have the same foot formation process. Dialects can use different boundaries for tone sandhi not because there is a typology of foot formation but because the dialects have different kinds of tone sandhi. In particular, the tone sandhi in
Shanghai involves tone deletion from unstressed syllables, and therefore the domains are strictly foot-based. (The fact that tone deletion happens in Shanghai but not in Mandarin or Xiamen is due to another reason. According to Duanmu (1990, 1999), Shanghai syllables are all CV and not inherently heavy. They receive stress primarily through Nonhead Stress. In contrast, most Mandarin and Xiamen syllables are inherently heavy (CVX), and each can form a bimoraic foot and has its own stress even without Nonhead Stress.) The tone sandhi in Mandarin does not involve tone deletion, and so it can continue beyond a foot. Finally, the tone sandhi in Xiamen does not depend on the tones of neighboring syllables but only on whether a position is pre-pausal. Therefore, except when foot boundaries create pauses in a deliberated style, feet are mostly irrelevant.

5.1.5. Further evidence for the stress-based analysis
The present analysis can also account for a number of word length effects, discussed in Duanmu (2000). A peculiar property of Chinese is that many words have two forms, a long disyllabic form and a short monosyllabic form. Some examples are shown in (74), where the redundant word in the long form is shown in parentheses in the gloss.

(74)  | Long          | Short       | Gloss
  | ji-shu       | ji          | ‘skill-(trick)’
  | gong-ren     | gong        | ‘worker-(person)’
  | shou-biao    | biao        | ‘(hand)-watch’
  | gong-chang   | chang       | ‘(work)-factory’
  | zhong-zhi    | zhong       | ‘to plant-(plant)’
  | da-suan      | suan        | ‘(big)-garlic’
  | shang-dian   | dian        | ‘(business)-store’
  | shu-cai      | cai         | ‘(vegetable)-vegetable’

If each word has two forms, a two-word compound should have four combinations. However, in a noun-noun compound, only three of the combinations are good. An example is shown in (75), where 1 is a monosyllabic word and 2 is a disyllabic word.

(75)  | [noun noun]          |
  | 2-2 ji-shu gong-ren  |
  | 2-1 ji-shu gong      |
  | *1-2 * ji gong-ren   |
  |   1-1 ji gong        |
  | ‘skilled worker’     |

Similarly, in a verb-object phrase, three combinations are generally good and one is bad, as shown in (76), where ‘?’ indicates a marginal form.
In the present analysis, the word length effects are due to foot structure. Consider the analysis in (77), where S is a stressed syllable and s is an unstressed syllable. By Nonhead Stress the first noun is always stressed. The second noun is stressed if it is disyllabic.

The problematic case is [1-2], where the first syllable has stress but is too short to be a foot (it is possible to delete the stress from the second noun so that the compound forms one foot, but the preferred solution seems to be to choose a disyllabic first noun, or to choose a monosyllabic second noun). Next consider [V O], analyzed in (78). By Nonhead Stress the object is always stressed. The verb is stressed if it is disyllabic.

Here the problematic case in [2-1]. Intuitively, the problem is that the object has more stress than the verb (because the nonhead has more information than the head), and so the object should not be shorter than the verb (because length is a major aspect in the manifestation of stress). Another way to look at it, suggested by Burzio (1994), is that a strong foot (one with two or more syllables) has greater attraction for stress than a weak foot (one with one syllable), and so the object will lose main stress to the verb.

Another set of word length effects are also discussed in Duanmu (2000) and illustrated in the examples in (79), where VV is a disyllabic verb, V a monosyllabic verb, OO a disyllabic logical object of the verb, and O a monosyllabic logical object of the verb.

