

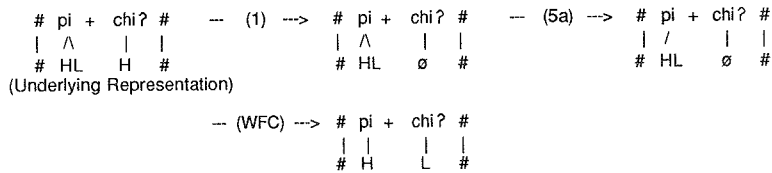
SHANGHAI TONAL PHONOLOGY
'RIGHTWARD SPREADING'
SOME ARGUMENTS BASED ON ACOUSTIC EVIDENCE

Takako Toda
Faculty of Asian Studies and Department of Linguistics
The Australian National University

ABSTRACT- In recent years, linguists have started looking to acoustics for evaluation of phonological questions (Ohala 1986; Ladefoged 1989). This paper presents some acoustic data pertaining to the Shanghai tone sandhi, and argues that for Shanghai the current phonological assumption 'rightward spreading' is not phonetically plausible.

INTRODUCTION

Most of the recent phonological analyses made of the tone sandhi in the Shanghai dialect (Zee and Maddieson 1979; Jin 1986, forthcoming) agree that it is best analysed as the rightward movement of the tone associated with the initial syllable. Recently the view of the autosegmental representation of tones, useful for the systematic analysis of Shanghai tonal phonology, is particularly well accepted. Using an autosegmental approach, the rightward spreading of a high falling tone is handled as follows (Jin 1986:11);



(1) Tone Deletion Rule 1; (5a) (De)Association Rule (Zee and Maddieson 1979:76)¹; (WFC) Well-Formedness Condition (Goldsmith 1979)

Here HL (high/low) indicates underlying representation of a high falling tone. In autosegmental phonology, a single segment can carry two tonal specifications in sequence. The separate level of tiers allows tonal melody (such as HL, LHL) to move independently from syllables. This is considered to be extremely useful to describe tonal phenomenon such as tone spreading, where L tone (second element of HL which is associated with the first syllable) moves onto the second syllable. The problem I see here, however, is that the final stage of this phonological process (i.e. output) does not tell us precisely what the phonetic realization is; a disyllabic contour, with H element at the onset and L element at the offset, can be realized in many ways - for example, [44-31], [53-21], [52-11] (when [51] represents a monosyllabic high falling contour). A number of different contour shapes are predicted and at least some of them should be auditorily distinguishable (i.e. linguistically significant). This is highly controversial since the current generative theory is based on agreement that such derivation should lead us to the correct surface form from the underlying input.

The crucial evidence for this argument is that Daishan (Toda 1989), a neighbouring dialect, demonstrates a contrast between [44-31] and [53-21], both of which are derived from the same input high falling tone [51]². When the duration is equalized, the latter has an identical shape with a monosyllabic high falling contour (i.e. [53-21]=[51]). This means that Daishan has a real *phonetic* tone spreading of a unitary high falling contour, and this corresponds to *phonological* tone spreading. [44-31], on the other hand, is regarded as a concatenation of two separate units, namely, [high level-mid falling] (Toda forthcoming)³. It is difficult to see how this contrast can be captured by using the current framework of autosegmental representation shown above.

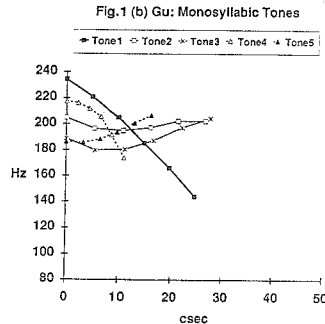
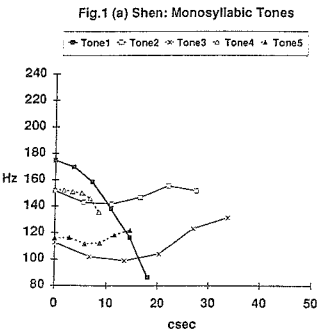
In this paper, I investigate, in acoustic terms⁴, the relationship between the monosyllabic high falling tone (i.e. the tone originally associated with first syllable) and the disyllabic sandhi output. The

durational difference between monosyllabic and disyllabic will be equalized in order to judge the extent to which the actual shape of the contours are similar. To my knowledge, no previous work of this kind has been done in statistically significant sample sizes. From the viewpoint of acoustics, the results show that the disyllabic sandhi output is best interpreted as a concatenation of a high level and a mid falling tone. This situation is therefore very similar to Daishan [44-31]. Given that both Shanghai and Daishan have a high falling tone [51] and sandhi output [44-31], if we regard this surface form [44-31] as a result of rightward spreading, how can we handle [53-21] in Daishan? In fact the Daishan data cannot be handled unless [44-31] is treated in different ways. One of the propositions for autosegmental phonology is a universalist approach, and thus I claim that treating Shanghai [44-31] as a result of rightward spreading goes against to this principle.

