The Role of Prepositions in Predicting Chinese Rhythm Structure

Pei Yulai, Qiu Jinping and Li Rong

Xi'an International Studies University

No.1 Wenyuan Nanlu, Guodu Educational Sci-tech Industrial Zone

Xi'an, China

{ peiyulai; yiying_qiu, lirong1210}@163.com

ABSTRACT. *The use of preposition to predict the rhythm structure of Chinese is not deeply studied in Chinese Text-to-speech. The preposition which is a marker to grammar structure and rhythm can tell some information about the segmentation and group of the sentence. This paper does the research on how to improve the accuracy on prediction of rhythm structure in Chinese by utilizing the preposition.*

Keywords: Predication of Rhythm Structure; the Use of Preposition; Chinese TTS

1. **Introduction.** This paper discusses how to use prepositions to improve the degree of accuracy in predicting prosodic segmentation and grouping in Text-to-Speech task. Text-to-Speech is a technique used to translate the text message generated or imported by the computer to intelligible influent speech sounds. It belongs to speech synthesis. Along with the wildly employment of the techniques ,such as Voce Wave form Concatenation and Pitch-Synchronous-Overlap-Add [1,2], the most effective and key way to improve the naturalness and intelligibility of synthetic speech is how to solve the problem in predicting the rhythm structure of the sentence. So, we need the answer to the question：How to more accurately predict the rhythm structure of the sentence? The lexical prosodic segmentation and grouping means foot，prosodic phrase and other rhythm structure. It usually divided into 3 layers: foot, prosodic phrase and intonational phrase [3-7].As showed in Example (1), where "space" means foot. Prosodic phrase separated by "#". Prosodic phrase is composed or directly realized by foot. "/" spaced prosodic phrase, prosodic phrase is composed or directly realized by prosodic phrase.

(1) a. Haolaiweu dianyingde # shengchan heliutong / shiyi yige# quanqiuhua shichang# weizhichengde

好莱坞　　电影的　　生产　　和流通　是以　一个　全球化　　市场 为支撑的

Hollywood movie production and circulation is supported by a globalized market.

b.Zhongguo # gaige kaifang yiqian / yefashengguo # xuduoci dizheng
中国　　改革 开放 以前　　也发生过　　许多次 地震
There are many earthquakes before the reform and opening up in China.

c.Wulun shenmeren # zhiyaota laiguo Xi'an / douhuibei zhezuo chengshi # suoxiyin.

无论　什么人　　只要他 来过 西安　都会被　这座　城市 所吸引

Xi'an City is attractive to anyone who has been there.

As for Text-to-Speech task, the most difficult problem in Predicting prosodic segmentation and grouping is that it can only rely on the input information of part of speech (POS) sequence and syllable number, which latter we refer as "POS-syllable number sequence". Such as this sentence "zhidingle diyu teda hongshui xijide mubiao（制定了 抵御 特大 洪水 袭击的 目标）" (Develop a goal against flood attacks), if it enters into the Text-to-Speech system, it only shows a sequence of 2v lue 2v 2b 2n 2v 1ud 2n[1], but no other information like meaning, sentence structure. So the system with the POS-syllable number sequence runs on weak support. Therefore, in the current situation which used statistics-based rules supplemented by technical conditions, how to more fully to tap the useful information underlying in the "POS-syllable Number Sequence"? And how to input more information of Text-to-Speech system under the existing conditions? This becomes the key to improve the accuracy of predicting prosodic segmentation and grouping. [3, 9-18] There are limited numbers of prepositions in modern Chinese which can make an exhaustive list. More importantly, prepositions appear in a very high frequency. We use "comma" as the mark of sentence, and randomly selected 6000 sentence from the People's Daily corpus, then 2000 sentence is divided into a group, and we got three groups A, B, C. Afterwards, we examine the "preposition" proportion of each group, and do final average. The result is that the sentences which contains preposition accounted for 31.4% of all the sentences. Further more, we consider the sentence with an "comma" are always shorter, such as example (2) , and these shorter sentence do not bother the prediction of the rhythm structure because of their being on unit of rhythm. If we only count the occurrence of the prepositions in the length of more than seven syllable "utterance unit", the ratio must be higher.

