

# 汉语韵律切分的语音学和语言学线索

曹剑芬

中国社会科学院语言研究所  
中国科学院心理研究所  
jianfencao@hotmail.com

## 摘要

本文通过感知试验、声学语音学实验和文本分析,考察韵律边界前后的语音学和语言学信息特征,为韵律短语切分提供更加科学可靠的、可以自动预测的标志。全文重点讨论以下三个方面的问题:第一,关于韵律词和韵律短语的界定;第二,停顿分布与句法结构和词性特点的关系;第三,实际言语中的韵律词组词规律。

## 1. 引言

韵律切分本质上是个话语韵律层次结构问题。当前,有关这个问题的研究已经成为语音学界和言语工程学界普遍关注的焦点。就以计算机文-语转换(TTS)语音合成为例,其基本任务就是把输入的文字材料转换为自然的话音输出。目前,虽然已经能够合成连续的语句,但是,自然度不理想。具体地说,就是节奏感较差,缺少轻重缓急和抑扬顿挫,还不能为公众所接受。因此,提高自然度是目前TTS合成的主要任务,而恰当的韵律切分,正确地把握话语的韵律结构,是加强合成语音节奏感、提高其自然度的关键。

确定话语的韵律层次,是一个相当艰巨的工程,尤其是要想做到自动切分,确实不容易。针对这个问题,国内外的语音学界和言语工程界已经从各个方面进行探索,提出了许多不同的方法。不过,还没有很好地把握语音、语法和语义等很好地结合起来考察,对语音学和语言学信息的运用还不够充分,尚有许多问题亟待解决。

本研究拟通过感知试验、声学语音学实验和文本分析,考察韵律边界前后的语音学和语言学信息特征,研究语言的表层表达跟深层结构以及语法中介之间的复杂关系,寻找各种跟口语的韵律结构对应的语音学和语言学线索,探讨实行自动韵律切分的策略。重点讨论关于韵律词和韵律短语的界定、停顿分布与句法结构和词性的关系以及实际言语中的韵律词组词规则,力求为自然语音处理时的韵律切分提供更加科学可靠的、可以自动预测的语音学和语言学信息。以便在计算机语音处理时,不但知道怎样利用各种超音段特征来实施不同层级的韵律切分,而且知道在什么地方对输入的文本自动进行韵律切分,从而构建恰当的韵律结构模型。

## 2. 测试语料及实验概况

本研究的实验考察对象是从一个语篇语音库里抽取出来的57个自然语句,句子的长度一般为9-20个音节,最长的一个句子包含了41个音节。

实验分析主要包括三个方面:第一,通过文本分析,标注了这部分语料的分词、词性和句法结构;第二,通过听辨测试,标注了不同等级的停顿,确定了各句的韵律切分;第

三,根据文本信息和韵律信息,分析归纳汉语的韵律结构特点。

## 3. 实验结果和讨论

### 3.1 韵律词及韵律短语的界定

在音系学里,韵律层级(Prosodic Hierarchy)从小到大依次为:莫拉、音节、音步、音系词、附着语素词组、音系短语、语调短语和韵律语句。一般情况下,都省略为三个基本层级:韵律词、韵律短语和语调短语。但是,对于这些韵律单元,各家的概念不尽一致,为了满足自然语音处理的需要,这里根据我们的实验研究结果,并参照相关文献,从便于操作的角度,对韵律词和韵律短语加以界定。

#### 3.1.1 韵律词:

- 1) 三音节以下的语法词或词组;
- 2) 其它凡是属于一个音步的结构。音步是指音系学中的基本节奏单位。一般由两个或三个音节构成。根据冯胜利[1],汉语普通话里的基本音步,或者叫标准音步,是由两个正常重音音节构成的,三个音节的音步叫做超音步,一个音节构成的叫做蜕化音步,还有一种叫做残音步,指两个音节中有一个是无调音节(atonic)的音步,即一重一轻的两音节结构,例如轻声词。为简化起见,这里把残音步并入基本音步。因此,韵律词也就是“语法上凝固的、节律上稳定的单音步”[2];
- 3) 跨度为1-3个音节,极大多数为2-3个音节,少数为单音节。单音节的韵律词一般是由单音节词延长而构成的所谓蜕化音步,多数为功能词,例如连词、介词等,也有少数是单音节的动词;
- 4) 具有类似语法词的连调模式和词重音模式。

