Speech Prosody 2016
31 May - 3 Jun 2106, Boston, USA

Influence of Dependency Parsing on the Prosody of Chinese Discourse

Yu Pang [1,2], *Yuan Jia*[2], *Aijun Li*[2], *Dawei Song*[1], *Ruifang He*[1]

[1] Tianjin University, Tianjin, China
[2] Institute of Linguistics, China Academy of Social Sciences, Beijing, China
conniecoding@gmail.com, summeryuan_2003@126.com, liaj@cass.org.cn,
dawei.song2010@gmail.com, rfhe@tju.edu.cn

Abstract

Dependency parsing has been a prime focus of Natural Language Processing. The present paper was attempting to elucidate the relationship between prosodic variation and syntactic structure within the framework of dependency parsing. The study adopted Harbin Institute of Technology (HIT) dependency parsing tool to annotate Chinese spoken discourse corpus. Duration variation and stress distribution patterns were employed as parameters to examine the interactions of the syntactic structures and the prosodic features. Preliminary results demonstrated that there was an intrinsic association between duration variation and dependency relation. Specifically, governor duration was more likely to be longer than dependent duration of dependency relations of Left-Adjunct, Mood-Tense, Attribute, Quantity and so on. Moreover, in "DE" Construction, Preposition-Object and Verb-Object, the dependent duration was much longer.

Index Terms: dependency parsing, duration variation, stress distribution

1. Introduction

Previous research in the area of prosodic study has converged on a set of prosodic correlates of discourse structure. With regard to prosodic research on standard Chinese, much attention has been paid to the concatenation of lexical tones, the effect of focus, and the realization of intonation based on following frameworks. The development of the study between Chinese discourse and prosody can be outlined as following frameworks.

HPG. Based on the point of view of hierarchical prosodic phrase grouping (HPG), which is a prosody framework to account for discourse prosody organization, Tseng [1, 2, 3] predicts cross-phrase F0 contours, duration patterns, intensity distribution and pause insertions in accordance with prosody organization. By using the adaptive threshold and modified normalization, their model is more robust. In particular, the better prediction achieved in pauses or breaks makes it possible to develop software towards locating and labeling prosody breaks not independently but in relation to prosody organization.

RST. On the basis of Rhetorical Structure Theory (RST) framework, Yang [4, 5] examines the relationship between prosodic variation and rhetorical structure in discourse. They propose that rhetorical structures of Chinese discourse, such as rhetorical hierarchy, rhetorical relations and nuclearity, are reliably conveyed by prosodic features. Besides, Zhang [6, 7] investigates the interface of rhetorical and rhythmic aspect of several Chinese reading discourses and states: the stress degree in both pre- and post- pausal positions has a significant impact on pause duration. What's more, the nucleus in pre-pausal and satellite in post-pausal positions can significantly lengthen the pause duration.

Anaphora. In terms of anaphora which is an interface between the phonetic level and the levels of syntax, semantics and pragmatics, Hou [8, 9, 10] conducts a pioneering study on the phonetic manifestation. Their main findings are: i) For zero anaphora, the boundary syllable after the zero anaphoric form is more accented; ii) For pronominal anaphora, the phonetic manifestation of the third personal pronominal anaphoric form "ta" is closely related with syntactic position; iii) For nominal anaphora, the nominal anaphoric form is less accented than the antecedent; the duration of the pause before the antecedent is significantly longer than that before the nominal anaphoric form.

Centering Theory. Depending on the framework of Centering Theory, Dong's works [11] systematically investigate the influences of different transition states on prosodic features of Chinese reading texts. These results suggest that the transition state has significant impacts on pause duration and stress degree combination pattern.

Dependency parsing is known as a syntactic or a shallow semantic analysis of Natural Language Processing. Wang [12] and Liu [13] conduct an interface study between syntax (dependency) and prosody (stress). Wang's work [12] statistically analyzes stress distribution across 24 dependency relations (*DRs*) for a Chinese spoken discourse corpus, proposing that the intonation stress is more likely to appear in DRs of Subject-Verb (SVB), Attribute (ATT), Adverbial (ADV) and Verb-Object (VOB). Liu [13] explores intonation stress distribution from a refined dependency parsing scheme with 26 DRs based on a spoken dialogue. Comparing Liu's [13] with Wang's [12] works, the stress distribution under spoken dialogues and spoken discourse across DRs is rather consistent, but the stress distribution within each relation is quite different.