(2)　a. shishishang /lb
　　　事实上
　　　In fact
　　b. Haolaiwu /ns　dianying/n
　　　好莱坞　　　电影

---

[1] This sequence contains both the information of the parts of speech and the number of syllables. Here, 2n indicates disyllabic nouns, 1ue indicates the Auxiliary word "Le", 1ud indicates Particle "de", 2b shows the two-syllable distinguishing word, 2v shows the disyllabic verb. Part-of-speech tags are from Peking University [8].

The Hollywood movie

  c. qixiang/n sumu/a
   气象   肃穆
   Status   solemn

 If the system find and fully utilize the role of preposition of predicting prosodic segmentation and grouping, the accuracy of prediction will be significantly improved. This paper will discuss that in detail in the following chapters.

2. Preposition Tips the Grammatical Structures. The research on predicting prosodic segmentation and grouping based on grammar structure has been studied fully [3, 14, 19,20]. If we can determine the syntax structure of a sentence, the task of predicting prosodic segmentation will become relatively simple. There are 85 prepositions in modern Chinese, including "an (按) (according to), ba (把), bi (比) (than), chao (朝) (towards), chen (趁) (take advantage of), cong (从) (from), dui (对) (to), genju (根据) (according to), gen (跟) (with), guanyu (关于) (on), he (和) (and), ju (就) (according to), ping (凭) (rely on), wang (往) (toward), wei (为) (in order to), xiang (向) (towards), yan (沿) (along with), yi (以) (by), yu (于) (into), youyu (由于) (since), zai (在) (in process of), zhao (照) (according to), zi (自) (since)". Different prepositions have different usages, but the usage of a specific preposition is relatively clear [4],This makes it possible to predict local syntax structure by using preposition depending on the preposition with some usage, for example, We can predict in example(3), inside the [......] is the adverbial-center phrase which using preposition as the adverbial.

 Although we are unable to predict the whole syntactic structure of the sentence, the map of rhythm structure does not usually depend on a complete syntactic structure.[13,21] In fact , the local prediction is very helpful.

 (3) a.yiding/d yao/vu [yong/p ziji/rr jianqiang/a de/ud yizhi/n] lai/vi baozheng/v

    一定 要 用 自己 坚强 的 意志 来 保证

   (We) must use our strong will to ensure.

   b.quanshi/n renkou/n yi/d [you/p jianguo/vn chuqi/t de/ud liubaibashierwan/m ren/n] zengjia/vdao/vq sanqianyibaijiushiqiwan/m

    全市 人口 已 由 建国 初期 的 六百八十二万 人 增加 到 三千一百九十七万

   The city's population has been increased from six million eight hundred and twenty thousand people in the early age of PRC to thirty-one million seven hundred and ninety thousand people.

   c.[you /p zhuguan/vn bumen/n] fuze/v zuohao/v jieshi/vn gongzuo/vn
    由   主管   部门 负责 做好 解释 工作

The department which is in charge is responsible for doing the interpretation.

d.haoran/nr [ba/p ziji/rr de/ud quanbu/m xinxue/n] sa/v [zai/p [wei/p jingdong/ns dadi/n] peiyu/v wenxue/n xinren/n shang/f]

浩然　　　把　　　自己　　的　　　全部　　　心血　　洒　　在　　　为
京东　　　大地　　培育　　文学　　新人　　上

Haoran put his entire heart and effort in educating new literary people of the Jingdong land.

e.[dui/p xini/ns aoyunhui/jn choubei/vn gongzuo/vn jinxing/vx le/ul] quanfangwei/n 、/wu quanjingshi/b de/ud baodao/vn

对　　　悉尼　　　奥运会　　　筹备　　　　工作　　　进行　　　了
全方位、　　　　全景式　　　的　　　报道

We conduct en all-round panoramic report on the preparations for the Sydney Olympic Games.