#### 3.1.2 韵律短语:

- 1) 由一个或几个韵律词组成,即“可能多音步”[2];
- 2) 跨度为 $7\pm 2$ 个音节[3];
- 3) 内部各个韵律词间可能出现韵律上的节奏分界,主要表现为韵律词末尾音节的延长(或伴有较短的无声停顿)和韵律词之间的音高(包括音调和音域)重新设置[4];
- 4) 具有相对稳定的短语调模式,即音阶的declination及其resetting;
- 5) 具有相对稳定的短语重音配置模式,即与句法结构相关的常规重音模式。例如,偏正结构一般是偏重,主谓结构一般是谓重,述宾结构一般为宾重,述补结构一般为补重[5];

#### 3.1.3 关于mini韵律词问题

这里主要涉及四音节结构的韵律层次问题。在汉语里,存在大量的四音节结构,它们多数为四字熟语或成语。在语法上,它们有的是词,有的是词组或短语;在语音上结合比

较紧密，总是一起说。对于这部分结构，在韵律上究竟怎样处理是个有争议的问题。根据王洪君关于韵律词和韵律短语的区分标准[2]，这类结构分为两大部分：一部分纳入韵律词，一部分纳入韵律短语。属于韵律词的有暗亮叠变式的“霹里啪啦”、“稀里糊涂”、“哆哆嗦嗦”、“絮絮叨叨”、“漂漂亮亮”，暗亮复合式的“乱七八糟”，等亮叠变式的“噼啪噼啪”、“啪啦啪啦”、“干干净净”，等亮复合式的“取长补短”、“眉开眼笑”、“远走高飞”、“一衣带水”。纳入韵律短语的有类词短语“格镍合金”、“全程旁听”、“认真学习”和自由短语“又唱又跳”、“天气晴朗”、“心情愉快”等。

根据本文的实验分析结果，这里倾向于把四音节结构处理为 mini 韵律短语。主要理由是，区分韵律词和韵律短语，固然需要考虑语法结构的特点；但是，更主要的，恐怕应当着眼于它们在实际言语中的语音特点。通常，一个韵律词相当于一个音步，它是音系规则作用的最小辖域，具有相对凝固的语音形式，例如相对稳定的连调模式和基于词重音对立的时长结构模式，内部不出现节奏上的分界现象。而韵律短语则不同，它往往由不止一个音步构成，是更高层次上音系规则作用的辖域，具有高于词层面的语调模式和重音配置模式，内部一般都会出现节奏上的分界现象，例如标志性音高运动模式和时长长延伸缩模式。假如从这个角度看，汉语里的四音节结构，不管是“稀里糊涂”、“乱七八糟”、“干干净净”、“取长补短”，还是“全程旁听”、“认真学习”、“天气晴朗”、“心情愉快”，尽管他们在语法结构上存在种种区别，但有一个共同的特点，就是都包含不止一个音步。这就超越了作为韵律词的音系规则作用的辖域。如果说像“稀里糊涂”、“絮絮叨叨”等几个有限的特殊结构还可以用一种特殊的连调模式来概括的话，那么，像“取长补短”、“眉开眼笑”、“一衣带水”、“全程旁听”、“天气晴朗”等，则明明包含两个不同的双音节连调模式，因而在语音上就必然既涉及音步之间的轻重配置问题，又涉及不同的连调模式之间的关系问题，而这些正是韵律短语所具有的层次结构特点。

