

# A Comparative Study of Sentential Stress Distribution in Mandarin Multi-Style Speeches

*\*Mingzhen Bao<sup>1</sup> & \*\*Min Chu*

*\*University of Florida, USA; \*\*Microsoft Research Asia, P.R.China*  
*joanneb@ufl.edu; minchu@microsoft.com*

## Abstract

This paper compares the distribution of sentential stresses among three speaking styles: *Lyric*, *Critical*, and *Explanatory*; and extends our previous study in the base phrase level to the sentence construction level and the prosodic word level. The results show that 1) The distributions of both rhythmic and semantic stresses act the same among styles within prosodic words, although the distribution tendencies change due to different structure properties of the words; 2) In the sentence construction level, the distribution tendency of rhythmic stress is quite similar across three styles in most construction types, while semantic stress presents more diversity among speaking styles. The *Explanatory* style shares a similar tendency with the *Neutral* style. The *Lyric* style differs from the *Neutral* style in constructions with the subject-predicate structure; the *Critical* style differs in constructions with the predicate-object, the adjunct-subject, and the adjunct-object structures. Generally, speaking styles have fewer effects on rhythmic stress distribution than on semantic stress. Such effects are more obvious in the sentence construction and the base phrase levels than the prosodic word level, where syntax plays a more crucial role in stress distribution.

## 1. Introduction

Stress has been defined as “the degree of loudness” from the viewpoint of speech perception [1]. It has been categorized into different types, among which sentential stress is the most salient one [2]. The definition of sentential stress varies in literatures. In Chu & Wang’s works [3, 4], sentential stress was classified into rhythmic stress and semantic stress. The former served the purpose of illustrating the rhythmic structure of an utterance and the later of making the speaker’s opinion or intention prominent. The validity of the classification has been proven by perceptual experiments. Their study shew that semantic stress tended to locate on the initial syllables within prosodic words, and rhythmic stresses on the final ones [4]. In another study on semantic stress in Mandarin, Wang *et. al.* [5] focused on stress distribution within three levels of units (e.g., sentence level; base phrase level; prosodic word level). Their analyses were carried out in 300 isolated reading utterances taken from Microsoft Mandarin TTS Speech corpus. They reported that the distribution tendency of semantic stress changed with the speech unit studied: within a sentence construction, semantic stress was more often distributed to the predicate or the object part (if there exists any) than to the subject part. However, such tendency did not hold within a base phrase. On the contrary, in a base phrase or a prosodic word, semantic stress was often found to be distributed to the adjunct part when they had an

adjunct plus a noun head or an adjunct plus a verb head structure.

Conclusions in [4] and [5] were drawn from the observation of independent sentences read with a neutral intonation (referred as a *Neutral* style in the remaining of this paper). However, in the real world speech, emotional coloring, which is realized by different choices of words, sentence types and ways to distribute stress and focus [6, 7], is an important part of the full meaning [8]. Gussenhoven’s study [9] of stress shift shew that stress shift as a rhetorical device was particularly frequent in propagandist speech, as used for instance in commercials and political speeches. Bolia and Slyh [10] worked on the relationship between perception of stress and speaking styles in monosyllabic or disyllabic words, and their result shew listeners correctly classified the utterance 58% of the time.

Chu and Bao [11] recently studied stress distribution tendencies within base phrases among different reading styles. Their results shew that the final-stressed tendency of rhythmic stress was valid in all the styles studied. The final-stressed tendency of semantic stress was shared in the prepositional (PP) and the predicate-object (PO) phrases among styles, while in cases of adjunct-verbhead (AV), adjunct-nounhead (AN) and subject-predicate (SP) phrases, variance occurred: the *Explanatory* style and the *Neutral* style shared a similar distribution tendency, having initial-stressed tendency in AN and AV phrases, but final-stressed tendency in SP phrases. Both the *Lyric* and the *Critical* styles differed from the *Neutral* style in phrases with the AV structure. Moreover, the *Lyric* style also had its own distribution of semantic stress in the AN and SP phrases.

In this paper, the comparison is extended to two other distinctive levels: the sentence construction level and the prosodic word level. The methodology used is rather similar to that used in [11] and is introduced in Section 2. The results and analyses are given in Section 3. Section 4 presents the final conclusion.

