



词类-音节数序列 与语句韵律结构生成

裴雨来 / 著



科学出版社

词类-音节数序列 与语句韵律结构生成

裴雨来 著



科学出版社

北京

内 容 简 介

本书主要讨论汉语文语转换中语句韵律结构生成相关的问题,具体研究涉及两方面内容:一方面是汉语语句韵律结构生成与哪些因素有关、语句韵律结构生成的具体过程和机制如何;另一方面,基于汉语语句韵律生成的本体研究,讨论词类、音节数等信息对语句韵律结构生成有什么贡献、产生贡献的机制是什么。

本书语言比较通俗,力避理论化倾向,对所涉及的主要概念都进行了举例说明,对读者基础知识要求不高。从内容上说,本书适合对汉语韵律或汉语节奏感兴趣的读者,也适合对汉语文语转换感兴趣的读者。

图书在版编目(CIP)数据

词类-音节数序列与语句韵律结构生成 / 裴雨来著. —北京:科学出版社, 2016.3

ISBN 978-7-03-047609-8

I. ①词… II. ①裴… III. ①词类-音节-关系-语句-韵律(语言)-结构-研究 IV. ①H0

中国版本图书馆 CIP 数据核字(2016)第 047365 号

责任编辑:常春娥 / 责任校对:张怡君

责任印制:张 伟 / 封面设计:铭轩堂

科 学 出 版 社 出 版

北京东黄城根北街 16 号

邮政编码:100717

<http://www.sciencep.com>

北京京华虎彩印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2016 年 3 月第 一 版 开本:720×1000 1/16

2016 年 3 月第一次印刷 印张:11 1/2

字数:230 000

定价:75.00 元

(如有印装质量问题,我社负责调换)

前言

汉语语句的韵律结构如何生成?关于这一问题,自二十世纪八十年代以来,多有研究,虽不同学者提出的具体模型有细微差别,但韵律结构生成依赖语法结构、意义等已是共识。没有语法结构信息,没有语义信息,韵律结构的生成就无法实现。这意味着,单纯依靠词类、音节数信息是无法正确生成语句韵律结构的。但语音合成等实际应用目前能够利用的基础信息仍只是词类、音节数等。词类、音节数与语句韵律结构生成间是否存在联系,存在什么联系;单纯依靠词类、音节数确定语句韵律结构的可能性是否存在;如果不能单纯依靠词类、音节数信息确定语句韵律结构,最少还需要增加哪些信息……语言应用实践向本体研究提出了一系列需求,本体研究有必要给予回应。于是,笔者希望通过深入研究来回答词类、音节数信息对语句韵律结构生成的贡献究竟有多少、产生贡献的根本机制是什么等关键问题。事实上这是本书研究的根本动力。

笔者首先仔细分析了汉语韵律结构相关的种种现象,对韵律结构的生成过程做了详细分析。首先,韵律结构生成的初始元素包括两个方面,一方面是语句的语法、语义信息,另一方面是语句中各成分的音节数。同时,笔者发现对于语句韵律结构的确定来说,语法结构和语义信息虽然是必要的,但其详尽程度却有很大的弹性。语句韵律结构确定允许语法、语义信息存在一定程度的误差或仅是局部范围内的信息。另外,汉语中事实上存在一些韵律结构与词类串对应的情况。

全书共分七章,逐步说明笔者对词类、音节数信息与语句韵律结构之间关系的认识。第1章主要阐述词类、音节数序列与汉语语句韵律结构生成间关系研究涉及的一些基本问题。本章原则上明确了语法信息、语义信息、词类信息、音节数信息等韵律结构生成中所扮演的角色,并描述了体现词类、音节数、语法、语义等方面信息的韵律结构生成模型框架。

第2章主要和词类-音节数序列与语法结构的关系有关。韵律结构的生成需要一定程度的语法结构信息,本章分别讨论词类-音节数序列中包含哪些语法结构信息进而辅助作用于韵律结构的生成和哪些现实条件降低了韵律结构生成对语法结构信息的依赖,从而带来词类、音节信息的贡献度提升。