同时，我们在这里谈论韵律词跟韵律短语的界定，恐怕还不得不考虑语音处理上的可操作性。目前，自然语音处理的实践已经发现，如果把四音节结构作为一个韵律词来整体处理，不但操作上比较困难，而且效果往往不理想，许多词语听起来别扭，主要就是由于这种韵律上的失调引起的。因为这些结构的音高和时长模式非常多变，它不仅涉及不同音节声调之间的彼此影响，而且还涉及不同层面上重音和语调的影响，因而很难建立有限的、能够适应不同语境的、相对稳定的音高和时长变化模式。而如果把四音节结构按韵律短语来处理，就可以根据这类单元所包含的音步数目及其类型（如：基本音步、超音步或蜕化音步），运用有限的韵律词模式和有限的高层调节规则，来实施跟短语调和轻重配置相关的音高和时长调节，从而比较方便地达到改善语音自然度的目的。

### 3.2 停顿插入与句法结构

话语的韵律切分问题，实质上就是言语过程中的组词断句问题，它既涉及听话时知觉上的停顿分布，又涉及说话时语音上的组块分合。

从对自然话语的分析来看，韵律的结构层次跟停顿的等级层次大体一致。因此，寻找韵律切分的规律也就是寻找停顿出现的规律。

一般的感知试验表明，对于通常的听者来说，能够觉察到的语句中的停顿一般只有三级；而进一步的心理知觉实验表明，语句中可以知觉到的停顿起码可以分为四级。而且，

初步的合成试验也发现，对于这第四级知觉上的停顿的适当应用，会对改善合成语音的自然度具有极大的贡献。于是，人们便自然会问，这些停顿的出现都有什么规律？能不能利用我们已经掌握的语法和词汇知识自动预测停顿的位置及其等级？这就涉及人们普遍关注的韵律跟语法的关系问题。迄今为止，国内外都已经投入了不少力量来探索这个问题，已有的研究已经证明，韵律跟句法不完全一致，这种不一致主要表现在它们的层级结构上。因此，我们不可能完全依靠句法的层级来预测韵律的层级。

然而，已有的研究同时也表明[6, 7]，韵律跟句法又是密切相关的，韵律的结构是以句法结构为基础的，例如，句法上不允许停顿的地方（如词内音节之间），韵律上就不可能停顿；而句法上的高层结构之间、特别是那些用标点符号标记的地方，韵律上就必定会出现停顿。不过，仅仅有这些还不够。现在人们最急切需要了解的是句子内部的韵律结构，希望通过研究句内的停顿跟句法结构和词性的关系，实施自动韵律切分。因为在书面文本上，句子内部通常没有标点符号；而在实际的话语中，语句内部，特别是较长的句子或短语内部，不但有停顿，而且具有不同的等级，正是这些不同等级的停顿标示着语句的韵律结构和层次。因此，我们必须着力研究语句内部的韵律停顿跟句法以及词性的关系，以便充分利用文本上具有的语言学信息本研究就是以此为目标的。

本研究有关语句内部停顿与句法结构的实验分析结果详见文末附表。

总的说来，语句内部停顿分布与句法结构之间没有始终一致的相关关系，这可以从上表的数据比较离散这个现象得到说明。不过，相对说来，各级停顿的分布还是具有一定的倾向性。兹概括如下，以供有关方面参考。

#### 3.2.1 一级停顿出现规律

一级停顿绝大多数出现在主语跟谓语之间，这在比较平衡的句子里尤为典型。

在不平衡的句子里，一级停顿的分布比较离散，跟二级和三级停顿都有较多的交叉，较多出在：

- 1) 前置状语跟主语之间；
- 2) 复合句的小句之间；
- 3) 复杂偏正结构的不同修饰语之间。

#### 3.2.2 二级停顿出现规律

二级停顿分布相对离散，跟一级和三级停顿都有较多的交叉，不过，主要集中在：

- 1) 比较复杂的偏正结构的不同修饰语之间；
- 2) 述宾之间，包括动宾之间、介宾之间或系表之间；
- 3) 在不平衡句里，常常出现在主语跟谓语之间；
- 4) 修饰语与中心语之间，包括定语跟中心语或状语跟中心语之间。