(76) [verb object]
    2-2 zhong-zhi da-suan
    *2-1 zhong-zhi suan
    1-2 zhong da-suan
    ?1-1 zhong suan
    ‘to plant garlic’

(77) [noun noun]
    2-2 (Ss)-(Ss)
    2-1 (Ss-s)
    *1-2 (S)-(Ss) first foot not binarity
    1-1 (S-s)

(78) [verb object]
    2-2 (Ss)-(Ss)
    *2-1 (Ss)-(SØ)
    1-2 s-(Ss)
    ?1-1 s-(SØ)

(79) [[OO VV] NN] *
    wheat harvest machine
    xiao-mai shou-ge ji-qi
    [VV OO] NN
    wheat-harvesting machine
    shou-ge xiao-mai ji-qi
In general, when the verb and the object are both disyllabic, their order must be [OO VV]. When they are both monosyllabic, their order must be [V O]. When the object is monosyllabic and the verb disyllabic, neither word order is good. Some exceptions can be found; for example, when N is zhe3 者 ‘person/the one’, both [OO VV N] and [VV OO N] can be used (He 2004). But for most N and NN the general patterns are as in (79).

In the present analysis (slightly different from that of Duanmu 2000), [[VV OO] NN] is bad because [VV OO] is a phrase, which should not occur inside a compound. On the other hand, [V O] can form a foot (as most disyllabic structures can), and hence a compound, because a foot can shield the syntactic information inside it, so that it is treated as a word. This is called the ‘Foot Shelter’ effect in Duanmu (2000), which is supported by the fact that exocentric compounds, such as huo-shao ‘fire-burn (baked wheaten cake)’ and kai-guan ‘open-close (switch)’, occur only in disyllabic structures. Next, [[VV O] NN] is bad because [VV O] is a phrase (and a bad phrase, see above), which is too long for a foot to shield its internal structure, and so it cannot occur inside a compound. Finally, [[O VV] NN] is bad because [O VV] is [1-2], which is a bad length structure for compounds (see above).

5.1.6. Summary
I have offered an analysis in which all dialects have the same foot formation process. Therefore, there is no need to assume that there are different foot types in Chinese dialects (Duanmu 1995, Iwata 2001). It is also worth pointing out that the predicted stress positions (the first syllable of a foot) agree with phonetic facts and other intuitive descriptions. For example, Zhu (1995) found that the rime of the foot initial syllable in Shanghai is phonetically nearly twice as long as that of noninitial syllables. In addition, Chao (1968), Xu (1982), and Hoa (1983) give similar judgments for stressed positions in Mandarin as the present analysis predicts (except for the pre-pause syllable, which is influenced by boundary effects). Finally, the present analysis follows a general theory of stress, and there is no need to make additional assumptions for grammatical words, pronouns, or the influence of emphatic/contrastive stress.

5.2. Chen (2000)
The analysis of Chen (2000) is an extension of Shih (1986) and is shared by Dell (2004). It makes the assumptions in (80) and (81).
Foot formation in Mandarin (ordered steps):

a. Grammatical words cliticize to the nearest lexical words.
b. Feet are formed cyclically for compounds.
c. Form a binary foot for the innermost disyllabic unit.
d. Start a foot boundary before emphatic/contrastive stress.
e. Form feet left-to-right for free syllables.
f. Join stray syllables with the nearest foot.

Application of T3S:

A T3S domain is a foot. T3S is optional between two feet.

The main difference between Chen’s analysis and mine is that for Chen foot formation does not depend on stress. For example, in Chen’s analysis a foot can be (s S), (S s), (s S s), (S s S), etc., where S is a stressed syllable and s is an unstressed one. For this reason, Chen introduces a new term—the ‘minimal rhythmic unit’ (MRU)—for what is traditionally called a syllabic foot (e.g. Chen 1979, Shih 1986), but it does not change the fact that the MRU is a new phonological entity. In addition, Chen needs separate assumptions to account for the difference between grammatical words and lexical words, and for the influence of emphatic/contrastive stress.

It can be shown that if we omit (80f), then Chen’s MRUs mostly correspond to the feet in my analysis. Consider the examples in (82).

