

## Using a Prosodic Labeling Text (PLT) in the Synthesis of Spoken Chinese

Wu Zong-ji,  
Institute of Linguistics, (CASS)  
E-mail: Linmc @sun.ihep.ac.cn

### ABSTRACT

The prosodic features of Spoken Chinese play the important roll of the naturalness, a list of prosodic labeling symbols represents all the prosodic features is given in this paper, and a paragraph of "Prosodic Labeling Text"(PLT) is also attached for example.

### INTRODUCTION

The naturalness of synthetic speech is not only depending upon the higher fidelity of its articulation and intelligibility, but also fits prosodic features, i.e., intonation, stress and tempo, which are played by the role of pitch, amplitude and duration respectively. The works of synthesis of Spoken Chinese(SC) are now entered the stage of naturalness researches. Since the synthetic researchers are no more satisfied to their conventional works on Chinese syllables and short phrases merely with citation forms, then the study of rules and patterns of tone-sandhi and intonation becomes popular in several laboratories. However, there are a lot of hard nuts to crack, for the coarticulation effects are not only effected in segmental level but also (even more) in supra-segmental ones.

In our laboratory, on the base of phonetic research, a series of experiments were held on tone-sandhi rules of poly-syllabic combinations in phrasal contours(PC), as well as of transposition processes on sentential intonations in SC, their results had been employed in synthesis programming.(Wu, 1982, 1985, 1988, 1988a, 1990, 1993, 1995a; Yang, 1994). Moreover, it is then noted that the naturalness of spoken speech are not merely characterized by pitch, the duration and stress are also played the important role(Wu, 1995). Therefore, prosodic rules of synthetic speech need to be fully investigated. Being tested in a synthesis system, a series of further comprehensive prosodic processing in the sake of naturalness had been dealt with. Based on these works, a plan of labeling all the prosodic words as well as the pause and rhythmic structures in SC utterance are put on the top of research agenda. Thus a system of "Prosodic Labeling Text"(PLT) of SC was brought to the schedule(Wu, 1996, 1996a). An inventory of prosodic labeling symbols and a paragraph of PLT are given in this paper for example.

### LIST OF PL SYMBOLS OF SC:

#### (1) PL symbols for segment of PC:

- [ / ]: segmentation of PCs.
- { } : boundary of syllables.
- [2S]: bi-syllabic tone-sandhi pattern.
- [2S0]: ditto with neutral-tone ending.
- [3S1]: tri-syllabic tone-sandhi pattern with 2+1 structure.
- [3S2]: ditto with 1+2 structure.
- [3S0]: ditto with neutral-tone ending.
- [4S2]: quadro-syllabic tone sandhi pattern with 2+2 structure.
- [4S1]: ditto with 3+1 structure.
- [4S3]: ditto with 1+3 structure.
- [C1],[F#2], etc.: register in semitone(ST) scale as the main key of PC in sentential intonation.
- [\*]: transposition of the main key, e.g. [\*+6]: 6 STs higher than the main key;[\*-2]: 2STs lower than the main key, etc.
- \*C: transposition canceled,(back to the main key).
- [R]: frequency range of PC in ST.

#### (2) PL symbols for monosyllables in PC:

[.]: separation between syllabic PLs in PC.

##### A. PL of tone:

[B]: tone-sandhi of the beginning syllable in PC.e.g.,[B-8]:ditto, with its key 8STs lower

than the main key.

[E]: tone-sandhi of the ending syllable in PC.e.g.,[E-5]:ditto, with its key 5STs lower than the main key.

[T]: transitional tone-sandhi contour between two syllables.e.g.,  
[T-7+3]: ditto, with its contour begun with a key 7STs lower than the main key and then raised 3STs in the ending.

[0]: neutral tone(syntactic) or neutralization(optional).

[<],[>]:tonal junctures, [<]:leftward; [>]:rightward.

B. PL of stress:

[^]: average amplitude in dB of the speaker's monosyllable in citation as the base point.  
e.g.,[^+4]: 4 grades higher than the base point,(each grade is the 25% of the base amplitude in dB).

C. PL of duration:

[~] average duration in ms of the speaker's monosyllable in citation as the base point.  
e.g.,[~+2]: 2gradeslongerthan the base point,(each grade is the 25% of the base duration in ms).

D. PL of pause:

[\$]: pause between 2 PCs, (in grade, each grade is 50ms).e.g.,[\$6]:  
pause up to 6 grades,(i.e., 50ms \* 6 = 300ms)

[&]: ending pause of a sentence.(length subject to circumstance).

E. PL of articulatory quality that effected the naturalness:

[ ]: silence gap of 40ms before a voiceless stop or a voiceless affricate when it is preceded by a syllable with voiced ending in a structure of immediate constituent.

[J]: juncture, the co-articulation effect between two syllables, phonetically termed "assimilation".