Determination of the syntactic structure by using prepositions needs to avoid the interference of "muti-possibility structures Sequence". "Muti-possibility structures Sequence" is the sequence with more than one possible structure. According to our research, there are three types of "muti-possibility structures sequence" associated with prepositions. The first one is "p+n+v+ud+n". The possible structure of this kind of sequence can be "[p+n]+v+ud+n" or "［p+n+v+ud+n］",which are respectively corresponded to (4a) and (4b). It is the different structure that makes their rhythm structure completely different. There is no pause in the Preposition and subsequent nouns in (4a) , but there is a pause in (4b). It is the first step to solve the problem to figure out all "muti-possibility structures Sequence". One is above, the others following.

(4) a.[cong/p Beijing/n] chufa/v de/ud lvxingtuan/n yijing/d dida/v shanghaijichang/n

从　　　北京　　　出发　　的　　　旅行团　　　已经　　　抵达
上海机场

The Tour group starting from Beijing has arrived at Shanghai airport.

b.[wei/p jinnian/t jijiang/d juxing/v de/ud aoyunhui/n] zhuru/v le/ul xin/a de/ud huoli/n

为　　　今年　　　即将　　举行　　的　　　奥运会　　　注入　　了
新　　的　　　活力

(It) injected new life for the upcoming Olympic Games.

The second muti-possibility structures Sequence is "p+n+ud+n". The possible structure of this kind of sequence can be "［p+n］+ud+n" or "[p+n+ud+n]",which are respectively corresponded to (5a) and (5b).

(5) a.quanshu/n fuyou/v jishi/m fu/qe [youguan/p ershiqi/m jie/q

aoyunhui/n] de/ud jingmei/a chatu/n

　　　　全书　　　　附有　　　几十　　　幅　　　有关　　　二十七　　　届
奥运会　　的　　　精美　　　插图

　　　　　The book contains dozens of beautiful pictures on the twenty-seventh Olympics.

　　　b.women/r [genju/p jishi/m ming/q mujizhe/n de/ud jieshao/n] dechu/v
le/ul jielun/n

　　　　我们　　　根据　　　几十　　　名　　　目击者　　　的　　　介绍　　　得出
了　　结论

　　　　　We came to the conclusion based on the testimony of more than a dozen witnesses.

The third muti-possibility structures Sequence is that the prepositional phrases appear in successive sequences. Though it is not too much, it makes trouble too. Its structure may be (6a), may also be (6b).

　　(6) a.ta/r [he/p ji/m ming/q duiyuan/n] [you/p Beijing/n] chufa/v
　　　　他　　和　　几　　名　　队员　　由　　北京　　出发
　　　　He and several members departed from Beijing.

　　　b.[zai/p diyu/v gezhong/r [dui/p ben/r diqu/n anquan/an] de/ud
weixie/vn fangmian/n] jiaqiang/v xiangmu/b zhichi/vn

　　　　在　　　抵御　　　各种　　　对　　　本　　　地区　　　安全　　　的
威胁　　　方面　　　加强　　　相互　　　支持

　　　　(we) strengthen mutual support to withstand a variety of regional security threats.

According to our research, more detailed classification and marking of all the prepositions greatly reduce the number of "muti-possible structural sequence". Predicting the local grammatical relations based on prepositions will obviously improve rhythm structure prediction accuracy.

3. **The Unique Performance of the Preposition in Rhythm Structure.** The unique performance of preposition in the rhythm structure means: except the preposition "zai (in process of), xiang　(towards), yu (into), zi (since)", such that located after verb or adjective as complement, other prepositions and its former components most likely appears respectively in two different prosodic group. In the other word , the prosodic boundary usually happens between the prepositions and its former components, such as "renmin # dui" , "weiyuanhui / cong",　 "zongshuji # zaiqinghua" and "jianghuazhong # xiangqueguo" in example (7). The every sentence in (7) are all long enough to segment two more groups, so the most likely prosodic boundary is between the prepositions and the former.

　　(7) a. pushide zhongguo renmin # dui zhexie guojia / shizhong huaiyou# youhaode ganqing

朴实的　中国　　人民　　对　这些　国家　　始终　　　怀有　　　友好的
感情

Chinese people have always cherished friendly feelings for these countries.

b. Pingshen weiyuanhui / cong shuwanfen zuopinzhong # xuanchule zuigaojiang.