## 2. Methodology

In order to compare the stress distribution tendency, speeches read in three styles are collected and annotated. Details of the speech corpus used in this study are introduced in Subsection 2.1 and the indicators for stress tendency are described in Subsection 2.2.

### 2.1. The speech corpus and the annotations

The speech corpus used in this study is the same as that used in [11]. Seven articles, including two lyric essays by famous Chinese writers, two remarks (one on a newly-published novel and the other on a newly-drawn policy) and three objective illustrations (respectively on weather, stock and rules of law), are read by the same voice talent who also read the independent

---

<sup>1</sup> The work was carried out as a visiting student in MSRA.

sentences used in [3, 4, 5]. She was requested to choose a proper reading style for each type of articles according to her understanding of these articles. The reading styles corresponding to the three types of articles are referred as *Lyric*, *Critical* and *Explanatory*. According to [12], the validity of the division of speaking styles can be demonstrated by speech rate. In our corpus, the *Lyric* style is presented the slowest rate and the *Critical* style the fastest. The *Explanatory* style is read with a mezzo speed which is very close to that of the *Neutral* one in [5]. Both sentential stresses and structural properties are annotated manually in this corpus.

### 2.1.1. Labeling of sentential stresses

Two well-trained graduate students who major in linguistics perform the stress labeling. They are asked to identify all sentential stresses and assign a type (rhythmic, semantic or both) to each stress. As a result, each syllable in the 7 articles obtains one of the four stress labels, US — unstress, SS — semantic stress, RS — rhythmic stress, SRS — both semantic and rhythmic stresses. More details on the labeling process can be found in [11].

In [11], a structural property is assigned to each base phrase. In this paper, similar structural properties are assigned to sentence constructions and prosodic words.

### 2.1.2. Labeling of structural properties for sentence constructions

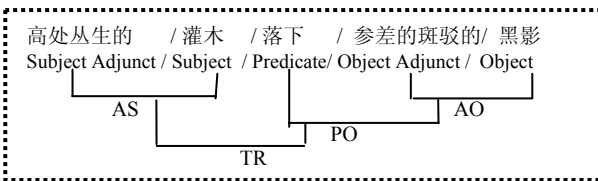
In order to investigate the distribution tendency of stress within sentences, six types of sentence constructions are studied, including adjunct-object (AO), adjunct-subject (AS), adjunct-predicate (AP), predicate-object (PO), subject-predicate (SP), theme-rheme (TR). All the seven articles are chunked manually and one of the following eight chunk types are given to each chunk identified. They are: sentence adjunct; subject adjunct; subject; predicate adjunct; predicate; complement; object adjunct; and object.<sup>1</sup>

### 2.1.3. Labeling of structural property for prosodic words

Since more than 80% of Chinese prosodic words have two syllables, the stress distribution tendency is only studied in disyllabic prosodic words in this paper. In Chinese, words normally have similar structures to phrases. All disyllabic words in the corpus are manually labeled with one of the eight structural properties, including subject-predicate (SP); adjunct-nounhead (AN); adjunct-verbhead (AV); predicate-complement (PC); predicate-object (PO); coordinative construction (CO); preposition phrase (PP); and root-affix (RA).

The structural labels obtained are illustrated in Figure 1.

(a) Sentence Construction Level:



(b) Disyllabic Prosodic Word Level:



<sup>1</sup>Each component of sentence constructions like PO, SP and TR, may involve more than one chunk.

Figure 1. An example of structural labeling in the sentence “高处丛生的灌木落下参差的斑驳的黑影” (“Tufty shrubs in the upland cast spotted irregular shadows.”)。

## 2.2. Indicators for stress tendency

For a certain type of sentence construction, the total number of syllables in one component may differ from that in the other. For instance, the subject adjunct chunk in AS (see Fig 1 (a)) has five syllables, while the subject chunk in the same construction has two. Therefore, the ratio between the two components does not show directly the initial-stressed or final-stressed tendency in that construction, because the ratio will be 5:2, if stresses are distributed normally among all syllables. Similar to the *SIW* used in [11], a *stress indicator for chunks (SIC)* is defined as the ratio of the number of stresses to the expected number of stresses as in equation (1) to describe the possibility for a certain type chunk to obtain sentential stresses.