<table>
<thead>
<tr>
<th>Present analysis</th>
<th>Chen’s analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td>mai (hao shu)</td>
<td>(mai (hao shu))</td>
</tr>
<tr>
<td>[buy [good book]]</td>
<td>[buy [good book]]</td>
</tr>
<tr>
<td>‘to buy good books’</td>
<td>‘to buy good books’</td>
</tr>
<tr>
<td>(wu-wu)-wu</td>
<td>((wu-wu)-wu)</td>
</tr>
<tr>
<td>‘five-five-five’</td>
<td>‘five-five-five’</td>
</tr>
</tbody>
</table>

In the [V O] case, both analyses build a foot for ‘good book’. In the present analysis, the verb remains outside the foot; it undergoes T3S not because it is in the same foot as the object but because T3S can apply beyond a foot. In Chen’s analysis, the foot can have two (or more) layers; the object is in the inner layer and the verb is in the outer layer. Similarly, in ‘five-five-five’, both analyses build a foot for the first two syllables. In the present analysis, the third syllable remains outside the foot, but in Chen’s analysis it is in the outer layer of a nested foot. In summary, Chen’s analysis assumes that all syllables must belong to a foot; as a result, foot types are rather complicated. In contrast, the present analysis assumes that some syllables can remain unfooted, and all feet are consistently left-headed.


Wang (2004) does not discuss T3S but does discuss foot formation. Her analysis of is summarized in (83), which echoes a similar proposal by Lu (1989).
Syllables form feet according to their syntactic ‘closeness’.
Syllables that are syntactically close belong to the same foot.
Syllables that are not syntactically close belong to different feet.
A final syllable may appear to have more stress, owing to boundary effects.

Wang argues that in languages like English, stress is lexically distinctive, and so foot formation is based on stress assignment. In contrast, in languages like Chinese, stress is not lexically distinctive, and so foot formation is not based on stress assignment, but based on the syntactic closeness between syllables. Some syllables may appear to have more stress, such as those before a pause, but the stress effect is secondary and has no bearing on foot formation.

There are some problems in this analysis. First, it probably overstates the fact that stress is lexically distinctive in English. For example, it can be argued that pairs like import-import are not distinguished by stress but by word categories (noun vs. verb), and word categories in turn has affected stress. If so, there are not many word pairs in English that are distinguished by stress alone. Second, it downplays the fact that stress can be lexically distinctive in some Chinese word pairs. Third, it does not explain how feet are formed in polysyllabic words or flat structures (such as a string of digits). Fourth, it seems to overlook the many similarities between English and Chinese that I have discussed above. In any case, Wang does not offer many examples and so it is hard to evaluate her analysis in detail.


(84) Every disyllabic unit forms a foot, whether it is a word or a phase.
   In words and compounds, build disyllabic feet from left to right.
   In phrases, build disyllabic feet from right to left.

Feng does not discuss T3S, but does discuss some word length effects. His analysis of [N N] and [V O] are shown in (85) and (86), where S represents a syllable and # is a word boundary.

(85) [N N] compounds: L→R foot formation
     2-2   (SS)#(SS)
     2-1   (SS)#S
     *1-2  (S#S)S   foot boundary not aligned with word boundary (?)
     1-1   (S#S)
(86) [verb object]: R→L foot formation
    2-2 (SS)#(SS)
    *2-1 S(S#S) foot boundary not aligned with word boundary (?)
    1-2 S#(SS)
    ?1-1 (S#S)

Feng says that [1-2] does not fit left-to-right foot formation and [2-1] does not fit right-to-left foot formation. However, the precise reason for the violation is not discussed. Clearly, it is not because the foot contains a word boundary, since [1-1] is fine. Perhaps the foot boundary must be aligned with a word boundary? The answer is also no, because in polysyllabic words the foot boundary need not align with a word boundary, as shown in (87).