THE DATA

Monosyllabic Contour Types:

The following section deals with the acoustic data of my informants; Shen, a male speaker in the late 20's and Gu, a female speaker in the early 30's. Both speakers were born and grew up in Shanghai. Fig.1 shows the acoustic data. Fundamental Frequency (Fo) was sampled at 20% intervals of the absolute duration of the Fo contour. The measurement includes the point of onset as 0% and offset as 100%. X axis shows the duration of the vowel of monosyllabic words (centiseconds) and Y axis shows the Fo in hertz (Hz). Each Fo contour is the mean of at least six tokens. Male speakers usually show a lower mean Fo value than female speakers. Shen's monosyllabic Fo is ranging between 85-175Hz while Gu's is ranging between 145-235Hz. (Note that the range in Fo, 90Hz, is the same between the two speakers.) The major differences concerning the shape of Fo contours are; 1) Shen's high falling Fo (Tone1) has a small level component, but Gu's shows a constant decrease in Fo after the onset. 2) the duration of high falling Fo is about half of long rising Fo (Tone2, 3) for Shen, but in the case of Gu, it is about the same as long rising Fo. 3) Gu's low rising Fo value is relatively higher for both long and short syllables than that of Shen. In other words, Gu's range in Fo is 40Hz (180-220Hz) when high falling Fo is disregarded, and thus is smaller than Shen's.

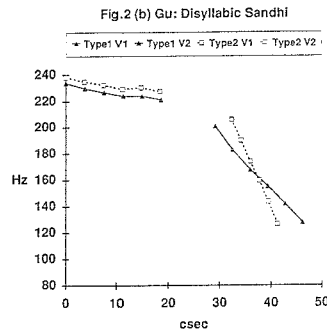
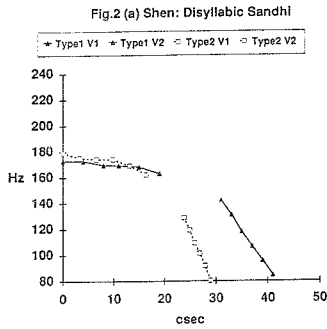


Disyllabic Contour Types:

In this section, the contour shape of Tone1 (i.e. monosyllabic high falling Fo) is compared with that of disyllabic sandhi output (i.e. disyllabic high falling Fo), which is phonologically analysed as rightward movement of a L tone. The relationship between these two contours is investigated from an acoustic point of view to determine to what extent they are similar. In other words, to what extent does phonological tone spreading actually correspond to phonetic tone spreading in Shanghai?

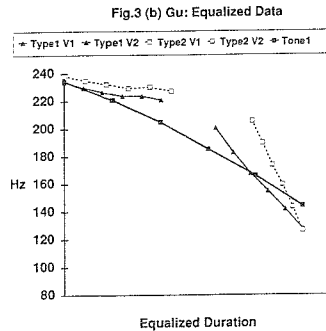
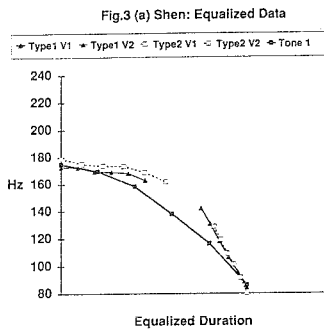
Fig.2 (a) and (b) show the disyllabic Fo shapes for Shen and Gu respectively. They represent the sandhi output shapes when the input tone in the first syllable is high falling (and any tones in the second syllable since it is deleted by Tone Deletion Rule). The first Fo tracing (i.e. high level component around 175Hz for Shen and 230Hz for Gu) represents the Fo of the first vowel (V1), and

the second F_o tracing (i.e. falling component starting around 135Hz for Shen and 200Hz for Gu) represents the F_o of the second vowel (V2) of the disyllable. The gap between the tracings of V1 and V2 shows the duration of the voiceless intervocalic consonant. Each figure has two sets of disyllabic F_o tracing; the F_o tracing shown in black marker (Type1) has long V2 (i.e. Tone 1, 2 and 3), while that shown in white marker (Type2) has short V2 (i.e. Tone 4 and 5). V1 is long for both types. Each F_o tracing is the mean value of at least six tokens.