$$SIC = N_r / N_p \quad (1)$$

where  $N_r$  is the number of stresses obtained by a type of chunk and  $N_p$  is the expected number of stresses for the type under the assumption that all stresses are distributed normally among all syllables in the corpus.  $N_p$  is calculated by (2) and (3).

$$N_p = N_w \times P \quad (2)$$

$$P = N_s / N_a \quad (3)$$

where  $N_a$  is the number of syllables in the studied corpus,  $N_s$  is the number of stressed syllables in the corpus,  $P$  indicates the possibility of a syllable to obtain a sentential stress under the normal distribution assumption,  $N_w$  is the number of syllables in a type of chunk.

$SIC > 1$  means that the possibility for the corresponding type of chunk to obtain stresses is above average, i.e. it has the tendency to obtain sentential stresses.  $SIC < 1$  means the opposite.

For illustrating the stress tendency within a certain construction, a *stress indicator ratio (SIR)* is defined as the ratio of  $SIC$  of the initial component to that of the final within a type of construction. If  $SIR > 1$ , the corresponding construction has the initial-stressed tendency, while, if  $SIR < 1$ , it has final-stressed tendency.  $SIR = 1$  means the two components have similar chance to be stressed.

Since the initial parts of disyllabic words always share the same number of syllables as the final parts, *stress ratio in word (SRIW)*, which is defined as the ratio of the number of stressed syllables in the two parts for a given word category, is used in studying the stress tendency with disyllabic words. If  $SRIW > 1$ , the type of words tends to be initial-stressed; if  $SRIW < 1$ , it tends to be final-stressed.

## 3. Results and Analyses

The stress indicators described in subsection 2.2 are calculated for sentence constructions and prosodic words. The results are presented in subsection 3.1 and 3.2 respectively.

### 3.1. Stress tendency within sentence constructions

$SIC$ s of RS distribution in chunks appearing in six sentence constructions are calculated for the three reading styles and given in Table 1(a). The corresponding  $SIR$ s of RS between the initial and the final components in each type of construction are given in Table 1(b).

Table 1. Stress indicators of RS in six constructions under three reading styles

(a) SIC for RS

Constr. type	Chunk property	Reading styles		
		Lyric	Criti.	Exp.
TR	theme	1.00	1.00	1.00
	rheme	1.00	1.00	1.00
PO	predicate	0.70	0.70	1.20
	object	1.20	1.30	1.00
SP	subject	1.20	0.60	0.70
	predicate	1.10	0.70	0.70
AO	adjunct	0.35	0.79	0.33
	object	1.41	1.91	1.50
AS	adjunct	0.00	0.32	0.42
	subject	1.33	1.61	1.18
AP	adjunct	0.47	0.67	0.28
	predicate	0.82	0.75	1.22

(b) SIR for RS

Construction type	Reading styles		
	Lyric	Criti.	Exp.
TR	1.00	1.00	1.00
PO	0.58	0.54	1.20
SP	1.10	0.86	1.00
AO	0.25	0.41	0.22
AS	0.00	0.20	0.36
AP	0.57	0.90	0.23

From Table 1, it is seen that the distribution tendencies of RS are quite similar across the three reading styles in most construction types, i.e. even distribution in TR and SP constructions and final-stressed tendency in AO, AS and AP constructions. *SIRs* for RS in AO and AS are mostly smaller than those in AP, which indicates chunks of AO and AS are more likely to be tightened up into one prosodic unit than those of AP. This conclusion is consistent with the one drawn in [13]. The main exceptions in RS distributions are in the PO construction. Final-stressed tendency appears in *Lyric* and *Critical* style, yet initial-stressed tendency in *Explanatory* style. *SICs* of SS distribution in the six sentence constructions are calculated for the three reading styles and given in Table 2(a). The corresponding *SIRs* of SS between the initial components and the final components in each type of construction are given in Table 2(b).