On the basis of researches mentioned above, it is obvious that the study on the influence of dependency parsing on the prosody, which has been adopted in interdisciplinary research, is mainly restricted in the relationship between stress and DRs. The empirical study on the interface of prosody and DRs, from the perspective of duration within each DR, is of fundamental importance but largely under-explored. The work reports in this paper aimed at solving these problems. Particularly, this paper does not only concern with the duration variation within each DR in a Chinese spoken discourse corpus, but also has an in-depth discussion the association between stress and duration within DRs.

2. Data

The materials used throughout this study are Annotated Speech Corpus of Chinese Discourse (ASCCD) which was built by the

 doi: 10.21437/SpeechProsody.2016-78

Phonetic Lab, Institute of Linguistics, Chinese Academy of Social Sciences (CASS). The reading materials included Narrative, Descriptive and Exposition discourse. The data were collected from ten Mandarin speakers (5 males and 5 females) in Beijing [14, 15]. The length of the discourse was approximately 550 words on average. Most of the sentences were complicated sentences. The C-ToBI system was used for annotation and four tiers were labeled.

The corpus contained prosodic annotation information including prosodic structure, duration and stress annotation [16]. Four hierarchical boundaries were annotated as prosodic word, prosodic phrase, intonation phrase and prosodic group boundary. Besides, they were denoted by "B1, B2, B3 and B4" respectively. As for the stress, we followed the process that Wang et al. [12] did. Three level stress were annotated for prosodic word stress, prosodic phrase stress and intonation phrase stress which were denoted by "S1, S2 and S3" respectively.

We used HIT dependency parsing tool. In the practice of annotation, two steps were followed: automatic parsing and proofreading manually. Table 1 listed some kinds of relations and their corresponding explanations.

Table 1. Some kinds of DRs.

Tags	Details	Examples
SBV	Subject-Verb	小明寻找失物(Xiao Ming finds lost)。(小明(Xiao Ming)←寻找(finds),SBV)
VOB	Verb-Object	总理接受采访(Prime accepts interview)。(接受(accepts)→采访(interview),VOB)
ATT	Attribute	小明找各种借口 (Xiao Ming finds a variety of reasons) 。(各种(a variety of)←借口(reasons),ATT)
ADV	Adverbial	他跑得很快 (He runs very fast) 。(很(very)←快(fast),ADV)
CMP	Complement	他跑得很快 (He runs very fast) 。(跑(runs)→得(de),CMP)
DEI	"DEI" Construction	他跑得很快 (He runs very fast) 。(得(de)→快(fast),DEI)
DE	"DE" Construction	心理上的准备 (mental preparation) 。(上(shang)←的(de),DE)
DI	"DI" Construction	他飞快地跑出去了(He ran out quickly.)。(飞快(quickly)←地(de),DI)
MT	Mood-Tense	我吃完饭了(I've eaten my dinner.)。(吃(eaten)→了(le),MT)
QUN	Quantity	北京是一座漂亮的城市 (Beijing is a beautiful city.) 。(一(a)←座(zuo),QUN)
AB	Attribute-Backwards	高级绿茶，中国制造(Senior green tea, Chinese manufacturing)。(绿茶(green tea)→制造(manufacturing),AB)
OF	Object-Forward	我渴了用英语怎么讲 (How to say I'm thirsty in English.) 。(渴(thirsty)←讲(say),OF)
…	…	…

Dependency parsing, which can signify the relation between two words, was built on word segmentation and parts of speech tagging. After this work, the dependency parsing would output the dependency relation between two units of word segmentation.

In the HIT dependency scheme, each relation was composed by three parts: the governor denoted as B, the dependent marked as A and the name of the relation-Label, expressed as (A←B, Label). For example, "小明寻找失物 (Xiao Ming finds lost)". The result of word segmentation was: "小明 (Xiao Ming) 寻找 (finds) 失物 (lost)". The dependency relation of "小明 (Xiao Ming)" and "寻找 (finds)" was SBV. "小明 (Xiao Ming)" was the dependent, and "寻找 (finds)" was the governor. The tag was: (小明 (Xiao Ming) ←寻找 (finds), SBV). Figure 1 displayed the dependency parsing result. After the automatic parsing for all 24 discourses, we proofread the result manually.