Equalized Fo Contours:

In order to compare the shape of the F_o contours, the duration of the disyllables with long and short second syllables (i.e. those shown in Fig.2) are equalized. The F_o contour of the monosyllable (i.e. Tone 1 shown in Fig. 1) is also shown in the same figure with bold line for comparison (See Fig.3).



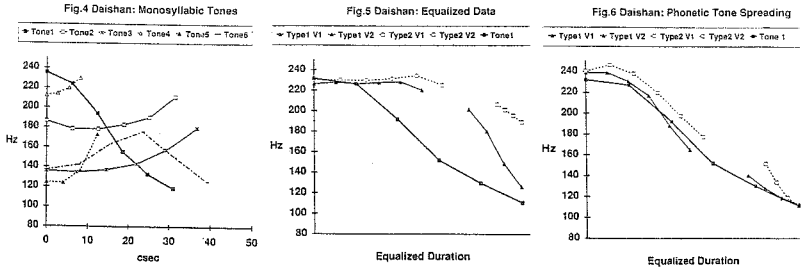
The F_o value at the onset and offset is nearly identical between monosyllabic and disyllabic contours for both speakers, except that Gu's monosyllabic F_o contour shows about 20Hz difference from her disyllabic F_o contours at the offset (0.1 < p < 0.05). It appears, therefore, that both speakers have target onset and offset values (for Shen and Gu, the target onset is around 175Hz and 235Hz, and offset is 80Hz and 140Hz respectively). The overall shape of the F_o contour, on the other hand, is noticeably different between monosyllable and disyllable for both speakers; the F_o tracing of V1 in the disyllable stays level (an average of 175Hz for Shen and 230Hz for Gu) while the monosyllabic F_o tracing shows a decrease in F_o after the onset (-5.0 Hz/csec for Shen and -3.5Hz/csec for Gu). The difference between the offset of the V1 in the disyllable (about 160Hz for Shen and 225Hz for Gu) and the corresponding point of F_o in the monosyllable (about 145Hz and 205Hz respectively) is about 30Hz

for both speakers, and this difference is statistically significant. Therefore these acoustic data indicate that the disyllabic output is different from the monosyllabic high falling Fo, and the disyllabic Fo is practically a concatenation of [high level-mid falling].

DISCUSSION

I have shown that the Fo shape of disyllabic sandhi output is different from the monosyllabic high falling Fo. In this section, I shall compare Shanghai data with the data of tone spreading seen in Daishan.

Fig.4 shows Fo contours of the monosyllabic tones in Daishan⁵. Now let us assume that Daishan high falling tone (Tone1) corresponds to that in Shanghai. From this tone, two types of sandhi output are derived in Daishan. Firstly, a concatenation of [high level-mid falling] can be an output (See Fig.5), which is very similar to the disyllabic Fo in Shanghai shown in Fig.3 (a) and (b). The only significant difference is that when the second syllable is short (Type2), the V2 Fo does not fall down to the same level as the offset of monosyllabic Fo and disyllabic Fo with a long second syllable⁶. Another type of sandhi output, which is also derived from the same high falling tone, is shown in Fig.6. The shape of disyllabic Fo contours is identical with the monosyllabic Fo contour when intrinsic factors (intervocalic consonant and intrinsic rising effect on Fo caused by the syllable final glottal stop) are disregarded (Toda 1989:73). This is evidence for a real phonetic spreading of a high falling tone over the disyllable. Given that this type of sandhi shows a real tone spreading, we can say that phonological tone spreading corresponds to phonetic spreading to describe this situation. If the autosegmental representation of rightward spreading, shown in page 1, is to account for the former output (i.e. [high level-mid falling]), it is difficult to describe a real phonetic spreading using the same framework of autosegmental phonology.



Under these circumstances, we have to assume that the autosegmental representation of rightward spreading, shown in page 1, accounts for the real phonetic spreading type. The concatenation type, [high level-mid falling], has to be represented in different ways in order to show the contrast with the real spreading. In particular, the representation has to capture the initial level component of [high level-mid falling].

