Interspeech 2018
2-6 September 2018, Hyderabad

Stress Distribution of Given information in Chinese Reading Texts

Yuan Jia[1] and Xiaoxiao Ma[1,2]

[1] Institute of Linguistic, Chinese Academy of Social Science, Beijing, China
[2] Nankai University, Tianjin, China

summeryuan_2003@126.com, nkumxxiao@163.com

Abstract

Using an information structure annotation System, namely RefLex Scheme, the present study annotates the information structure of Chinese reading discourse and explores the relationship between information status and stress distribution. Our analysis results show that given information could bear stresses as well as new information. Specifically, the stress distribution of given information is significantly affected by the sub-category of information status on the referential level, i.e., r-given, while r-given-generic and r-given-displaced show different stress distributions. However, the sub-category of information status on the lexical level exhibits no such effect. Besides, the given information of proper nouns and personal pronouns on the lexical level can attract stresses. The reason is that a proper noun often serves as the topic of a sentence and a personal pronoun usually processes a center shift. Furthermore, the inconsistency of information status on both referential and lexical levels causes the stress on the given information unit.

Index Terms: information status, stress, given information.

1. Introduction

Language, as the most important tool of human communication, has the function of information transmission. Information structure refers to the distribution of language information. According to Halliday's theory of Information Structure, in the process of communication, each sentence can be viewed as a combination of information units. Each information unit consists of old information and new information. New information is “not recoverable from the preceding discourse”, while old information is “recoverable either anaphorically or situationally” [1].

Information structure and prosodic characteristics are closely related. In previous studies such asHalliday [2], Chafe [3], Brown [4]and Gussenhoven [5], they mainly studied English. It is assumed that new information will be stressed but given information is not. However, some scholars proposed that information status does not correspond to stress completely. Nooteboom and Kruyt [6] studied Dutch, and found that it is acceptable for listeners to stress expressions referring to new entities, whereas expressions that refer to given entities are also stressed sometimes. In monologue, Nakatani [7] found in American English, given entities expressed by pronouns could be stressed. In this case, she concluded, stress signaled focus center shifts. Sityaev’s [8] found that there is no one-to-one relationship between accentuation and information status in English. Many referents, which are given information units, may be accented too.

In response to the above inconsistent findings, this paper probes into the relationship between information status and stress in Chinese reading texts, and reaches the same conclusion as in Nooteboom [6], that given information can be stressed in some circumstances. We then further study the stress distribution of given information and investigate the conditions under which given information can be stressed. This will help gain a deeper understanding of the prosodic features and information structure of Chinese discourse. The rapid development of speech engineering has put forward a higher demand for phonetic research, which aims not only at the correctness of synthetic speech but more at the naturalness of the speech. This requires a systematic and comprehensive study of the prosodic features of language. In addition, our analysis of stress distribution is not based on the acoustic parameters, but on human auditory perception. Therefore, the experimental data obtained through the study would also help improve the naturalness of synthetic speech in the field of speech engineering.

2. Annotation system

2.1. Information structure labelling system: RefLex scheme

In our study, we labelled the information status of Chinese reading texts using an improved version of the Reflex scheme [9], which is adapted to the characteristics of Chinese, for Lexical-Level (L-level) and Referential-Level (R-level) separately. The information status of a word on L-level was determined by its form, while that on R-level was determined by its meaning. The judgment criteria we used is explained in more detail in following paragraph. Since one expression may refer to different entities and one entity may be expressed by different expressions, it is necessary to distinct the information status from both L-level and R-level.

From a cognitive perspective, Riester & Baumann [9] classified the information status, according to the degree of activation in listener’s mind. On L-level, information status of the reference is classified into three categories: new information (L-new), accessible information (L-accessible) and given information (L-given). Specifically, if one expression is identical, synonymous, hypernymic, or holonymic with/to an expression in previous texts, it is labeled as L-given. For R-level, information status of the referents are also classified into three categories: new information (R-new), bridging information (R-bridging) and given information (R-given): if one referent is present in previous discourse, it is labeled as R-given. Then, the information status on both levels further branches into finer categories in different contexts.

2.2. Chinese prosodic labeling system: C-ToBI (Chinese-Tone and Break Indices)

We used the C-ToBI system for Chinese prosodic annotation. The annotation file includes 4 tiers: (1) Shengyun (initial-

10.21437/Interspeech.2018-1602

final); (2) Pinyin; (3) Break indices; (4) Relative stress. The Relative stress tier marks the stress level of each prosodic unit, falling into four levels: 0,1,2,3, which respectively represent the unstressed syllables, the heaviest syllable in a prosodic word, the heaviest syllable in a minor prosodic phrase, and the heaviest syllable in a major prosodic phrase. The annotation of the relative stress tier is judged based on annotators' perception, and the consistency of the annotation was up to 87% [10]. Nonetheless, there are still personal differences in annotation, which were resolved through maximum voting.