Figure 1: The *result of dependency parsing*.

3. Methodology, results and analysis

In this section, the study systematically examined the variation of duration in each DRs, and possibly influential factors to duration of dependency units. The variation of duration was extracted first and then statically analyzed within each DRs. Besides, we discussed the association between stress and duration within dependency relations, comparing the percentage of the governor duration which was longer than dependent duration with the percentage of the stress that the governor gained within each dependency relation.

We annotated duration for each syllable. When it came to the duration extraction, we paid particular attention to the data normalization process: first, according to the DRs, dependency units were classified; second, we extracted the dependency units with the same numbers of governor words and dependent words; third, in order to analyze the variation of duration, we normalized the dependent units which had different words. In terms of the stress, we extracted stress distribution corresponding to the DR.

For an illustrative purpose, we measured four kinds of percentages for comparison. These percentages were defined as following:

- P_{gd}: The percentage that governor duration was longer than dependent duration in dependency units.
- P_{gs} : The percentage that governor gained stress in dependency units.
- P_{dd}: The percentage that dependent duration was longer than governor duration within dependency units.
- P_{ds} : The percentage that dependent gained stress in dependency units.

3.1. Distribution of the DRs

To study the influence of DRs on the prosody of Chinese discourse, the first step was to comprehend the distribution of each dependency relation in the ASCCD corpus. ASCCD corpus had 54130 DRs. The distribution of each DR was

showed in Figure 2 where the first number meant the number of the dependency relation and the second number represents the percentage that the relation accounted for. As shown in Figure 2, ATT relation reached 19% across all the relations, which happened in the most. ADV and VOB relations took up 17% and 13% respectively. SBV relation took up 10% etc. This result was consistent with Wang's [12] research.

Moreover, the result (Figure 2) described that the DRs of ATT, ADV, VOB and SBV accounts for more than half of the total. Obviously, they played a crucial role within all these DRs.

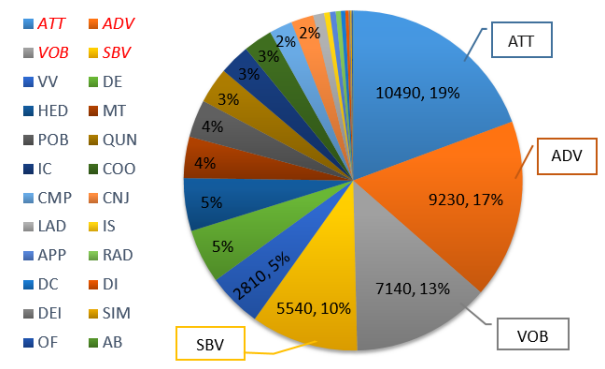

Figure 2: *Distributions of dependency relations.*

3.2. Duration variation within twenty-four dependency relations

To better observe the duration variation, all of P_{gd} taking place in each relation were counted and the result was illustrated in Figure 3. However, there were some relations removed because the total numbers of them were too small, like DEI, DI, OF, SIM and AB relations. In addition, the HED relation was also removed since no duration variation showed up at the governor (a sign of the ending of one sentence) in HED relation.

In Figure 3, the abscissa represented the rest 18 DRs, and the ordinate represented P_{gd} took place in each relation. It was displayed clearly that duration variation within each relation. If the percentage was higher than 50%, it meant that the governor duration was more likely to be longer than dependent duration. On the contrary, if the percentage was lower than 50%, it represented that the dependent duration was more likely to be longer than governor duration.