评奖　　委员会　　从　数万份　作品中　　　选出了　最高奖

The committee reviews the best one from tens of thousands of entries.

c. HuJintao zongshuji # zaiqinghua daxue jianghuazhong # xiangqueguo keyan gongzuoze / zhimingle # zhongguo daxuede # fazhan fangxiang

胡锦涛　总书记　　在清华　大学　　讲话中　　　向全国　科研
工作者　指明了　　中国　大学的　发展　　方向

Mr. President Hu Jintao shows the future of Chinese University in the Tsinghua University speech.

We need to pay attention that " the most likely prosodic boundary " does not mean the prosodic boundary must appears here, and it just means more occurrence of prosodic boundary over the other words between the two components in this sequence. If the sentence is shorter, it does not need to be divided into two basic unit of rhythm. And the prosodic boundary does not occur naturally between the preposition and its former components, for example, in (8), the entire sentence only constitute a rhythm structure, there is no prosodic boundaries between "nansheng (boy)" and "xiangbeizou (go north)".

(8) a. #nansheng   xiangbeizou#
　　男生　　　向北走
　#boys　　　go north#
　The boys head towards north.
b. #niaowang tianshangfei#
　　鸟往　　天上飞
　The bird flies to the sky.

This unique performance of preposition in the rhythm structure is very significant to predict rhythm structure of a sentence. According to prosodic boundary based on preposition, we can divide the sentence rhythmic structure into the two sections, which reduces the length of the prediction target. The further prediction only needs to handle these two sections around the prosodic boundary, which in fact, greatly reduces the difficulty in predicting the rhythm structure. As the following sentence (9) is another example of predicting the rhythm structure based on prepositions:

(9)a.guowai/s didui/a shili/n tongguo/p gezhong/r tujing/n jinxing/vx sixiang/n wenhua/n shengtou/vn

国外　　　敌对　势力　通过　各种　　途径　进行　　思想
文化　渗透

Rhythm Structures：

guowai didui shili/ tongguo gezhong tujing/ jinxing# sixiang wenhua shengtou

The foreign hostile forces infiltrate their ideology and culture through a variety of channels.

b.anzhao/p jiangzeming/nr tongzhi/n ti/v chu/vq de/ud "/wyz san/m ge/qe daibiao/v "/wyy de/ud yaoqiu/n gaishan/v mingshengg/n

按照　　江泽民　　　同志　提　出　　的　　　　　三
个　　代表　　　　　的　　要求　改善　　民生

Rhythm Structures：

anzhao# jiangzeming tongzhi tichude/ sange dabioade yaoqiu / gaishan mingsheng

According to the requirement of "Sangedaibiao(The Three Representations)"proposed by Comrade Jiang Zemin, we improve people's wellbeings.

c.ba/p gongzuo/vn luoshi/v zai/p jiceng/n

把　　工作　　落实　在　基层

Rhythm Structures：

bagongzuo# luoshizai jiceng

To work in the basic level

d.zhubu/d shixian/v you/p zhongyang/n banxue/vi

逐步　　实现　　由　　中央　　　办学

Rhythm Structures：

zhubu shixian/ you zhongyang banxue

Gradually realize that the central government runs the schools.

4. **Conclusions.** Prepositions play an important role in improving the accuracy of predicting the prosodic segmentation structure. We should exploit the prepositions role properly and must include it into the entire sentence rhythm structure prediction systems, such as considering how to utilize prepositions in complex environments, how prepositions collocate other function words and so on. As in example (10a), we need to consider the characteristic of the adverb "zhengzai (in the process of) " can be attached to other rhythm component. If only considering the most likely prosodic boundary which preposition and its former components appear, such as "dui (on)" and the former "zhengzai (progressive)", we can not get the correct rhythm structure as in (10b) but the bad rhythm structure of the sentence (10c).

(10)a.women yao zhengshi shuzi jishu zhengzai dui dianying chansheng de gemingxing tiaozhan.

我们　要　正视　数字 技术　正在　对　电影　　产生　　的　革命性
挑战

b.women yaozhengshi | shuzi jishu | zhengzaidui dianying | chanshengde gemingxing tiaozhan.

我们　　要正视　数字 技术　正在对　　电影　　　产生的
革命性　　挑战

c.women yaozhengshi | shuzi jishu zhengzai | duidianying chanshengde | gemingxing tiaozhan.