From Table 2, weak final-stressed tendency ( $SIR < 1$ ) is observed in TR, PO and SP constructions in most reading styles, i.e., when SS is distributed to these constructions, it often goes to the rhemes, objects or predicates. Exceptions lie in the TR, SP constructions under the *Lyric* style and the PO construction under the *Critical* style, where SS is uniformly distributed. Regarding the constructions of heads and their adjuncts, SS is also evenly distributed in the AP construction in all reading styles. However, within AO and AS constructions, the distribution varies across speaking styles. Adjuncts tend to be stressed in *Critical* style, while objects and subjects tend to be stressed in *Lyric* and *Explanatory* styles.

Table 2. Stress indicators of SS in six constructions under three reading styles

(a) SIC for SS

Constr. type	Chunk property	Reading styles		
		Lyric	Criti.	Exp.
T R	theme	1.00	0.80	0.90
	rheme	1.00	1.10	1.10
PO	predicate	0.50	1.00	0.60
	object	1.30	1.10	1.20
SP	subject	1.10	0.70	0.80
	predicate	1.20	1.00	1.10
AO	adjunct	1.05	1.20	1.23
	object	1.35	1.06	1.17
AS	adjunct	0.81	1.13	0.69
	subject	1.07	0.47	0.97
AP	adjunct	0.65	1.07	1.00
	predicate	0.64	0.92	0.99

(b) SIR for SS

Construction type	Reading styles		
	Lyric	Criti.	Exp.
TR	1.00	0.73	0.82
PO	0.38	0.92	0.50
SP	0.92	0.70	0.73
AO	0.78	1.12	1.04
AS	0.75	2.41	0.72
AP	1.02	1.16	1.01

Comparing relevant items about SS under the *Neutral* style in [5] with those under the three styles mentioned above, we find that the SS distribution under the *Explanatory* style resembles most to that of *Neutral* style: final-stressed tendency in PO and SP; normally distributed in AO and AP. The *Lyric* style shares similarity with the *Neutral* in all the cases except those in SP. The *Critical* style gains much more difference: SS is evenly distributed in PO; while initial-stressed tendency appears in AO and AS.

### 3.2. Stress tendency within prosodic words

*SRIWs* are only calculated for word types with more than 10 observations in the speech corpus. The results are listed in Table 3, in which, (a) is for RS and (b) is for SS. Syllables that obtain SRS are calculated repeatedly in both categories.

From Table 3, final-stressed tendency is consistently presented for RS and initial-stressed tendency for SS across all the three reading styles. The initial-stressed tendency for SS in AN and AV words indicates that adjuncts tend to gain more SS than heads within prosodic words. This observation is rather different from those obtained on the sentence construction level and the base phrase level. For PO words, Results are only available for *Explanatory* style and initial-stressed tendency is observed, which also differs from the final-stressed PO constructions in Table 2. All observations are in agreement with results drawn in [4].

Table 3. Stress tendency of prosodic words under three reading styles

(a) SRIW for RS<sup>1</sup>

Word type	Reading styles		
	Lyric	Criti.	Exp.
CO	0.24	0.06	0.22
AN	0.16	0.21	0.07
PO	-	-	0.43
AV	-	-	0.06

(b) SRIW for SS<sup>2</sup>

Word type	Reading styles		
	Lyric	Criti.	Exp.
AN	2.91	4.75	8.57
AV	+∞	1.33	2.63
PO	-	-	1.86
RA	+∞	-	-

#### 4. Conclusions and Discussions

This paper investigates the distribution of sentential stresses among three reading styles. The results show 1) in the sentence construction level, the final-stressed tendency for SS in TR, PO and SP is mostly valid, except cases of TR and SP under the *Lyric* style; PO under the *Critical* Style. Adjunct chunks in AO, AS and AP do not tend to acquire more SS except under the *Critical* style. Generally speaking, no special RS tendency occurs in TR, PO and SP in this level, but in the domain of adjunct-head constructions, final-stressed tendency for RS is observed. 2) Within disyllabic prosodic words, the final-stressed tendency for RS and the initial-stressed tendency for SS under different reading styles are in consistency with those of neutral sentences

Together with [11], the distribution of sentential stress in three levels has been discussed. Results show that speaking styles have fewer effects on RS distribution than on SS distribution. Such effects are more obvious in the sentence construction and the base phrase levels than the prosodic word level, where syntax plays a more crucial role in stress distribution.