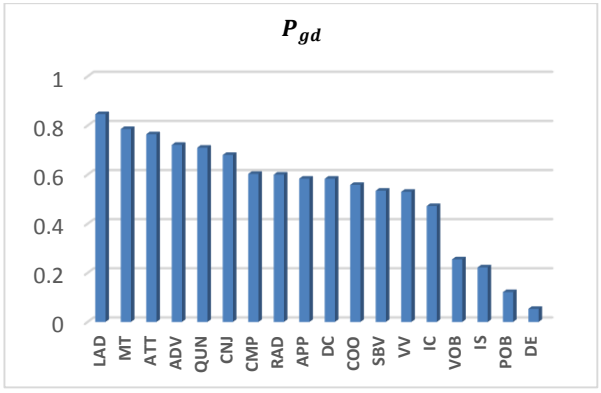

Figure 3: *The percentage of governor duration is longer than dependent duration (P_{gd}).*

The result in Figure 3, showed a significant difference between governor duration and dependent duration within each DRs. In Left Adjunct (LAD), Mood-Tense (MT), Attribute (ATT), Adverbial (ADV), Quantity (QUN), Conjunctive (CNJ), Complement (CMP), Right Adjunct (RAD) and Appositive (APP) relations, governor duration were more possibly longer than dependent duration. On the contrary, In Verb-Object (VOB), Preposition-Object (POB) and "DE" Construction (DE) relations, dependent tended to get longer duration. In Coordinate (COO), Subject-Verb (SBV) and Verb-Verb (VV), the governor and dependent had the same possible for the duration variation. The statistical result also stated that there was an intrinsic association between duration variation and DRs.

For LAD, MT, ATT, ADV, QUN, CNJ dependency relations, Figure 3 displayed that the governor had extremely high possibility to get longer duration than the dependent. It can be seen that modal particles and adjuncts, such as "并且（and）" and "吗（ma）", hardly got longer duration (from LAD, CNJ and MT and relation). It was consistent with previous research [2]. The same phenomenon that the governor duration was longer than the dependent took place in ATT and ADV relations. Such a conclusion was also in accord with most scholars' ideas [17, 18, 19] that the adverbial modifier was inclined to gain more information. In numerical phrase, duration distribution was more likely to be shorter at the numerals (from QUN's data).

In VOB, POB and DE relations, the governor had less chance to get long duration than dependent. Compared with governor, the longer duration tended to appear at the objects in verb-object structure to some extent (from VOB's data). For the POB's data, it can be seen that long duration was more likely to be shown at the objects (dependent). As for the DE relation, it always appeared with ATT relation. The result suggested that longer duration was inclined to appear in the attribute in ATT structure.

As for the CMP, RAD and APP relations, the probability that the governor duration was longer than dependent duration was a little more than 50%. It showed that in RAD relation, the adjuncts words(dependent), such as "等等（and so on）", hardly got long duration. In consideration of verb-complement structure, there was more possibility for verbs (governor) to gain long duration. Based on APP's data, it can be noted that longer duration was prone to appear at the former position.

The same possibility that the governor and the dependent gained longer duration hold in COO, SBV and VV relations. There were three relations between two clauses, namely, IC, DC and IS. However, present statistics had not considered the clause relation.

Combined with the result in Figure 2, the total percentage of relations which had significant variation of duration (LAD, MT, ATT, ADV, QUN, CNJ, VOB, POB, and DE) reached as high as 69%, which demonstrated that there was an intrinsic association between duration of each dependency unit and dependency relation.

3.3. Duration variation and stress distribution

This part examined the variation of duration in the combination of stress within 18 DRs. Figure 4 showed that P_{gd} and P_{gs} took place in each relation and proposed such a result that the percentage of duration variation and the percentage of stress distribution had substantially the similar trend within some DRs.

It was clear to see that in most of the dependency relations (LAD, MT, QUN, CNJ, CMP, RAD, DC, COO, IC, VOB, IS, POB and DE), the variation of duration and stress distribution between governor and dependent had similar percentage distribution. In other words, the higher P_{gd} meant the more possibility the governor had to obtain stress, which conformed to previous studies [17, 18, 19].