(87) (duomi)(nija) ‘Dominica’
    (jieke)(siluo)(fake) ‘Czechoslovakia’

Consider also the example in (88), which is an actual name of a park.

(88) (Shijing#Shan)#(Gongyuan) *(Shijing)#(Shan#Gong)yuan
    [[Shijing Mountain] Park]

If we apply left to right foot formation, the result is (SS)#(S#S)S, which is incorrect. The correct foot structure is (SS#S)#(SS). Perhaps what Feng wants to say is something like (89).

(89) Feet are built cyclically, first for words and then for compounds and phrases.
    In words and compounds, there cannot be a free syllable on the left.
    In phrases, there cannot be a free syllable on the right.

But why is there such a difference between words and phrases? Feng (1998) suggests that left-to-right foot formation is the most ‘natural’ process for words, but one wonders why it is not the case for English words (consider potato, America, Chicago, Japan, etc.), where the foot does not start from the left. Nor is there an obvious reason why foot formation for phrases should be right-to-left. On the other hand, if foot formation is based on stress, there is an explanation. In a compound, Nonhead Stress is on the left, and so it should start with a disyllabic foot. In [V O], Nonhead Stress is on O, and so it should end in a disyllabic word.

Feng (2004) raises two questions for the stress-based analysis. First, Feng points out that some disyllabic compounds have initial stress, such as zhuel-zi0 ‘table’, and some have final stress, such as ying2-hang2 ‘bank’. If they both form one foot, stress cannot be relevant. However, as I have discussed above, compounds like ying2-hang2 ‘bank’ are heavy-heavy, which appear to have final stress only before a pause. Otherwise, they have initial stress, as do heavy-light compounds, such as zhuel-zi0 ‘table’. It is worth noting that there are no light-heavy compounds, which follows from an analysis that assumes left-
headed feet, but not from an analysis that allows right-headed feet. Second, Feng points out that in a [[V O] NN], compound, [V O] is a foot, yet according to Nonhead Stress, O has more stress, which is inconsistent with left-headed feet. The answer to this problem is that a foot can shield its internal syntax, so that the disyllabic [V O] is treated as a word in this environment, and thus Nonhead Stress does not apply. The supporting evidence is that, while in a [V O] phrase O indeed has more stress than V, in [[V O] NN] V has more stress than O. Thus, neither of Feng’s criticisms of the stress-based analysis seems compelling.

6. Concluding remarks
I have discussed some patterns of tone sandhi in Chinese and argued that a full understanding of tonal phenomena requires an understanding of stress. In addition, I have offered a general theory of stress and shown that it plays a key role in foot formation and the syntax-phonology interface, without additional assumptions for grammatical words, pronouns, or the influence of emphatic or contrastive stress. Moreover, I have argued that different kinds of tone sandhi need not imply different rules for foot formation. Rather, different kinds of tone sandhi may be conditioned by different parts of a general prosodic structure, which is present in all dialects.

Many linguists still doubt that foot formation in Chinese is based on stress, but there seems to be little evidence for the alternatives (See Duanmu 2004 for more discussion). It is worth noting that the predicted stress positions in the present analysis agree with phonetic facts and other intuitive descriptions. If the present analysis is correct, it supports the view that tone and stress can exist in the same language (Liberman 1975, Pierrehumbert 1980, Goldsmith 1981), and offers a demonstration in a true tone language. Hopefully, it can help resolve similar controversies in Japanese and African languages (Odden 1999, Kaji 1999), and offer an alternative approach to the syntax-phonology interaction within them (Odden 1987, 1995).

*Acknowledgments: For helpful comments and discussions, I thank the audience at the 2004 Symposium on Cross-Linguistic Studies of Tonal Phenomena in Tokyo, and in particular my discussant Professor Ray Iwata. I also thank Professor Shigeki Kaji and other organizers for hosting the symposium.
References


Fry, Dennis B. 1955. Duration and intensity as physical correlates of linguistic stress. JASA 27:765-768.