Differences in stress perception among styles fit into the theory of ornate form, or rhetorical dualism [14]: ideas exist wordlessly and can be dressed in a variety of outfits, depending on the need for the occasion. To deliver the attitude of a speaker through speaking styles, listeners and speakers share a stress system as a convention in which listeners know to go to stressed items to find information which the speaker is particularly attentive to produce.

The *Lyric* style helps to express personal emotions in a rhythmic way. It weakens the effect of syntactic constraints and highlights the importance of rhyming. To satisfy the poetic

requirement, syllables with semantic stress are arranged deliberately evenly.

The *Critical* style is adopted to make comments, where the selection of verbs catches the speaker's attention. In the process of information delivery, semantic focus is presented by having more stresses in the predicate part. Consequently, the final-stressed tendency for SS in PO is trailed off under this style.

In *Explanatory* style, the speaker's task is to present messages clearly and concisely with an objective tone. This is also a regular way to deliver independent neutral sentences where syntactic constraints work actively. Hence, the tendency of SS distribution under this style is very close to that under the *Neutral* style.

#### 5. References

- [1] Trager, G.L.; Smith, H.L.J., 1957. An Outline of English Structure. *Studies in Linguistics: Occasional Paper 3*, American Council of Learned Societies, (5<sup>th</sup> printing)
- [2] Zhong, X.B.; Yang, Y.F., 1999. Foreign Researches on Prosodic Features and Stresses. *Acta Psychologica Sinica*, No. 4. (in Chinese)
- [3] Chu, M.; Wang, Y.J.; He, L., 2003. Labeling Stress in Continuous Mandarin Speech Perceptually. *In the Proceedings of the 15th International Congress of Phonetic Sciences*.
- [4] Wang, Y.J.; Chu, M.; He, L., 2003. Location of Sentence Stresses within Disyllabic Words in Mandarin. *In the Proceedings of the 15th International Congress of Phonetic Sciences*.
- [5] Wang, Y.J.; Chu, M.; He, L., 2003. Pilot Study of Semantic Stresses in Mandarin. *In the Proceedings of the 6th National Conference of Modern Phonetics*. (in Chinese)
- [6] Ken de Jong. 2000. Attention Modulation and the Formal Properties of Stress Systems. *The Proceedings from the Main Session of the Chicago Linguistic Society's Thirty-sixth Meeting, Volume 36-1*.
- [7] Edward P. J. Corbett. 1965. A Method of Analyzing Prose Style with a Demonstration Analysis of Swift's A Modest Proposal. *Reflections on High School English: NDEA Institute Lectures 1965*, pp106-124.
- [8] Laurence Perrine. 1956. *Sound and Sense*. Harcourt, Brace and Company, Inc.
- [9] Carlos Gussenhoven. 1983. Stress Shift in Dutch as a Rhetorical Device. *Linguistics* 21, pp 603-619.
- [10] Rober S. Bolia and Raymond E. Slyh. 2002. Perception of stress and speaking style for selected elements of the SUSAS database. *Speech Communication* 40, pp 493-501.
- [11] Min Chu; Mingzhen Bao. 2004. Comparison of Sentential-Stress Allocation within Base Phrases among Different Reading Styles. *Speech Prosody 2004*.
- [12] Justine F. et al, 2001. Prosodic Variation with Text Type. [http://chardonnav.elis.rug.ac.be/papers/2000\\_0001.pdf](http://chardonnav.elis.rug.ac.be/papers/2000_0001.pdf)
- [13] Chu, M.; Wang, Y.J.; Bao, M.Z., 2003. Local Grammatical Constraints and Length Constraints for Forming Base Prosodic Phrase in Mandarin. *In the Proceedings of the 6th National Conference of Modern Phonetics*. (in Chinese)
- [14] Louis T. Milic. 1965. Theories of Style and Their implications for the Teaching of Composition. *College Composition and Communication*, XVI, 66-69, 126

<sup>1</sup> Blank cells in Table 3 indicate no enough observations are available for certain cases.

<sup>2</sup> "+∞" means stress is always distributed to initial syllables without an exception.