#### 3.2.3 三级停顿出现规律

三级停顿跟二级停顿的分布存在一定的交叉，但主要集中在：

- 1) 修饰语与中心语之间；
- 2) 动宾之间、介宾之间或系表之间。

#### 3.2.4 四级停顿出现规律

四级停顿的分布范围也很广，但主要集中在：

- 1) 不带“的”或“地”短语或词组的两词之间；
- 2) 带“的”名词短语或词组的“的”字之后，以四字格为最常见。

### 3.3 停顿分布与词性的关系

- 1) 人名 (尤其是外国人名)、地名之后;
- 2) 连词“和”、“及”、“跟”、“与”、“同”(连接两个复杂成分时)之前和之后(之前的常常大于之后的);
- 3) 助词“的”、“地”、“得”、“着”、“了”、“过”(后随多音节结构时)之后。

### 3.4 实际言语中的韵律词组词规律

如前所说, 话语的韵律切分问题实质上就是言语过程中的组词断句问题, 既涉及知觉上的停顿分布, 又涉及说话时语音上的组块分合。因此, 寻找韵律切分的规律, 同时也是寻找组字成词的规律。

根据本实验获得的文本信息和语音信息, 我们发现, 在实际的话语中, 为了满足韵律上的需要, 单音节的语法词都会或者通过延长该音节, 或者前附或后附到另一个标准音步上, 或者跟前后其它的单音节词组合, 设法凑足一个音步, 从而构成一个韵律词。不过, 这种韵律词的组词并不是任意的, 而是遵循一定的原则, 具有一定的规律(不计那些跟句法词一致的韵律词)。

实际言语中的韵律词组词的主要原则是: 第一, 单音节词处于重读地位, 或者因为左邻右舍已经是完整的音步而无所依附时, 就通过延长而构成独立的音步; 第二, 单音节词彼此粘附或与其它双音节词组合成音步时, 虽然具有很大的自由度, 但仍然要受各自词性的制约, 根据不完整的统计, 大致规律可以概括如下:

- 1) 单音节的动词(如: 达、离)、介词(就、用、以、比、从)、副词(如: 近、已)、连词(如: 因、但、而、与、可、)通过延长而构成韵律词(一个音步);
- 2) 单音节名词+“里”/“外”/“际”(如: 省里、组里、省际、部际、省外、所外);
- 3) 单音节名词/形容词+单音节或双音节名词/形容词=常用的简称(如: 建行、央行、北大、南航、二外、女一中、铁四中);
- 4) 双音节名词+单音节名词=三音节地名(如: 海南省、四川省、海淀区、三元里);
- 5) 双音节动名词+单音节名词(如: 抵押人、收押期、中介人、生意人、制造商);
- 6) 单音节或双音节数词+量词(如: 一批、两位、十二所、三十个);
- 7) 双音节动名词+时间词(如: 开拍前、到达后);
- 8) 双音节动名词+动词(如: 预言道、挥手说);
- 9) 单音节动词+单音节形容词(如: 抓紧、走开、打断、关严、放松);
- 10) 单音节/双音节动词+“前”、“后”/“了”、“着”、“过”/趋向动词“到”、“来”、“去”、“进”、“出”、“上”、“下”(带或不带“了”)(如: 睡前、走后; 死了、创造了; 吃着、制造过; 来到、回去、出来、看出、走进了、走上了、打下了);
- 11) 单音节的能愿动词+单音节动词(如: 应让、应是、愿作、会给、会出);
- 12) 动词“有”+“个”、“些”、“点”/“了”、“着”、“过”(如: 有个、有些、有点、有了、有着、有过);
- 13) 单音节的副词+单音节的形容词(如: 较低、最细、不甜、很大、最好、很老、较白);
- 14) 单音节副词+单/双音节动词/动补结构(如: 正娶、也有、再如、刚结束、才踏上);