A SAMPLE OF PLT:

A SC text is spoken in Standard Chinese by a female, recorded and transcribed in "Pinyin" and "tone-letter", segmented into 7 PCs as the following:

SCT: (1)wo3 zhun3 bei4, (2)cong2 bei3 jing1, (3)chu1 fa1, (4)jing1 guo4,  
我 准 备 从 北 京 出 发, 经 过  
(5)nan2 jing1, (6)shang4 hai3, (7)dao4 hang2 zhou1.  
南 京 上 海 到 杭 州.

(I prefer to start from Beijing, via Nanjing, Shanghai to Hangzhou.)

PLT: (PC1) 3S2.G4.R12. /{B-10+10}wo3 {/>}zhun {E-8-4.~+1}bei4/ \$2  
(PC2) 3S2.\*-2.R17. /{~+2}cong2 { }bei3 {>}`jing1/  
(PC3) 2S.\*-4.R12. /{~-2}chu1 {~+2}fa1/ \$12  
(PC4) 2S.\*0.R12. /{~-2}jing1 {~+1}guo4/  
(PC5) 2S.\*+2.R17. /{>.\*-12.~+4}nan2 {>.\*~2}`jing1/  
(PC6) 2S.\*+1.R12. /{\*+1}shang4 {>.\*E-12-8}hai3/ \$5  
(PC7) 3S2.\*-9.R12./{~-2}dao4 {T}hang2 {E-9-2}zhou1/ &

DISCUSSION

There were two main problems in tonal analysis which were long bewildered Chinese phoneticians: i.e., the vast variants of the tone-sandhi in PCs among sentence; and of the frequency ranges in sentence among different speakers and different moods; which are all difficult to analyze and patternize. Poly-syllabic tone-sandhi rules up to quadro-syllabic combinations had been established by the author since the eightieth and tested in synthesis works. It may more or less solve the tone-sandhi's problem. As for the range of intonation, a study on the frequency ranges of the four tones of SC under different register groups had been shown that it is rather consistent in melodic scale instead of frequency one in normal speech(Wu,1994,1995). This can be used as a basic reference for predicting any PC with a common range width in normal sentences; or may generate any prominent PC in a sentence or any sentential intonation in a synthetic work with different range widths, just by

transposition processing(Wu,1993). Moreover, prosodic features other than pitch i.e., stress and duration are also playing the important role for naturalness. It is true that the tone pitch is the chief prominent cue in intonation, however, experiments on speech synthesis had shown that in exclamatory sentences, the absolute register of a prominent PC raised to a certain extent is not significant in loudness unless a certain degree of amplitude and/or duration are added to an appropriate proportion(Wu,1995). These rules are normalized and labeled to meet the programming requirement. Of course this list of PL are rather complex and are less practical yet for actual automatic processing in prosodic synthesizing. Further grading and simplifying of these labeling will be improved later on.

#### REFERENCES

- Wu,Z.J.(1982) "Rules of intonation in Standard Chinese", Reprints of Papers for the Working Group on Intonation, 13th Int. Cong. Linguists, Tokyo.
- Wu,Z.J.(1985) "Tone-sandhi rules of tri-syllabic combinations in Standard Chinese", Bull. of Chinese Linguistic Society, Com. Press.(in Chinese)
- Wu,Z.J.(1988) "The basic tone-sandhi patterns in Standard Chinese intonation", Essays on Linguistics, Festschrift for Prof. Wang Li, Com. Press.(in Chinese)
- Wu,Z.J.(1988a) "Tone-sandhi Patterns of quadro-syllabic combinations in Standard Chinese", ARPR-PL,CASS.
- Wu,Z.J.(1990) "Can poly-syllabic tone-sandhi patterns be the invariant units of intonation in Spoken Standard Chinese? Proc of 1st. Inter. Conf. on Spoken Language Processing, Kobe, 1990.
- Wu,Z.J.(1993) "A new method of intonation analysis for Standard Chinese: Frequency transposition processing of phrasal contours in a sentence",RPR-IL,CASS.1993.
- Wu,Z.J.(1994) "Further experiments on spatial distribution of phrasal contours under different range registers in Chinese intonation, Proc. Int.Symp.on Prosody, Yokohama.
- Wu,Z.J.(1995) "Predictability of different attitudinal intonation in Standard Chinese, 13th ICPhS, Stockholm, 1995.
- Wu,Z.J.(1995a) "A new method for intonation processing of Standard Chinese based on the relation between tonology and musicology" (in Chinese)(in Press)
- Wu,Z.J.(1995b) "Tentative planning of prosodic rules for the naturalness of synthetic spoken Chinese",RPR-IL,CASS.1994-1995.
- Wu,Z.J.(1996) "A planning of prosody-by-rule for improvement of naturalness in spoken Chinese synthesis",Proc. of Conference of Chinese Language and Letter in Computer Era, Tsing-Hua Univ. Press.
- Wu,Z.J.(1996a) "A design of 'Prosodic Labeling Text' used on the synthesis of Standard Chinese", 11th Conference of phonetic research, 1996, Beijing.(in Chinese)
- Yang,S.A.(1994),The Chinese Speech Synthesis Technique Oriented Acoustic-Phonetics, Document Press of Social Sciences, Beijing.