

汉语普通话语音节奏的初步研究

曹剑芬

摘要 本课题通过对新闻广播语料的声学分析,研究汉语普通话的语音节奏。本文的重点在于揭示基本的节奏单元以及不同层次上各单元之间的分界信息。初步的实验结果表明,(1)从结构上看,普通话的节奏可以大致分为基本节奏单元、节奏群和句子三个主要层次。基本节奏单元主要是双音节和三音节组块,相当于韵律音系学上的标准音步和加强音步,而节奏群和句子则是基本单元组块的扩展和组合。(2)从功能上看,这些组块可以标志连续话语中各语音单元之间的疏密关系,大致相当于韵律词、短语和句子三个层面的实体。(3)从声学效应上看,基本节奏单元内部各音节间不存在韵律上的分界信息,而在不同层次上的节奏单元之间,确实存在着显著的韵律边界信息,主要包括边界前音节韵母的延长和/或停顿间歇以及特殊的边界标志性音高移动方式。但是,作为典型的声调语言,普通话的节奏边界信息除了具有一般语言的共性以外,还有它自己的特性。主要表现为,基本节奏单元以双音节和三音节组块为主,基本节奏单元之间边界标志性音高移动的具体形式,要取决于边界后起首音节的声调特征。

1. 引言

何谓节奏?从听觉感知的角度说,节奏是一种由语音的时长、音高或者轻重的规律性再现而产生的韵律感知模式(Trask 1996);从发音生理的角度看,节奏是说话时的一个个韵律节落(保尔·帕西 1933)。通常,人说话,总是想尽可能准确地表达自己的想法,以便听话人能够正确地理解他的意思。根据言语产生和感知的相关研究发现,幼儿呀呀学语时,每次连续发出的音节数从不超过7个,而人的整块记忆的跨度通常也是 7 ± 2 个音节(Kohno & Shima 1989; Kohno & Tanioka 1990),这是言语产生和理解的一种自然的时域控制机制。所以,即使说话人憋足一口气,也许能够一下子说出好几十个音节来,但是那样肯定不利于听话人理解。由此可见,语音的节奏不仅由于发音生理的制约,更是出于语言表达和听觉理解的需要。语义的准确表达和正确理解,都要求把连续的话语分割为若干片段或组块,而语音的节奏就起这种分组的作用。

从客观实际来看,话语的节奏跟一个语言的语音信息分布方式有关,它反映出语流中各个语音单元之间的疏密关系。通常,一段连贯的话语总是由若干基本的信息单元构成,由于各单元在话语中的位置和地位的不同,它们又可依据彼此关系的紧密程度而结合为较大的组块,从而构成话语节奏的不同层次。人的耳朵对于这种信息相当敏感,通常感知为话语的不同程度的断连和停延。所以,一个语言里各个语音单元之间的断连停延及其有机结合,是语音节奏的核心。它遵循一定的规律,具体情况随语言而定。研究语言的节奏,实际上就是研究这种连与断及其相互呼应的关系。可以想象,在计算机语音处理中,假如能够恰当地处理好语音信息的断连停延问题,必然会极大地提高口语处理系统的性能和质量。于是,一个关键的问题,就是怎样客观地获得话语的节奏特征信息,以便为话语的自动切分和标记提供理论依据。

语音节奏特征的自动检测跟语音单元之间的连接和分界信息密切相关。一个语言的韵律

边界，固然要受它的句法和语义结构的制约，但是，单靠句法和语义信息，不可能准确地确定韵律上的分界，而一套超音段特征的规律性变化，却是韵律分界的一个很好的说明物 (Swerts etc. 1997)。当然，语音的分界信息同样也体现为音段层面的相应变化。例如，在不同层次的语音单元之间，音段间协同发音现象的有无及其程度的不同就体现了彼此疏密关系的差异 (许毅, 1986)。本文主要从超音段特征角度探讨普通话的节奏特点，至于音段层面的问题将另文讨论。

许多语言的相关研究已经发现，与语音的韵律边界相关的超音段特征，主要表现为停顿和/或边界前的元音延长以及特殊的边界标记性音高移动 (Blaauw, 1994)，其中，边界前的元音延长尤其具有重要的作用。例如，在荷兰语里，对于韵律边界的感知来说，甚至单边界前的元音延长就是一种足够的征兆 (de Rooij, 1979)。那么，对于汉语的韵律边界来说，是否也存在这样的相关信息呢？作为一种典型的声调语言，它又有什么样的特点？这正是本文将要讨论的重点。

2. 材料和方法

作为一种初步探索，本研究使用相当有限的实验材料，包括男女两个播音员约 3 分钟的新闻广播语料，考虑到文章的篇幅，这里只出示其中的两个实例。新闻广播语言通常为陈述语气，一般不包括特殊的重读和语调，因而可以从中窥测最基本的节奏特点。通过声学分析，测量语句中各项超音段特征，包括时域和频域两个方面。时域的参量主要包括词语的时长和停顿间歇时长；频域的参量主要是词语的最大和最小基频值。然后，对测得的参量进行分析，观察普通话语音节奏的层次以及各层节奏边界上的基本时长流变格局和音高移动规律，主要包括不同层面上停和/或延的使用规律以及边界前后音高阈限的沉浮和音域范围的涨缩行为。