第3章主要探讨音节数和音步结构生成之间的关系。根据音步生成的节律常规及笔者对音步结构与语法结构关系的研究,语句的音步生成,要解决的事实上是单音节如何归步的问题。单音节可以归入前音节、归入后音节,也可以独立组成音步。本章主要说明具体文本中,单音节成分以哪一种方式实现音步的规律。

第4章具体讨论词语两兼的词类序列。现代汉语中有一些词类序列,既可以是词,也可以是短语。这些序列是词还是短语会极大地影响其所在语句的韵律结构生成结果。这些词、语两兼的词类序列究竟会对韵律结构生成造成多大影响是笔者必须回答的问题。本章通过分析含动复合词这类常见的词语两兼结构,讨论了确定语句韵律结构除了词类、音节数信息外,最少还需要哪些信息的问题。

第5章主要是在前几章研究成果的基础上,借助具体模型进一步明确词类、音节数信息在音步结构生成中的贡献。总体来说,确定语句的音步结构并不需要掌握全句的语法结构,而只需要在了解很小的某个局部范围的语法关系后,就可以进行。另外,真实语料中单音节词比例并不高于实际情况,事实上大大提高了词类-音节数序列在音步结构生成中的贡献。

第6章主要说明词类、音节数信息对停延段结构生成的贡献。停延段结构的生成,对语法等信息的依赖度远高于音步结构生成,所以根据词类、音节数等简单信息生成语句停延段结构的正确性降低,基于停延段生成模型,本章论述了词类、音节数之外,还需要哪些信息。同时,本章也说明了一些可以降低停延段结构生成对语法结构信息依赖的手段。

第7章是本书最后一章,主要讨论本书所讨论的汉语语句韵律结构生成模型对汉语文语转换来说的应用可能。

本书语言比较通俗,力避理论化倾向,对所涉及的主要概念都进行了例举说明,对读者基础知识要求不高。从内容上说,本书适合对汉语韵律或汉语节奏感兴趣的读者,也适合对汉语文语转换感兴趣的读者。

本书的写作出版得到科学出版社阎莉女士、常春娥编辑及其他工作人员的辛勤努力,没有他们的支持和体谅,本书是不可能与读者见面的。谨向他们表示衷心的感谢!

由于本人水平和时间限制,书中难免存在疏漏和不足之处。欢迎各位读者批评指正。

裴雨来

2015年12月

目 录

前言

第 1 章 语句韵律结构的生成框架	1
1.1 韵律结构生成研究述评	1
1.1.1 Shih 的两分句法树模型	3
1.1.2 董秀芳的 X 杠树形模型	4
1.1.3 初敏等的局部语法约束和长度约束	6
1.1.4 王洪君的句块模型	10
1.2 语法结构在韵律结构生成中的地位	11
1.3 体现词类、音节数及语法结构信息的韵律结构生成分析	14
1.3.1 对韵律结构生成过程的基本思考	14
1.3.2 韵律规律对韵律结构的要求及韵律特征核查	16
第 2 章 词类-音节数序列与语法结构	33
2.1 语法结构信息的提取和利用	33
2.1.1 在一定范围内正确预测语法关系	33
2.1.2 根据韵律特征分化词类	38
2.1.3 韵律结构预测对语法结构的容错性	41
2.2 非语法结构信息的辅助作用	45
2.2.1 必须同处一个韵律单元和最不可能同处一个韵律单元的词类序列	45
2.2.2 虚词在停延段生成中的去结构依赖作用	48
2.3 结语	51
第 3 章 单音节成分对音步结构生成的贡献	53
3.1 引言	53
3.2 单音节成分作用范围与音步的关系	54
3.2.1 动词性成分及其作用范围与音步的关系	57
3.2.2 修饰性成分、介词等的作用范围与音步结构的关系	59
3.3 单音节成分与相邻成分的可能性等级判断	61