3. Corpus and method

3.1. Corpus

The research object of this paper is a dataset of eighteen Chinese reading texts. They are selected from the ASCCD corpus (Annotated Speech Corpus of Chinese Discourse), supported by the phonetic reference room of the Language Institute of the Chinese Academy of Social Sciences. The dataset includes materials concerning texts, sound and phonetic annotation. The texts were carefully selected by linguists, covering different writing styles such as narration, argumentation, prose. The sound materials were produced by ten subjects (five male and five female) from Beijing, proficient in standard Chinese mandarin, and were annotated by phonetics experts. The sound file was digitized as the WAV format with a sampling rate of 16khz, 16bit resolution, and on a dual channel.

3.2. Experimental method

In this paper, we use Praat [11] script to extract the annotated information structure and its corresponding stress distribution. We get 10825 data on L-level, with 5076 labeled given information; and 10908 data on R-level, with 4776 labeled given information. Then, we use SPSS 23 to further analyze the data collected and organized with Microsoft Excel.

Firstly, we examine the relationship between stress distribution and information status. Given that the annotation of the 0-level stress and 3rd-level stress across different annotators reaches a high consistence, we only take into account these two levels. Then, we further investigate the stress distribution of given information on R-level and on L-level respectively. Finally, we investigate the influence of information status inconsistency on given information's stress distribution.

4. Results

4.1. Information status and stress distribution

To study the relationship between information status and stress, we first sum up the stress distribution of different information status on L-level and R-level, and the results are shown in Figures 1.

Note that the figures in this paper, including Figures 1 to 3, show how different information status receives 0-level and 3rd-level stress respectively. In each figure, the horizontal axis indicates the information status, and the ordinate axis indicates the percentage of 0 level, 3 level stress that different information status receives.

From figure 1, We can see that on both L-level and R-level, the proportion of the 0-level stress (unstressed word) increases in turn for new information, accessible/bridging information and given information. However, the ratio of 3-level stress decreases in turn. This result shows that accessible/bridging information and new information are more inclined to receive stress, especially higher level stress, than given information does.

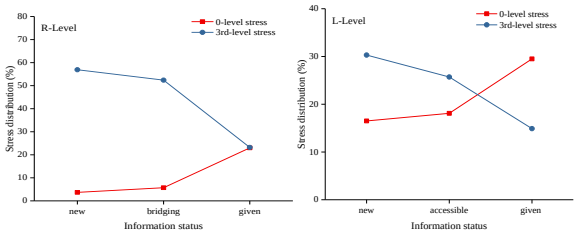

Figure 1: *Stress distribution on R-level (the left) and L-level (the right).*

The Chi-square test shows that there are significant differences in stress distribution of different information status on R-level (χ_2 (2 n=4688) = 1023 P<0.01) and on L-level (χ_2 (2 n=4921) = 459 P<0.01).

4.2. Given information and stress distribution

As is shown above, given information can be stressed, but we would like to find out under what condition, it can be stressed. Thus we further study the stress distribution of given information on L-level and R-level separately.

4.2.1. Given information and stress distribution on R-level

On R-level, given information branches into three categories: r-given, r-given-displaced, r-given-generic. If an expression is co-referential with an antecedent in the previous discourse, it is r-given; if the co-referential antecedent of an expression occurs earlier than the previous five sentences, the expression is r-given-displaced; if an expression refers to the generic definite or indefinite of the reference that is introduced, it is r-given-generic.

Accordingly, we anticipate that there will be a significant difference between the sub-categories of given information and stress distribution. r-given-displaced references can receive more 3rd-level stress than r-given references do, which corresponds with human memory rules: immediate memory holds the information for less than 1 second, and then it either forgets or passes the information on to short-term memory; the retention time of short-term memory is only 5-20 seconds without retelling, and the longest is not more than 1 minutes. As is mentioned above, a r-given-displaced reference is five sentences away from a previous reference, thus speakers tend to stress it so as to remind the listener of the entity mentioned before. By Chi-square test, we analyse the relation between stress degree and the sub-categories of given information. The statistical results (χ_2 (2 n=2211) =106.236 P<0.01) are in line with our conjecture. Besides, we investigate the stress distribution pairwise. Stress distribution of r-given and r-given-generic reference show a significant difference (χ_2 (1 n=2073) =23.195 P<0.01). The same is found for r-given and r-given-displaced references (χ_2 (1 n=1907) =73.661 P<0.01) and for r-given-generic and r-given-displaced references (χ_2 (1 n=442) =105.971 P<0.01). The stress distribution of given information is shown in the following figure.