One possible solution would be to apply an extra association line. In this model, rule (4), that is, inserting an association line between an initial H and the final syllable, would cause the second syllable to assimilate in tone, and change the relative domains of the tonal segments (see Goldsmith 1979:205). This model accounts for the level component of the tone associated with the first syllable. The derivation of the disyllabic sandhi output is as follows;

$$\begin{array}{c}
\# \text{ pi } + \text{ chi? } \# \quad \text{--- (1) ---} \rightarrow \quad \# \text{ pi } + \text{ chi? } \# \quad \text{--- (2) ---} \rightarrow \quad \# \text{ pi } + \text{ chi? } \# \\
\begin{array}{c} | \quad \wedge \quad | \quad | \\ \# \text{ HL } \quad \text{H} \quad \# \end{array} \quad \begin{array}{c} | \quad \wedge \quad | \quad | \\ \# \text{ HL } \quad \emptyset \quad \# \end{array} \quad \begin{array}{c} | \quad / \quad | \quad | \\ \# \text{ HL } \quad \emptyset \quad \# \end{array} \\
\text{(Underlying Representation)} \\
\\
\text{--- (3) ---} \rightarrow \quad \# \text{ pi } + \text{ chi? } \# \quad \text{--- (4) ---} \rightarrow \quad \# \text{ pi } + \text{ chi? } \# \\
\begin{array}{c} | \quad | \quad | \quad | \\ \# \text{ H } \quad \text{L} \quad \# \end{array} \quad \begin{array}{c} | \quad \backslash \quad / \quad | \quad | \\ \# \quad \text{H} \quad \text{L} \quad \# \end{array}
\end{array}$$

(1) Tone Deletion Rule 1; (2) (De)Association Rule (Zee and Maddieson 1979:76); (3) Well-Formedness Condition (Goldsmith 1979); (4) Association Rule

This solution still has some problems: Firstly, this representation suggests that the onset of V2 has the same value as V1 (i.e. H tone), and it does not show to what extent the onset of V2 assimilates to the offset of V1. If this representation suggests that the output is [high level-high falling] instead of [high level-mid falling], there are some discrepancies: a) Neither acoustic nor auditory values of the tone on the second syllable correspond to values found on the monosyllabic high falling tone. If we compare V2 Fo shape with monosyllabic high falling Fo (Tone1), the difference is apparent. b) When compared with other monosyllabic Fo, the onset of V2 Fo corresponds to the value of mid rising Fo of both speakers. This means that the onset of V2 Fo has mid value, instead of high. c) The phonetic value which corresponds to the phonological H tone is about 175Hz for Shen and 235Hz for Gu, and the mean onset of V2 is about 135Hz and 205Hz respectively. The difference is 30-40Hz, that is, more than one third of the Fo range of the both speakers. The difference is statistically significant, and it appears that we can not attribute the lower onset of V2 to intrinsic Fo perturbation caused by the intervocalic consonant. The above data show that the tone on the second syllable is not best described in terms of HL. Although in the phonological level, in contradiction to the phonetic reality⁷, S2 can be analysed as HL if we can prove the onset of S2 is underlyingly the same as the H tone. Unfortunately, we do not have any ways to prove it within the present data shown above. The fact that we need to have M (mid) level in order to account for all the underlying monosyllabic tones in Shanghai (Zee and Maddieson 1979), however, seems to suggest the necessity of more than three levels of underlying representation in explaining the rightward spreading in terms of the movement of tonal elements. Secondly, although the modified version describes the output shape more accurately than the previous version because the rule (4) was introduced, this rule has little phonetic motivation. It does not explain why this particular output should be derived in such a way. Thirdly, and most importantly, the controversy arises from the fact that we can not explain the distinction between these two types of output (the concatenation type [44-31] and the phonetic spreading type [53-21]) in terms of underlying tones or stress. The underlying tone is Tone1, and the underlying stress is on the first syllable for both types. Thus it appears that there is no way that the distinction can be explained in terms of applying derivational rules to underlying representations in order to account for the surface forms.

CONCLUSION

In this paper, the phonetic reality of 'rightward spreading' in Shanghai phonology was investigated. The findings were as follows: 1) The shape of the disyllabic Fo contour was different from that of the monosyllabic Fo contour. This means that the phonological description 'tone spreading' does not correlate with the acoustic spreading in Shanghai as it does in Daishan. 2) The current model does not capture this reality and therefore needs to be modified. An alternative phonological solution was suggested by applying an extra association line and a rule which would cause the second syllable to assimilate in tone, and change the relative domains of the tonal segments. This model accounts for the level component of the tone associated with the first syllable. 3) Even in the modified version, some problems still remain. Although the description of two types of sandhi can be improved to show the contrast to a certain extent, the nature of the distinction cannot be explained in terms of applying derivational rules to underlying representations. This is because these two types of output have the same underlying condition. It was concluded that it is difficult to capture this situation within the current framework of autosegmental phonology, and generative phonology in general.