3.3.1 单音节成分的表现与可能性等级的对应	61
3.3.2 词类-音节数与可能性等级	63
3.4 单音节对相关成分的选择限制情况与可能性等级	67
3.5 结语	68
第4章 词语两兼的词类-音节数序列	69
4.1 词语两兼的序列——含动复合词	69
4.1.1 含动复合词的类型	71
4.1.2 对含动复合词的初步观察	74
4.1.3 含动复合词的各级组成成分	78
4.1.4 含动复合词中动词性成分的作用范围与音步结构间关系	81
4.2 含动复合词的自动识别	84
4.2.1 有助于识别含动复合词的词汇信息	84
4.2.2 不太可能构成含动复合词的词类-音节数序列	85
4.2.3 需要语义等信息才能识别的含动复合词类型	88
4.3 结语	93
第5章 词类-音节数信息对音步结构生成的贡献	94
5.1 音步结构生成过程	94
5.2 音步结构生成对语法关系的依赖	98
5.3 语句音步结构生成模型	103
5.3.1 可能性蕴含等级的具体实现	104
5.3.2 音步结构预分第一步	107
5.3.3 音步结构预分第二步	111
5.3.4 音步结构确定阶段	115
5.4 结语	118
第6章 词类-音节数信息对停延段结构生成的贡献	120
6.1 停延段结构生成模型	120
6.1.1 韵律特征标注规则和可能性等级标注规则	121
6.1.2 停延段结构预分操作	123
6.1.3 停延段结构确定操作	126
6.2 预测停延段结构的有效信息	129
6.2.1 词类-音节数序列能够提供的有效信息	129

6.2.2 确定停延段结构不需要完全识别语法信息的内在原因	132
6.2.3 影响停延段结构预测的语法结构	135
6.3 如何利用虚词帮助预测停延段结构	143
6.3.1 助词“的”在确定停延段结构中的作用	143
6.3.2 利用介词预测停延段结构	145
6.4 结语	148
第3章 文语转换中的词类-音节数信息运用	149
7.1 语句韵律结构生成模型及词类-音节数信息的贡献	149
7.2 韵律结构生成模型与文语转换应用	151
参考文献	155
附录	163
附录1 音步结构生成的部分验证语料	163
附录2 北京大学词类标记集	167
附录3 含动复合词部分词表	172

语句韵律结构的生成框架

1.1 韵律结构生成研究述评

汉语语句韵律结构生成的研究已有近几十年的历史，前人的研究详细讨论了汉语语句韵律结构的生成规律，笔者认为，语句韵律结构的生成与语法结构、意义等有不可分割的关系，从理论上来说，单纯依靠词类、音节数这些信息是无法正确生成语句的韵律结构的。但语音合成等实际应用能够利用的语句信息仍只限于词类、音节数等信息。语言研究有必要回应应用研究的需求，这就必须回答词类、音节数信息对语句韵律结构生成的贡献究竟有多少、产生贡献的根本机制是什么等关键问题，这是本书研究的根本动力。

某一语句的词类、音节数信息可以整合在一个序列中。比如，句子“我爱北京天安门”的词类信息为“r(人称代词)+v(心理动词)+n(专有名词)+n(地点名词)”，音节数信息为“1+1+2+3”，整合为一个序列即为“1r+1v+2n+3n”。本书中，笔者将这种整合词类、音节数信息的序列称为“词类-音节数序列”。词类-音节数序列的词类信息，可以根据通常的词类系统，也可以采用更为细致的中文信息处理词类标注系统，因二者并无本质差别，所以本书在使用时并不做具体说明。详情请见例(1)，其中(1a)、(1b)句是经过分词及词性标注的语句，(1a')、(1b')是其对应的词类-音节数序列¹，其中“2mq”代表2音节的数量词，“2a”代表2音节的形容词，其余的依此类推²。

(1)a. 一个/mq 巨大/a 的/ud 问题/n 我们/rr 不/df 能/vu 忽略/v

a', 2mq 2a 1ud 2n 2rr 1df 1vu 2v

1 音节数的信息对于计算机来说，非常容易获得，所以并不给预测韵律结构带来困难。而且，从学界对韵律句法、韵律词法的研究来看，词的音节数对于预测韵律结构是至关重要的。