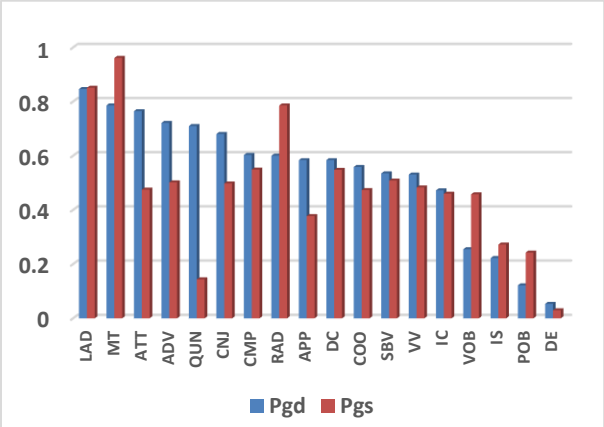

Figure 4: P_{gd} *and* P_{gs} *take place in each relations.*

However, there were some conclusions of relations which were unusual, such as ATT and ADV, it was obvious that P_{gd} accounted for larger proportion than P_{dd}, but it had the close proximity with P_{gs} and P_{ds}. This result was quite out of accord with most scholars' view that adverbial modifier was inclined to gain stress in the adverbial structure for the cause of adverbial modifier which tended to gain more information [17, 18, 19]. The surprising result may be caused by the factor that the adverbials located at the first place of one sentence hardly obtain stress [20]. In numerical phrases, the stress was more likely to occur at the numerals (dependent), but the governor was more likely to gain longer duration. The same phenomenon was found in APP relation, where the stress was prone to appear in the latter position (dependent), whereas the former position (governor) had more probability to get longer duration than the latter position. In SBV and VV relations the governor and the dependent had the same possibility in duration assignment, but the stress distribution tended to show up at the latter word. In order to find the deep relation between stress and duration within VV and SVB relations, we found that the stress tended to show up at the latter word when the duration of the latter word was more likely to be longer than another word.

Based on the results mentioned above, it was undoubted that duration variation made contributions to correctly predict stress distribution within most of the relations, indicating that the longer the duration was, the more possibility it obtained stresses.

4. Discussion

The obtained results in this paper showed that the duration variation was significantly related to dependency relations, which it enhanced the knowledge of the relationship between syntactic and prosody and also provided a new approach to consider dependency relations and its application.

From the figures above, it was observed that the variation of duration and stress distribution in the governor or the dependent had a similar percentage distribution in most of the dependency relations. From the perspective of duration, we analyzed the distribution of stress. The longer duration the governor or the dependent gained, the more possible it was to obtain stresses. The conclusion was explained by Duanmu's works [17, 18, 19], which proposed that the modifier was a tendency to gain more information. It made a contribution for the governor or dependent to gain longer duration. The longer duration caused the result that the governor or dependent had more chances to obtain stress. However, there were a few surprising results reflected in the statistics, which might be resulted from the following situations.

The study might suffer from several biases in the following aspects. First of all, the skewness of data might be caused by comparatively small corpus. For instance, the numbers of DEI, DI, OF, SIM and AB relations were too small. Secondly, due to the inevitability of subjectivity, it is impossible to avoid all of mistakes, regardless of double check being processed during the labeling the dependency relation. Last but not the least, the skewed result might stem from the fact that the accuracy and consistency of stress degree annotation were influenced by several factors, such as vowel quality and tones. To sum up, it was built on pure perception of trained and experienced linguistics, but it was unavoidable for the subjectivity.

5. Conclusions

This paper explored the variation of the duration within 24 dependency relations based on a Chinese spoken discourse corpus. The principal discoveries were as follows. First, there was an intrinsic association between duration variation and dependency relation. It showed that governor duration was more likely to be longer than dependent duration at dependency relations of LAD, MT, ATT, ADV, QUN, CNJ, RAD , and dependent duration was more likely to be longer than governor duration at dependency relations of DE, POB and VOB. Secondly, duration variation made contributions to correctly predict stress distribution mostly. The result indicated that the longer duration accounted for, the more possible it tended to obtain stresses. However, some conclusions in relations were unusual, such as ATT and ADV.

The research, as a preliminary experiment of the whole study, revealed and evoked more questions than what had been solved. The forthcoming study is trying to expand the data even to spontaneous speech for get more reliable results, and also to combine the rhetorical structure and information structure to explore the relation between the acoustic features and syntactic or semantic meanings for discourse analysis or understanding.

6. Acknowledgements

This research is supported by National Program on Key Basic Research Project (973 Program) under Grant 2013 CB329301, as well as the Innovation Program of Chinese Academy of Social Sciences "Key Laboratory of Phonetics and Speech Science".