由于语料有限，不便进行统计分析，本文采用实例解剖的方法，根据相关的时长和音高变化行为，揭示普通话新闻广播语言的节奏特点。

从句法和语义上看，下面第一段女声的新闻语料由四个句子组成，第二段男声的新闻语料由两个句子组成，而每一个句子又包含若干短语。为便于叙述，我们根据主观听辨试验时，多数人认为具有比较明显停顿印像的位置，把实验材料分为若干片段，按顺序用数字标码如下。

女声：

- (1)中央军委主席江泽民(2)日前签署命令
- (3)颁布实行(4)中国人民解放军(5)军事交通运输条例
- (6)这是我军第一部(7)全面规范军事交通运输工作的基本法规
- (8)它的颁布实行(9)标志着我军的军事交通运输工作(10)进入了法制化管理的新阶段

男声：

- (1)红塔人扶贫(2)走的是开发扶贫的路子
- (3)他们说(4)不能只是冬送寒衣春送粮(5)而要从生活科技和农业基础设施上下功夫

3. 结果和讨论

3.1 结果

3.1.1 时长

根据测得的原始时长数据，用每一片段的音节韵母平均时长作为标杆，计算出各个音节韵母的归一化时长。表 1 和表 2 给出了男女声各一个实例的数据。

表 1 女声韵母归一化时长(%)及片段间无声间歇(G)时长(ms)

(1)中0.63央1.37 军0.83委1.06 主0.80席1.03 江1.03泽0.52民2.15 (G:134.4ms)
(2)日0.80前1.35 签0.88署0.60 命1.01令0.96 (G:459.4ms)
(3)颁1.26布0.85 实0.61行1.86 (G:68.8ms)
(4)中1.03国0.89 人0.67民0.73 解0.67放1.12军1.56 (G:130.6ms)
(5)军1.18事1.18 交1.03通1.14 运0.97输0.53 条1.04例0.63 (G:606.3ms)
(6)这0.60是0.71 我0.99军1.18 第0.64一0.88部1.75 (G:37.5ms)
(7)全1.19面1.39 规1.16范1.67 军1.04事0.72 交0.91通0.92 运1.15输0.74 工1.02作0.67的0.79 基0.68本0.88 法1.11规1.18 (G:546.8ms)
(8)它0.55的1.41 颁1.31布0.69 实0.63行1.60 (G:253.1ms)
(9)标1.35志0.75着1.06 我1.38军1.28的1.73 军1.07事0.60 交1.00通0.95 运1.14输0.58 工1.03作1.16 (G:300ms)
(10)进0.97入0.71了0.91 法1.32制0.79化1.15 管1.35理0.65的0.60 新0.68阶0.67段0.93 (G:1112ms)

表 2 男声韵母归一化时长(%)及片段间无声间歇(G)时长(ms)

(1)红1.25塔0.67人1.05 扶1.00贫1.90(G:178ms)
(2)走1.07的0.52是1.42 开1.58发1.58 扶0.80贫0.80的0.45 路0.87子0.00 (G:719ms)
(3)他0.39们0.66说1.08(G:528ms)
(4)不0.56能0.81 只1.36是1.47 冬1.29送1.29 寒1.62衣1.93 春1.05送1.05 粮1.66(G:484ms)
(5)而0.72要0.83从1.18 生1.35活1.36 科1.12技0.81 和0.72农0.72业0.97 基0.97础0.00 设1.01施0.89上0.52 下0.56功0.85夫0.85(G:2000ms)

3.1.2 音高

根据测得的原始音高数据，分别用每一个音节的音高上限和下限值为坐标，标出各个片段的音高上下限移动的大致轨迹。图 1 和图 2 分别为男女声各一实例的音高移动示意图。

3.2 讨论

本文的研究对象是新闻广播语言，一般为陈述语气，没有特殊的重音。因此，作为节奏的超音段特征线索，这里主要考察各单元边界上的时长伸缩流变和音高升降沉浮，并据此界定基本节奏单元。时长方面，主要观察单元间的无声间歇和边界前的韵母延长，音高方面主要观察音高阈限的沉浮和音域范围的涨缩。

3.2.1 时长的伸缩流变

3.2.1.1 停顿与无声间歇

停顿是一种主观感知概念，是听感上与语音单元之间的间断相关的边界信息，它的客观相关物主要是无声间歇(叶军 1996)。从我们这里的表 1 和表 2 给出的数据来看，无论是男声还是女声，每一个话语片段的末尾都有或长或短的无声间歇段，从几十、几百到几千毫秒不等。如若仔细分析，这种无声间歇的时长差异在一定程度上反映出语音韵律结构层次的不同。