2 字母所代表的词类具体请参看“附录2”的北京大学词类标记集。

b. 好莱坞/ns 电影/n 的/ud 生产/vn 和/c 流通/vn 是/vl 以/p
一个/mq 全球化/vn 市场/n 为/vl 支撑/v 的/ud

b'. 3ns 2n 1ud 2vn 1c 2vn 1vl 1p
2mq 3vn 2n 1vl 2v 1ud

本书讨论的语句韵律结构主要限于音步、停延段等韵律单位的结构¹。笔者采纳王洪君(2008)对韵律单位术语的区分,将语句表层语音形式所组织得到的表层韵律结构层级称为“音步、小停延段、大停延段”三级,而没有采用“韵律词、韵律短语/小韵律短语、语调短语/大韵律短语”等概念²。下(2)是本章所说语句韵律结构示例,其中“空格”隔开的都是音步,比如,“中国、在这里、和奠定、电影的、是以、为支撑的”等;“#”隔开的是小停延段,小停延段由音步组成或直接实现,比如,“中国共产党的、在这里、开创和奠定、是以一个、为支撑的”等;“/”隔开的是大停延段,大停延段由小停延段组成或实现,比如,“中国共产党的领导机关、在这里开创和奠定、生产和流通、是以一个全球化市场为支撑的”等。

(2)a. 中国 共产党的 # 领导 机关 / 在这里 # 开创 和奠定 / 人民 共和国的
基业

b. 好莱坞 电影的# 生产 和流通/ 是以 一个# 全球化 市场# 为支撑的

如果词类、音节数等简单语句韵律结构间存在某种函数对应关系,即存在词类-音节数序列与韵律结构间的关系函数“ $F(x,y)$ ”,根据输入的词类-音节数序列 LINE,就可以得到与之对应的韵律结构 LINE,那么词类、音节数信息在语句韵律结构生成中的作用无疑是充分的,但实际情况并不如此,在词类、音节数信息与语句韵律结构间存在语法结构层面³。大量事实证明,相同的词类-音节数序列完全可能对应不同的韵律结构,相同的语句韵律结构也完全可能对应不同的词类-音节数序列⁴。

比如,在以下两组例句中,语法结构不同,词类、音节数等相同的序列,却体现出完全不同的韵律结构:(3)、(4)的词类-音节数序列——(3a)、(4a)——可能是(3b)、(4b)的韵律结构,也可能是(3c)、(4c)的韵律结构。

1 某一个语句具体实现的韵律结构的形式存在较大的自由度(参见初敏等2004),尤其是实现为哪种停延段结构,可能性更多。这套系统的较高目标是可以预测出各种可能的停延段边界,初步目标是正确预测多种停延段可能性中的一种。

2 不同研究领域的不同研究者,对语句韵律层级的命名有所不同(参见吴洁敏1998,贺琳等2001),本书所使用的音步、停延段对应于语音合成工程学界常用的韵律词、韵律短语。

3 如果是这样,基于词类-音节数序列预测语句的韵律结构也会比较容易,但事实上并不如此。在本书对语句韵律结构生成过程的分析说明中,语法结构作为中间层面,起着重要的作用。

4 不考虑表达特殊语义的情况。

(3) a. $1a+2n+2v+2a+1ud+2n$

b. $1a+2n \# 2v \ 2a+1ud \ # \ 2n$

老仆人 整理干净的 房间

c. $1a+2n \# 2v / 2a+1ud \ # \ 2n$

小城镇 占有 重要的地位

(4) a. $2m+2vn+2n+2v+2n+2v+1ul$

b. $2m \ 2vn \ 2n / 2v \ 2n \ # \ 2v+1ul$

许多 批判 材料 / 当作 垃圾 # 处理了

c. $2m \ # \ 2vn \ 2n \ 2v \ 2n / 2v+1ul$

一些 # 办公 用品 采购 项目 / 失败了

看来,词类-音节数序列与韵律结构间并不存在某种直接的对应关系。句子的语法结构不同对韵律结构的生成起着重要的影响。为了更清晰地说明这一问题,笔者首先从韵律结构的生成模型分析开始。