- 15) 单音节的副词+介词+人称代词(如: 就把我、也以他、又将你);
- 16) 单音节副词+形容词/名词(如: 正旺盛、最清静、近几年、头两天);
- 17) 双音节的副词+单音节的动词(包括能愿动、助动、系动)(如: 一定能、必定要、或许会、特别是、尤其要、可以凭);
- 18) 双音节副词+“地”(如: 轻松地、迅速地、简单地、自信地);
- 19) “还是”/“就是”+人称代词(如: 还是我、还是他、就是你、就是他);
- 20) “也”/“还”+“将”+“于”(如: 也将于、还将于);
- 21) 单音节形容词+单音节名词(如: 强将、弱兵、少妇、老夫、期房);
- 22) 单音节形容词+双音节名词/动名词(如: 反倾销、抗病毒、高消费、低脂肪);
- 23) 双音节形容词/名词+“性”(如: 积极性、灵活性、功能性、动词性);
- 24) 代词+单/双音节名词或数词/量词(如: 这一、那一、那时、这类、哪种、哪几个);
- 25) 单音节指代词+双音节名词(如: 该房屋、该男子、那孩子、这报纸、那老师);
- 26) 单音节疑问代词+动词(如: 谁知、哪有、怎想到);
- 27) 双音节动词/形容词+“的”(如: 生产的、制造的、买回的、悲伤的、漂亮的);
- 28) 单/双音节的名词、代词+“的”(如: 人的、军队的、这里的、谁的、哪里的);
- 29) 单音节的连词+单音节的介词/系词/助动词/连词(如: 则由、则应、但应、又跟、又因);
- 30) 单音节连词+双音节名词/代词(如: 与内地、和父母、同军队、跟他们、跟那些)。

上述研究结果既可以为提高 TTS 合成语音自然度提供实践对策, 又可为提高自动识别的正确率提供理论依据。

### 参考文献

- [1]冯胜利, 《汉语的韵律、词法与句法》, 北京大学出版社, 1997年;
- [2]王洪君, “汉语的韵律词与韵律短语”, 《中国语文》, 2000年第6期;
- [3]曹剑芬, “汉语节奏的声学语音学特性”, 吕士楠、初敏编《现代语音学论文集》, 金城出版社, 1999年;
- [4]曹剑芬, “汉语口语的节奏”(英文), 第6届国际口语处理会议论文集, 北京, 2000年;
- [5]郑波等, “汉语的常规重音”(英文), 《第6届国际口语处理会议论文集》, 北京, 2000年;
- [6]叶军, “停顿的声学征兆”, 《第三届全国语音学研讨会论文集》, 北京, 1996年;
- [7]叶军, 《汉语语句韵律的语法功能》, 复旦大学博士研究生学位论文, 1999年。

附表：语句内部的停顿分布（这是对一个朗读语料库中随机抽取的 57 个句子的分析测量数据）

停顿级别及分布	一级	二级	三级	四级
句法结构				
主谓之间	23	11	3	9
小句之间	3			
前置状语与主语之间	5	3		
并列宾语或表语之间	2			
动宾之间、介宾之间或系表之间	2	16	29	8
修饰语与中心语之间	2	10	48	3
偏正结构的不同修饰语之间	3	17		
并列谓语之间	2			
前置宾语与动词之间	1			2
宾语与补语之间	1	2	2	
小句主谓之间		2	3	
并列成分 1 跟连词之间		3	2	
连词跟并列成分 2 之间		1		2
动补之间		1	4	2
较短名词短语之间			4	
转折连词跟所连词语之间			2	
带“的”词组的“的”字之后				13
不带“的”词组的两词之间				47
动词词组的状中之间				10
四字熟语的两两之间				8
年份的前两位数跟后两位数之间、数词与量词之间或复合数词的两两数位之间				8
较短介宾结构的介宾之间				4
年份跟月份或季节之间				2
多音节序数词的“第”跟数词之间				1

## PHONETIC AND LINGUISTIC CLUES IN CHINESE PROSODIC SEGMENTATION AND GROUPING

*Jianfen Cao*

Institute of Linguistics, Chinese Academy of Social Sciences

### ABSTRACT

This paper tries to discuss the interrelation between prosody and syntax by clarifying some syntactic constraints in Chinese prosodic segmentation and grouping. The main attention is paid to search for (1) possible correlation between

prosodic breaks and syntactic construction; (2) possible correlation between prosodic breaks and POS; and (3) the role of syntactic and lexical information in prosodic word chunking.