NOTES

1 The Association Rules are as follows : 1. Delete the association line between the second of two tones and the first syllable in any compound. 2. Insert an association line between an initial L and the final syllable in a bisyllabic or trisyllabic compound (Zee and Maddieson 1979:76).

2 Both [44-31] and [53-21] have stress on the first syllable. Recent assumption of a metrical approach to analysis of tone sandhi of Chinese dialects (Wright 1983) does not account for this situation either.

3 Liao's (1983) acoustic data in Suzhou show a similar contrast between a concatenation of [high level-mid falling] and a spreading of a unitary high falling contour. The former is derived from a high level tone, not a high falling tone, as in the latter case. This is further strong evidence that [high level-mid falling] is a separate lexical sandhi output from the spreading of unitary high falling contour.

4 For the sake of brevity, observations in this paper are limited to the high falling tone. Also it must be acknowledged that the data presented here were only of two speakers and much more work needs to be done with a number of speakers. Ideally speaking, acoustic data of trisyllabic sandhi should also be compared.

5 Daishan has a convex (rising-falling) tone, which is not found in Shanghai.

6 This is probably due to the interaction between the shorter duration and the syllable final glottal stop (which accompanies short syllables).

7 The nature of the tone in the second syllable is interpreted as 'default falling' regarding a similar sandhi pattern observed in Zhenhai (Rose and Toda 1989). In Zhenhai data, it is 'possible to predict the F_0 onset value to within 4Hz from the S1 F_0 trajectory extrapolated for the duration of the intervocalic consonant. The F_0 onset of S2 seems therefore to be determined by phonatory inertial effects from the preceding syllable, and the rest of the F_0 shape (and acoustic shape in general) probably reflects a simple relaxation in vocal cord tension and/or sub-glottal pressure.' (Rose in print). With regard to the Shanghai data presented here, it is possible to predict an S2 onset value to within an average of 8.5Hz (Shen 6.5Hz; Gu 10.5Hz).

REFERENCES

- GOLDSMITH, J. (1979) "The aims of autosegmental phonology", Dinnsen D.A. ed., *Current Approach to Phonological Theory*, 202-222.
- JIN, S. (1986) *Shanghai Morphotonemics*. Indiana University Linguistics Club.
- (forthcoming) "Toward a systematic account of Shanghai tonal phonology", to appear in Wang J.L. & Smith M. ed., *Studies in Chinese Phonology*.
- LADEFOGED, P. (1989) "Representing phonetic structure", *UCLA Working Papers in Phonetics* 73.
- LIAO, S. (1983) "Suzhouhua danzidiao, shuangzidiao de shiyanyanjia", *Yu Yan Yan Jiu* 2, 41-83 (in Chinese).
- OHALA, J. (1986) "Consumer's guide to evidence in phonology", *Phonology Yearbook* 3, 3-26.
- ROSE, P. (in print) "Acoustics and phonology of complex tone sandhi: An analysis of disyllabic lexical tone sandhi in the Zhenhai variety of Wu Chinese", *Phonetica*.
- ROSE, P. & TODA, T. (1989) *Typology of Tone Sandhi Rules in Wu Dialects*. Paper presented at the Australian Linguistic Society Conference, Melbourne, September 1989.
- SHERARD, M. (1972) *Shanghai Phonology*. Ph.D dissertation, Cornell University.
- TODA, T. (1989) "Acoustic quantification of tone spreading in Daishan and Mende", *Cahiers de Linguistique Asie Orientale* Vol.XVIII, No.1, 63-80.
- (forthcoming) "Daishan tone sandhi: Acoustically based description and phonological analysis", to appear in *Proceedings of the International Conference on Wu Dialects*. Chinese University of Hong Kong.
- WRIGHT, M.S. (1983) *A Metrical Approach to Tone Sandhi in Chinese Dialects*. Ph.D Dissertation, University of Massachusetts.
- ZEE, E. & MADDIESON, I. (1979) "Tones and tone sandhi in Shanghai: Phonetic evidence and phonological analysis", *Glossa* 14. 1, 45-88.