韵律结构的生成,在前人的研究中也被称为韵律结构的动态组织。目前,讨论现代汉语韵律结构动态组织/生成的模型主要有四种,分别是 Shih (1986) 的结构主义的两分句法树(IC)模型;王洪君的基于功能语法的句块模型(王洪君 2002, 2005a, 2008);董秀芳(2003)的基于生成语法的X 杠树形模型;初敏等(2004)的基于线性局部语法结构模型。笔者先依次讨论 Shih 的方案、董秀芳的方案、初敏的方案,最后讨论王洪君的方案。

1.1.1 Shih 的两分句法树模型

Shih 的方案是在音步生成分析中最有影响的方案。该方案有如下 3 条顺序使用的规则(Shih 1986, 另请参看王洪君 2008)。

- I IC 音步(直接成分音步): 实现为结构最内层的两个直接成分的两个单音节组成音步;
- II DM 音步(Duplex meter, 双拍步): 从左到右连接未成对音节形成双拍音步, 但不包括句法分支方向相反的两成分;
- III f' 音步(超音步): 根据句法分支的方向把剩下的单音节和邻近的双拍音步连接形成超音步。

比如,在下面(5a)中,“小红纸伞”最内层的两个直接成分分别是“纸”“伞”,他们都是单音节,组成音步“纸伞”;余下的成分都不能构成直接成分音步,“小”“红”

构成 DM 音步，结构最后的音步结构为“小红 | 纸伞”。

(5b)中的“取”“车”可以构成 IC 音步，“车”“费”也可以构成 IC 音步，余下的成分语法分支方向相反，不能构成 DM 步，两个单音节成分“人”“交”分别按语法分支方向组成超音步——“人”与“取车”构成超音步，“交”与“车费”构成超音步，最后得到语句的音步结构为“取车人 | 交车费”。

(5)a.



b.



Shih 的操作方案是以语句的树形图为基础来制定规则的，在其方案中，更关注相关项是否是直接成分、语法分枝方向是否一致。但现代汉语中，相同的语法分枝，完全可能是不同的语法关系。所以 Shih 的方案似乎暗示只要语法分枝的方向相同，不论二者间是什么语法成分，都可以构成音步，比如(5a)的“小”与“红”并不具有任何语法关系；(5b)中的“取车”与“人”是便正关系，“交”与“车费”间是述宾关系。但事实上语法关系与音步组织关系密切，忽略语法关系，有时会得到错误的音步结构。比如在(6)中，按 Shih 的方案，(6a)的音步结构为“买小 | 红纸 | 雨伞”、(6b)的音步结构为“爸不 | 想去”，但这种音步结构恐怕相对“~买 | 小红 | 纸雨伞”“~爸 | 不想去”而言，被认可的可能性更低。

(6)a.



b.



之所以 Shih 的方案处理(6)会出现错误，就是因为其方案忽略了语法关系。从(6)中可以看出，主谓关系和述宾关系与定中关系和状中关系相比，在音步组织过程中的表现并不相同。

1.1.2 董秀芳的 X 杠树形模型

董秀芳(2003)提出如果组成成分与组合整体语法范畴相同，组成成分中某一成分可跨语法结构界构成韵律结构。董秀芳在其研究中指出，在实际话语中，音步的组织

要受到句法结构的制约，形成结构控制下的音步模式，她在 X 杠阶模型(X-Bar)框架基础上提出了关于音步模式跨层的假设，指出在图 1-1 中“如果短语 X 是短语 XP 的一个直接组成成分，当短语 X 与短语 XP 句法范畴不同时，短语 X 中的成分在韵律上就不能被分开，即不能中插音步界，比如“红苹果”是“买红苹果”的一个成分，“红苹果”是名词性成分，“买红苹果”是动词性范畴，二者句法范畴不同，所以音步结构只能是“(~买)(红苹果)”；当短语 X 与短语 XP 句法范畴相同时，如果短语 X 中的成分在韵律上越出其结构边界，可以造成更好的音步组合，那么短语 X 可以在韵律上被分开，与相邻短语中的成分组成音步，比如“白馒头”是名词性短语，“大白馒头”同样是名词性短语，二者句法范畴相同，所以“白馒头”中的成分可以跨出结构边界，与“大”组成音步，得到音步模式“(大白)(馒头)”。