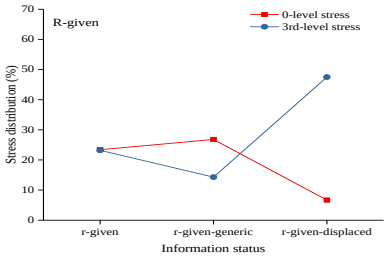

Figure 2: *0-level stress and 3rd-level stress that given information receives on Referential Level.*

According to Figure 2, we see that the percentage of 0-level stress that r-generic-displaced reference receives is 6.7%, lower than that of r-given, r-given-generic reference receives. The percentage of 3rd-level stress is 47.5%, which is the highest among the three categories. In summary, we observe that, on Referential Level, the stress distribution of given information is significantly affected by the information status sub-category of reference and language context.

4.2.2. Given information and stress distribution on L-level

For a more focused investigation, we confines the annotation object of L-level to nouns and pronouns. On L-level, the given information branches into four sub-categories: given-same, given-syn, given-super, given-whole (l-given-same, l-given-syn, l-given-super, l-given-whole for short). l-given-same expresses the recurrence of same word; l-given-syn indicates the "relation between words at the same hierarchical level"; l-given-super indicates that a word is lexically superordinate to previous word in the sense that the markable is a hypernym of the antecedent expression; l-given-whole shows that a word is lexically superordinate to previous word in the sense that the markable is a holonym of the antecedent. [9]

Similarly, from a cognitive perspective, we predict that different sub-categories of given information receive stress differently. Compared with other three categories of given information, l-given-same references may receive stress least possibly. An intuition is that it will be easier for people to understand a word which has appeared before than to understand a different word which relates to a previous word semantically. However, the result of Chi-square test (χ_2 (3 n=2,279) = 5.143 P>0.05) shows that there is no significant difference between stress distribution and sub-categories of given information on the lexical level. This drives us to look at the stress distribution of given information on L-level from different perspectives. We try to find out if there is some commonality among given information marked 3rd-level stress.

To this end, we mainly investigate how proper nouns and personal pronouns labeled as given information receive 3rd-level stress. In the 18 texts, 96 proper nouns were labeled as given information, and 68.8% of them received 3rd-level stress. Table 1 presents some examples.

Table 1: *Frequency of word labeled given information in the text and frequency of them receiving 3rd-level stress.*

Word	Given information	3rd-level stressed
Liu Xiaoguang	10	9
Shu Yi	10	6
Sun Qingfu	8	5
Song Sumei	8	6
America	9	4

Why do proper nouns bear 3rd-level stress at such a high percentage? We think that accenting 'given' information may have something to do with topicality [8]. For instance, the most frequently repeated referent "Liu Xiaoguang", in text *Liu Xiaoguang wakes from "hibernation"*, is mentioned 10 times as given information unit, and 9 of them were stressed, while 8 of the 9 are subjects and topics of the corresponding sentences that they appear. Nooteboom and Kruyt [6] proposed that, sometimes, speakers would like to use intonational focusing to highlight the theme or topic of a sentence.

As for the personal pronouns, in the 18 texts, 63 were labeled given information, and 42.9% of them received 3rd-level stress. Some examples are presented below.

(1) Yin1 er2 lang2 shi4 zui4 ke3 pa4 de1 dong4 wu4.
Wo3 shi4 yong3 sheng1 ming2 ke4 yu2 xin1 de0.
Wolf is the most horrible animal, and I'll keep it in my mind all my life.

(2) Cao2 xun1 xuan4 shu1 zai4 xia4 de0 guo4 fen4.
Wo3 ying2 zai4 ba3 wo4 zhu4 le0 shi2 ji1.
Cao Xunxuan lost the competition for he played it too radically, while I won because I grasped the chance.