7. References

[1] C. Y. Tseng, "Corpus phonetic investigations of discourse prosody and higher level information," *Language and Linguistics*, vol. 9, no. 4, pp. 659-719, 2008.

[2] C. Y. Tseng, "An F0 analysis of discourse construction and global information in realized narrative prosody," *Language and Linguistics*, vol. 11, no. 2, pp. 183-218, 2010.

[3] C. Y. Tseng, and B. L. Fu, "Duration, intensity and pause predictions in relation to prosody organization," in *INTERSPEECH 2005 - 6th Annual Conference of the International Speech Communication Association. 2005*, pp. 1405-1408.

[4] X. Yang, J. Zhao, Y. Yang, et al. "The roles of pitch and duration in sentence accent of Chinese discourse," *Chinese Journal of Acoustics*, vol. 2, no. 02, pp. 224-240, 2012.

[5] X. Yang, and Y. F. Yang, "Prosodic realization of rhetorical structure in Chinese discourse," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 20, no. 4, pp. 1196-1206, 2012.

[6] L. Zhang, Y. Jia, and A. J. Li, "A preliminary research on rhetorical structural and prosodic features in Chinese reading texts," in *ISCSLP 2014 - 9th International Symposium on IEEE, Chinese Spoken Language Processing*, 2014, pp. 265-269.

[7] L. Zhang, Y. Jia, and A. J. Li, "Analysis of Prosodic and Rhetorical Structural Influence on Pause Duration in Chinese Reading Texts," in *Speech Prosody 2014*.

[8] L. Y. HOU, and Y. JIA, "Phonetic manifestations of pronominal and nominal anaphora in Chinese reading texts," *Journal of Tsinghua University (Science and Technology)*, vol. 6, pp. 022, 2013.

[9] L. Y. HOU, and Y. JIA, "Phonetic manifestation and influential factors of pronominal anaphoric word "TA" in Chinese reading texts," *in 2012 International Conference on IEEE, Speech Database and Assessments (Oriental COCOSDA)*, 2012, pp. 112-117.

[10] L. Y. HOU, Y. JIA, A.J. Li, "Phonetic manifestation and influence of zero anaphora in Chinese reading texts," in *INTERSPEECH 2013*, pp. 1424--1428

[11] Y. Q. Dong, A.J. Li, Y. Jia, "Prosodic features of Chinese reading texts within the framework of Centering Theory," in *NCMMSC 2015*.

[12] Y. Wang, A. J. Li, Y. Jia, "Stress distribution based on dependency parsing of Chinese discourse," in *COCOSDA 17th - 2014 Oriental Chapter of the International Committee for the IEEE Co-ordination and Standardization of Speech Databases and Assessment Techniques*, 2014, pp. 1-5.

[13] L. X. Fei, A.J. Li, "Intonation Stress Distribution of Chinese Spoken Dialogue under Dependency Relations," in *IACL 2*3. August, Korea, 2015.

[14] L. X. Fei, A.J. Li, Y. Jia and Y. Zu, "Syntactic annotation under dependency scheme on Chinese spontaneous speech," In *COCOSDA 2014 - 17th Oriental Chapter of the International Committee for the IEE Co-ordination and Standardization of Speech Databases and Assessment Techniques*, 2014, pp. 1-6.

[15] A.J. Li, et al. "Speech corpus of Chinese discourse and the phonetic research," in I*CSLP 2000*.

[16] A. J. Li,and Y. Q. Zu. "Corpus Design and Annotation for Speech Synthesis and Recognition," *Advances in Chinese Spoken Language Processing*, 2006, 243-268.

[17] S. Duanmu, *A formal study of syllable, tone, stress and domain in Chinese languages*, Beijing: Massachusetts Institute of Technology, 1990.

[18] S. Duanmu, "Syllable Structure and Stress," *The Handbook of Chinese Linguistics*, 2014, pp. 422-442.

[19] S. Duanmu, "Stress and the development of disyllabic words in Chinese," *Diachronica*, vol. 16, no.1, pp. 1-35, 1999.

[20] Y. J. Wang, M. Chu. L. He. "A Preliminary Experimental Study on Chinese Focus and Semantic Stress Distribution," *Chinese Teaching in the World*, 2006, pp.87-98.