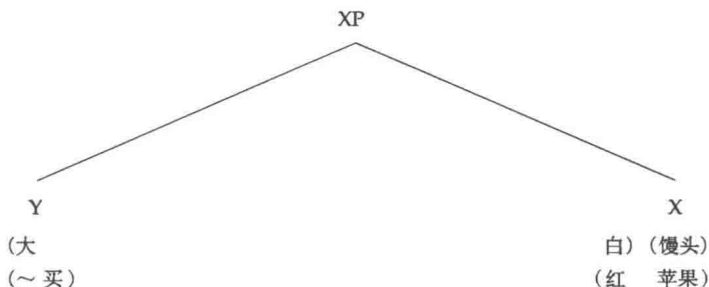
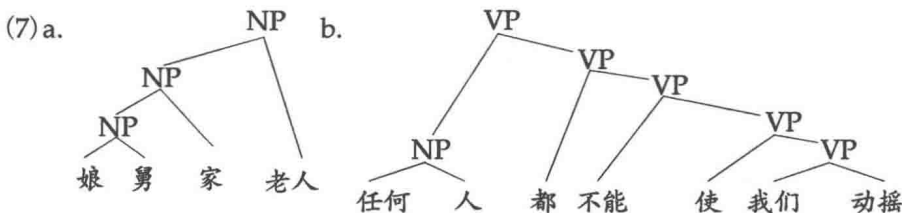


图 1-1 X 杠阶模型基础上的音步构成分析

董秀芳的这一方案的意义也许并不在于提出一种全面分析语句音步动态组织的模型，而更在于其指出了音步组织与语法范畴的关系——音步组织可以在相同语法范畴内根据节律常规进行调整，即在相同语法范畴内，可以打破语法的边界，按右向音步模式组织音步，得到跨层音步——跨层音步打破结构音步，语句最终的音步结构选择“有标记音步”最少的那种。

董秀芳方案的目的并不在于全面分析语句音步组织的组织过程——虽然可以在一定程度上达到这一点，所以当以预测语句音步结构的标准来审视这一方案时，就会发现非常多的细节被忽略了。比如如下几个例子，(7a)中每一个投射的结构都是名词性结构，“娘舅”首先构成结构音步，但接下来的单音节“家”究竟是遵循右向音步模式与“老人”构成音步，还是根据结构与“娘舅”构成音步，在董秀芳方案中并没有说明。与结构相比，董秀芳的方案似乎更偏向于右向音步模式，于是很可能得到“*娘舅 | 家老人”这一错误结果；(7b)中“都不能使我们动摇”按结构音步，有太多单音节残音步，很显然不好，而且其单位语法范畴都是动词性，只能采用右向音步

模式，于是根据董秀芳的方案得到的结构是“*都不 | 能使 | 我们 | 动摇”，这显然也是错误的。



除此以外，王洪君(2008)还指出董秀芳的方案在处理如下问题时也会有困难。如(8)，两句中的“我”与其后成分间语法关系相同，但(8a)只有采用右向音步模式得到的音步结果才是好的，(8b)采用右向音步模式得到的结构却是错误的，董秀芳的方案无法解释这一现象。

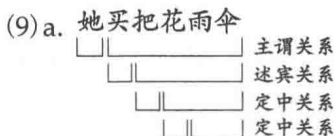
- (8) a. [我[想[买[糖火烧]_N]_{VP}]_{VP}]_S
 (我想买)(糖火烧)
 *(我)(想买)(糖火烧)
- b. [我[很[想[买[糖火烧]_N]_{VP}]_{VP}]_{VP}]_S
 *(我很)(想买)(糖火烧)
 (我)(很想买)(糖火烧)