We try to explain why given personal pronouns were stressed by speakers under the centering framework elaborated by Grosz, Joshi and Weinstei [12]. For example, in (1), the referent "lang2" refers to "lang2 qun2" that was mentioned earlier, thus it is the "backward-looking center" of the first sentence; meanwhile, it is also the "forward-looking center" of the second sentence. However, in second sentence, the reference "wo3" is the subject and also the "backward-looking center" of the second sentence. Thus the "backward-looking center" of the two sentences shifts from "lang2" to "wo3". As a result, the two sentences are not so coherent semantically. In order to communicate smoothly, speakers may stress the personal pronoun to indicate that there is a center shift so that listeners can pay more attention to it. In addition, in (2), the stressed word "wo3" is the contrastive focus. Our result can support previous claims that personal pronouns tend to be deaccented; yet, when accented, the signal centre shifts from one reference to another [8]. As Fang [13] mentioned, the contrastive focus component is always accompanied by mandatory contrastive stress in spoken language.

On the one hand, on R-level, a personal pronoun is usually labeled as given information, however when it refers to a person which is introduced for the first time, we label it new information. On the other hand, personal pronouns on L-level would be labeled as given information if the reference has appeared before, disregarding whether it refers to the same person or not. This may cause a dilemma that a personal pronoun is marked as given information on L-level, but is marked as new information on R-level. We will discuss the impact of the inconsistency in following subsections.

4.3. Influence of information status inconsistency on stress distribution

Although the distinction between information status is made according to the degree of activation of information in mind, in our study, information status on L-level was determined by their forms whereas on R-level by the meaning of the referents respectively. This may cause the inconformity of an expression. For example,

(3) zhe4 wei4 nv3 shi4 xiang4 Shu1 Yi3 shuo1 ming2 yuan2 yin1: wo3 jiao4 Wang1Cun2 yi1. Ta1 jiao4 Peng2Jun1

This lady explained to Shu Yi that: I am Wang Cunyi, she is PengJun.

"ta1" in this sentence refers to the person "Peng2 Jun1", which is introduced for the first time in the text, so we labeled it as new information on R-level. However, the word "ta" has appeared in previous text, it refers to another person "Wang1 Cun2 yi1", we labeled it as given information on L-level.

We expect that this phenomenon will influence the distribution of stress. To verify the hypothesis, we further study the relationship between stress distribution and information status of the reference on L-level, when it is labeled as given information on R-level, and study the relationship between stress distribution and information status of the reference on L-level, when it is labeled as given information on R-level.

The Chi-square test shows that there are significant differences under the conditions mentioned above. For given information on R-level, their information status on L-level significantly influences the stress distribution (χ_2 (2 n=2,287) = 94.817 P<0.01). Similarly, for given information on L-level, their information status on R-level significantly influences the stress distribution (χ_2 (2 n=2,165) = 19.176 P<0.01) too. Then, for R-given reference, the effects of information status (L-new, L-given, L-accessible) on stress distribution are investigated, and we find that the stress distribution between every pair of information status is statistically significant. Stress distribution of L-given and L-accessible reference shows a significant difference (χ_2 (1 n=1603) = 65.101 P<0.01), and the same is found for L-given and L-new references (χ_2 (1 n=2110) = 40.415 P<0.01) and for L-new and L-accessible references(χ_2 (1 n=661) = 29.576 P<0.01). These results indicate that, for R-given references, their information status on L-level affects the stress distribution.

As for L-given reference, the cross-effects of the three kinds of information status (R-given, R-new, R-bridging) on stress distribution are investigated. Significant differences are found in stress distribution between R-given and R-bridging (χ_2 (1 n=1892) = 12.448 P<0.01) as well as between R-given and R-new (χ_2 (1 n=1789) = 9.616 P<0.05). However, no significant differences are found between R-new and R-bridging (χ_2 (1 n=640) = 0.000 P>0.05). The results indicate that for L-given references, their stress differs greatly when they are labeled differently as R-given or R-bridging and when they are labeled differently as R-given or R-new. Figure 3 shows the stress distribution of references labeled as different information status on L-level and on R-level. The horizontal axis in Figure 3 represents the information status of R-given reference on lexical level (the left), the information status of L-given reference on referential level (the right). The ordinate axis indicates the percentage of 0 level, 3 level stress that different information status receives in both figures.

We can see that R-given references labeled as L-accessible receive 0-level stress with the least proportion (5.5%), while they receive 3rd-level stress with the highest proportion (64.5%), much higher than that they receive when labeled as L-new and L-given. At the same time, we can see that L-given references receive more 0-level stress than 3rd-level stress, no matter they are labeled as R-new, R-bridging or R-given on referential level. This indicates that compared with given information on R- level, given information on L-level are more inclined to be unstressed.