　　　　我们　　要正视　　数字 技术 正在　　　对电影　　产生的

革命性　　　挑战

We must face the revolutionary challenge which the influence of the digital technology on the film.

REFERENCES

[1] Cai Lianhong, Speech Synthesis based on wave form compilation and application on Chinese TTS, Mini-Micro Systems, vol.15, no.10, pp.11-16, 1994.

[2] Chu Min, Shinan Lv, A Chinese text-to-speech system with high intelligibility and high naturalness, Acta Acustica, Vol. 21, No. 4 Suppl, pp 639-647, 1996.

[3] Wang Hongjun, Chinese non-linear phonology (revised) , Beijing: Peking University Press, 2008.

[4] Shih, Chi-Lin, The prosodic domain of tone sandhi in Chinese. Ph.D thesis of University of California, 1986.

[5] Richard Hogg, C.B. McCully，Metrical phonology: a course book, Cambridge: Cambridge University Press, 1987.

[6] Feng Shengli, On natural foot, Studies of the Chinese Language, Vol. 32, No. 1, 40-47, 1998.

[7] Duanmu, San, The phonology of standard Chinese. Oxford: Oxford University Press,2007.

[8] Yu Shiwen, Xuefeng Zhu, Hui Wang, Huarui Zhang, Yunyun Zhang, Dexi Zhu, Jianming Lu, Rui Guo, The Grammatical Knowledge-base of Contemporary Chinese—A Complete Specification, Beijing: Tsinghua University Press, 2002.

[9] Ying Hong, Lianhong Cai, Research on the segmentation of the prosodic phrase based on driven by the structural auxiliary word, Journal of Chinese Information Processing, Vol. 13, No. 6, pp. 41-64, 1999.

[10] Cao Jianfen, Prediction of prosodic organization based on grammatical Information, Journal of Chinese Information Processing, Vol. 17. No. 3, pp. 41-46, 2003.

[11] Cao Jianfen，The use of grammar structure in the prosodic prediction and the TTS, The Research on The Phonetics, Beijing: The Commercial Press，2007.

[12] Tao Jianhua， Statistic and rule based prosody model of Chinese speech synthesis, Acta Acustica, Vol. 20, No. z2, pp. 359-361, 2003.

[13] Chu Min, Wenjia Wang, Mingzhen Bao, The local constrains of grammar and length on the rhythm structure, Yuyanxue Luncong (33) , Beijing: The Commercial Press, 2004.

[14] Dong honghui, Jianhua tao, Bo xu, Chinese prosodic phrasing with a constraint-based approach, Journal of Chinese Information Processing, Vol. 21, No. 1, pp. 54-59, 2007.

[15] Zhu Weibin, Shinan Lv, Semantic-based speech synthesis——survey and perspective on the speech synthesis technology, Transactions of Beijing Institute of Technology, Vol. 27, No. 5, pp. 408-412, 2007.

[16] Pei Yulai, Jinping Qiu, Hongjun Wang, Shinan Lv, Chinese sentence prosodic structure prediction based on the sequence of the parts of speech, Journal of Tsinghua University (Science and Technology), Vol. 49, No. s1, pp. 1339-1343, 2009.

[17] Pei Yulai, The relation between POS-NOS and prosodic structure and prosodic structure's prediction, Linguistic Sciences, Vol. 10, No. 6, pp. 574-585, 2011.

[18] Zhu Weibin, Linguistic models in speech synthesis: the state of the art, Contemporary Linguistics , Vol. 11, No. 2, pp. 159-166, 2009.

[19] Wang Hongjun, The association between the rhythm structure and the grammar, semantic and pragmatic, Yuyanxue Luncong (26) , Beijing: The Commercial Press, 2002.

[20] Wang Hongjun, The re-study on the association between the rhythm structure and the grammar, Acta Acustica, Vol. 22, No. z1, pp. 428-433, 2005.

[21] Chen, Y. Matthew, What must phonology know about syntax?, In Phonology-syntax Connection, by Sharon Inkelas and Draga Zec, Chicago and London: The University of Chicago Press, 19-46,1990.