本书认为，以上例(7)说明单纯根据语法范畴是无法完全解决音步组织的问题的，还需要语法结构的信息，比如，“娘舅”与“家”间存在语法关系，而“家”与“老人”间不存在语法关系；“使我们动摇”是关系非常紧密的一个整体，整体上类似于一种格式。音步组织与语法范畴有关系，但似乎语法范畴只是与音步组织有关的一个方面，如果考虑到语法范畴可以由语法结构推导得出，那么似乎可以认为音步组织与语法范畴间的关系只是音步组织与语法结构关系的一种表现，语法结构才是根本。(8)则说明除了语法范畴之外，语用、语义的信息同样对语句的音步结构造成重要影响。

1.1.3 初敏等的局部语法约束和长度约束

初敏等的局部语法约束和长度约束方案认为，语句的节律组织是“局部语法关系的相对松紧”及“节律组织中的长度约束”共同作用的结果。这一方案最大的特色在于“局部语法关系”的提出。初敏认为，“基本节律单元的组织属于短时决策，主要受制于局部的语法关系，而与这个局部在整个句子中的语法地位关系不大”，初敏等

还进一步指出,“一个成分是否需要同相邻成分组织在同一个节律单元中,或应当同前后哪个成分组织在同一个节律单元中,主要取决于它与左右临近成分之间的语法关系的疏密程度的比较”。于是在线性平面上的成分间语法关系的松紧对比对于节律组织是至关重要的,套合的语法关系无法提供相邻词间的松紧关系的信息,所以必须对其进行处理。初敏等将套合的语法关系描述为线性列上两个成分之间的局部语法关系,比如(9a)的套合语法关系在初敏等的方案中被转化为(9b)的线性成分之间的关系。



在(9a)中“把”的中心语是“花雨伞”,而“把”与“花”、“把”与“雨伞”间本不存在任何语法关系,几个成分层层套合;但在(9b)中,“把”与“花”、“把”与“雨伞”间存在语法关系。“把”与“花”作为线性序列中相邻的两个成分,二者间具有了一定的语法关系。另外,笔者还注意到,在(9a)中,描述成分间关系的术语都是语法关系,比如“主谓关系”“述宾关系”等,但在(9b)中,描述成分间关系的体系已经有了变化,部分引入了纯词类-音节数序列描写的描述方式,比如“买”与“把”间在(9b)中被描述为“动数”关系。

初敏等提出了各种局部语法关系的松紧度等级,比如(9b)的主谓关系是“1”级、动数关系是“4”级、并列定语是“3”级、定中关系是“5”级。另外,初敏等还指出“有些语法关系只有在两个成分都为单音节时才表现得比较紧”。根据松紧度等级及音步及基本节律单元的长度约束进行节律组织。

音步组织的过程如下。

- I. 优先合并: 通过比较相邻语法关系的松紧度决定若干连续单音节词典词中,哪些应当优先合并成一个音步。如果两个词中有一个不是单音节词,它们之间的松紧度受长度影响降低两级。即除了比较紧的关系,不倾向于生成超过三音节的音步。
 - II. 从左到右顺序两两合并: 不满足优先合并条件的其余所有连续单音节词典词从左到右两两合并成双音节音步。即,不倾向保留过多的单音节音步。
- 具体操作如例(10)。

(10) 词典词边界 她 / 买 / 把 / 小 / 花 / 雨 伞

	↓	↓	↓	↓	↓	
局部语法关系	主谓	动数	并列	并列	定中	
		/代	定语	定语		
松紧度	1	4	3	3	3(5-2)	符合优先合并条件的先合并，
	0		1	3	3	其余从左到右合并。
音步	她	买把	小花	雨伞		

定中关系的等级本是“5”，但因为“花雨伞”中“雨伞”是双音节，所以降低为“3”级，于是语句中的最高级别为“动数”关系的“4”级，他们优先合并，得到“买把”。“买”与“把”合并后构成双音节，“把”与“小”间等级同样降低两级，由“3”级变为“1”级，于是“小+花+雨伞”间都是“3”级，按从左到右的顺序两两合并得到“小花|雨伞”。“她”与“买把”间的等级降低两级后，只能为“0”。如果不合并，所以得到语句最终的音步结构为“她|买把|小花|雨伞”。