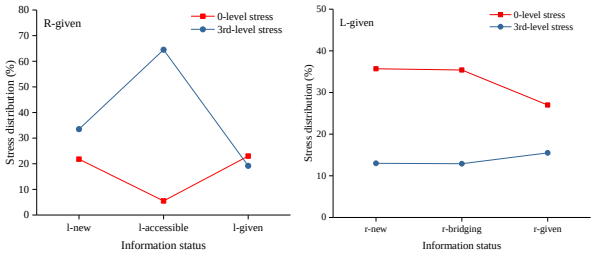

Figure 3: *The stress distribution of R-given references on L-level (the left) and of L-given references on R-level (the right).*

5. Discussion & Conclusion

Based on the statistical analysis of a spoken corpus in Mandarin Chinese, we have found that information structure does affect the distribution of stress, yet there is no one-to-one correspondence between the two, and some given information can also be stressed.

Through a series of further analysis, we find that, first, on R-level, categories of the given information also affect the distribution of the stress. This is consistent with the characteristics of human memory, but there is no significant difference in the stress distribution across different L-given categories. Second, on L-level, the given information of the proper nouns and personal pronouns can be stressed. This is because that a proper noun often serves as the topic of a sentence and a personal pronoun processes a center shift. Third, the inconsistency of information status on both referential and lexical level causes the stress on the given information unit. An implication of these results is that the influence of information structure on stress distribution could be related to the cognition of human speech. We think it is sensible to classify the status of information from the cognitive perspective. For example, as demonstrated in Figure 4, since given information refers to the entities activated already in human's mind while accessible information refers to the entities that are half-activated, thus the difficulty of inference increases. As a result, speakers would tend to stress it to remind the listeners. Therefore, stress is a useful strategy to attract listeners' attention. This complies with cognition rules. Meanwhile, for the sake of understanding, speakers may use the words appearing before to readjust the information structure of the discourse rationally. However, when they are not allowed to use the words repeatedly, speakers would adopt other strategies (such as stressing the words) to reduce listeners' reasoning process and difficulty, in order to deliver the information accurately.

Finally, there are some aspects for further research. The annotation of the stress level of given information is based on annotators' perception. As a consequence, the subjectivity is inevitable. How to improve the consistency of annotations is of importance in our further study. Besides Chinese reading texts whether there is a similar relationship between information status and stress in natural spoken language still needs to be explored.

6. Acknowledgements

This research was supported by National Key R&D Program of China (2017YFE0111900), Key Project 'Interaction of Grammar, Semantics and Prosody' of National Social Science foundation under grant 16AYY016 and Innovation Program of Chinese Academy of Social Sciences.

7. References

[1] M. A. K. Halliday, “Notes on transitivity and theme in English (part 2),” *Journal of Linguistics*3: 199-244,1967.

[2] M. A. K. Halliday, “Intonation and Grammar in British English,” *The Hague: Mouton,*1967.

[3] W. Chafe, “Language and consciousness,” *Language* 50:111-133, 1974

[4] G.Brown, “Prosodic structure and the given/new distinction,” In Cutler, A. & R. Ladd (eds.).Prosody: Models and Measurements. 67-77. Berlin: Springer Verlag,1983.

[5] C. Gussenhoven, “On the Grammar and Semantics of Sentence Accents,” *Dodrecht: Foris,*1984.

[6] S. G. Nooteboom and J. G. Kruyt, “Accents, focus distribution, and the perceived distribution of given and new information: An experiment,” The Journal of the Acoustical Society of America, vol. 82, no. 5, pp.1512-1524, 1987.

[7] C, Nakatani, “Constituent-based accent prediction,” *Proceedings of ACL/COLING,*939-945.*Montreal: ACL,*1998.

[8] D. Sityaev, “The relationship between accentuation and information status of discourse referents: A corpus-based study,” *UCL Working Papers in Linguistics*, 2000.

[9] A. Riester and S. Baumann, RefLex scheme–Annotation guidelines,2014.
URL: http://www. ims. uni-stuttgart. de/institut/mitarbeiter/arndt.

[10] Z. G. Yin, “On the rhythm of read speech in Mandarin,” Doctoral dissertation of Chinese Academy of Social Science, Beijing, China, 2011.

[11] P. Boersma and D. Weenink. "Praat: doing phonetics by computer (Version 4.6.33)." (2007).

[12] B. Grosz, A. Joshi & S. Weinstein (1995) Centering: A framework for modeling the local coherence of discourse. Computational Linguistics 21(2): 203-225.

[13] M. Fang, “Syntactic representation of Chinese contrast focus”. Zhongguo Yuwen, 1995(4):279-288.