初敏方案的基本节律单元的组织规则比较简单，“主要是通过比较相邻音步间的松紧度来确定哪些应当优先合并。由于受到长度约束，当两个韵律词优先组合后，它们与其左右两侧的其他韵律词的松紧度会降低两级”。具体如(11)所示，首先是等级最高的“3”优先组合，得到“每一个百姓”和“离开政治”，然后是等级为2的“从来没有”合并，最后是等级为“1”的“事实是”合并。单音节“都”前后等级都是“0”，所以独立构成一个基本节律单元，最后得到的基本节律单元组织结果。

(11) 音步序列 事实|是 |, 每一个|百姓|都|从来|没有|离开|政治

	↓	↓	↓	↓	↓	↓	↓	↓
局部语法关系	主谓	没	数量	没	并列	普通	间接	动宾
		关系	定中	关系	状语	状中	状中	
松紧度	1	0	3	0	1	2	1	3
					0		0	
基本节律单元	事实是	每一个百姓	都	从来没有	离开政治			

初敏等的方案虽然具体规则的设计上仍存在不足，但本书认为规则设计的不完整并不影响这一方案本身的价值，如果方案合理，补足规则并不是很难的任务。应该说初敏等的方案在预测语句节律组织方面做了非常大胆的尝试，提出了很多极有借鉴意义的分析。不过，这一方案本身存在着一些比规则不充分更本质、更关键的问题——方案的框架仍有缺陷。从其预测分析得到的结果来看存在着两点不足。第一个问题是局部语法关系并不局部。局部语法关系中不但包含以词类来标记的内容(比如“动+数”)，还包括利用语法关系来说明的内容。而初敏等的模型中不包含如何在局部范围内获得语法关系的模块，所以，局部语法关系看似可以不考虑全句的语法结构，但事实上并无法摆脱对语句整体语法结构的依赖。比如句(12)，如果只根据相邻词间的

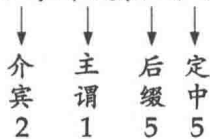
语法关系松紧度, (12a)、(12b)都只能得到错误的节律组织结果。而这样的问题, 只有参考语句更大范围的语法结构, 才能够解决。初敏等的方案中并不包括如何得到类似(12)情况的词与词间语法关系的方法, 也没有说明其局部语法关系的“局部”究竟是多大范围, 确定“局部范围”的条件是什么。这些不足在更大规模语料的分析中, 会体现得更为明显。

(12)a. 计划经济时代形成的部门体制



→ *计划经济时代 | 形成的部门体制

b. 根据学科发展的特点



→ *根据学科 | 发展的特点

初敏等方案的第二个问题是缺乏弹性。在此方案中, 相邻成分间语法关系的松紧决定着某两个成分是否合并成一个更大的韵律单元。于是得到了以上(10)、(11)两例中的情况: 在(10)中单音节代词做主语独立构成一个基本节律单元; 在(11)中单音节副词做状语独立构成一个基本节律单元——很显然, 这两种基本节律单元都不是理想的结果, 至少还需要进一步的处理。另外, 王洪君(2005a)还指出与(10)语法结构完全相同的(13), 如果节律结构与(11)相同, 得到的结果并不好, 而正确的结构应该是“从来|没有||离开||德谟克拉西”。

(13) 每一个百姓都从来没有离开德谟克拉西



*…… 从来 | 没有 || 离开 | 德谟克拉西

笔者认为, 初敏等的方案缺乏弹性的根本原因在于其方案中的“局部语法约束”与“长度约束”的结合出了问题。局部语法约束与长度约束究竟是什么关系, 如何在节律组织的分析框架中体现这种关系, 使二者的作用能够得到准确的反映和体现, 这是至关重要的(本书的韵律特征核查操作可以成功解决这一问题)。