例如，就第一段新闻而言，在进行听辨试验时，百分之百的听音人认为，在片段(2)、(5)、(7)和(10)的末尾有大的停顿，这些正好对应于各个句子的末尾；有百分之八十的听音人认为，在片段(1)、(3)、(4)和(8)、(9)的末尾有较小的停顿；而只有百分之三十的听音人认为，在片

段(6)的末尾也有一个小停顿。对照表 1 的相关数据,我们发现,片段(2)、(5)和(7)末尾的无声间歇的确较长,都在五、六百毫秒上下,这些跟听感上较大的停顿是相适应的,尤其是片段(10)末尾的无声间歇,长达一千毫秒以上,因为它既处于句末,又处于整个话语段落的末尾。而比较起来,片段(1)、(3)、(4)、(8)、(9)末尾的无声间歇则要短得多,多半在二百毫秒左右。不过,其内部差异相当大,从六十多毫秒到三百毫秒不等。我们发现,这种差异往往同边界前的韵母延长互相补充,所以在听感上基本都对应于较小的停顿。至于片段(6)的末尾,只有 37.5ms 的无声间歇,它实际上是边界后“全”的声母的持阻闭塞段,而不是真正的停顿间歇,所以,大多数听音人对此没有停顿感。

再就男声的情况来看,在听辨时,百分之百的听音人认为,片段(2)和(5)末尾的停顿较大,百分之七十的听音人认为,片段(1)末尾的停顿较小,至于片段(3)和(4)末尾的停顿,只有百分之三十的听音人认为是较小的,而大部分的听音人则认为这两个停顿较大。从表 2 的无声间歇的数据来看,感知上停顿的大小层次基本跟声学上无声间歇的大小层次一致,即片段(2)和(5)末尾的无声间歇最长,所以听起来停顿最大;而片段(1)末尾的无声段最短,所以听起来停顿最小。不过,限于目前的实验数据,我们还无法判断,究竟多大数量级的无声间歇会被感知为较大的停顿,而低于这个水准的就被感知为较小的停顿。例如,对于男声片段(3)和(4)末尾的无声间歇时长,就其相对关系而言,它们介乎片段(1)和(2)、(5)之间,所以既可能被感知为较大的停顿,也可能被感知为较小的停顿;而就其绝对值来看,它们又都接近于女声的较大停顿情况下的无声间歇时值,因此,大部分的听音人把它们感知为较大的停顿。同时,主观听辨上的停顿感,不仅同无声间歇密切相关,而且还同边界前音节的韵母延长有关系,所以,实际的情况要比我们想象的复杂得多。详细情况将在 3.2.2 节里分析。

3.2.1.2 延宕与边界前韵母延长

延宕也是一种主观感知概念,是听感上与语音单元之间的关联相关的边界信息,它的客观相关物是边界前的音节延长(许毅 1986),主要表现为韵母的延长。总观表 1 和表 2 所列的韵母归一化时长数据,并对每一话语片段末尾音节的韵母时长跟非末尾音节的韵母时长进行比较,可以看出,在汉语普通话里,的确也存在边界前音节的韵母延长现象。不过,不能一概而论,究竟会不会延长或者延长的多少,则取决于边界层次的不同,具有一定的规律。客观的数据表明,凡是句中的短语末尾,例如女声的片段(1)、(3)、(4)、(6)、(8)、(9)以及男声的片段(1)、(3)和(4)的末尾音节的韵母都显著地延长了,有的甚至比韵母的平均时长长出一倍以上,例如女声的片段(1)末尾“民”字的韵母归一化时长为 2.15,男声片段(1)末尾“贫”字的韵母时长为 1.90。但是,作为句子末尾的情况就不同了。表 1 和表 2 的相关数据都表明,凡是句末的音节,其韵母几乎都没有延长。例如,女声片段(2)、(5)、(10)末尾的音节,其韵母不但没有延长,反而都有不同程度的缩短。又如男声的片段(2)和(5),其末尾韵母也都毫无例外地缩短了,尤其是片段(2)末尾的“子”字,其韵母干脆都丢失了。作为唯一的例外,只有女声片段(7)末尾的韵母略微延长,但并不显著。

其次,如若仔细观察每一片段内部的韵母时长分布情况,可以发现,韵母延长现象同样也存在于片段内部较低层次的语音单元边界之前。譬如说,女声的片段(1)由“中央”、“军委”、“主席”和“江泽民”四个较低的语音单元组成。在这些单元之间,又存在着一定的分界标志,边界前的韵母延长就是其中之一,在这里,各单元末尾音节的韵母时长分别为 1.37、1.06、1.03。这种现象在其它片段内部以及男声的各个话语片段内部也都普遍存在。不过,这种较低层次上的边界前延长现象比较复杂,它还要受到语音单元间关系疏密程度的制约,因而并不十分整齐划一。同时,它们还跟边界标记性音高移动行为相互作用,具体情况将在 3.2.2 节里讨论。

总的说来,句内各个片段末尾音节韵母的延长最为显著,片段内各个语音单元末尾的韵母延长次之,而句子末尾音节的韵母基本上不延长,而且大多数还要相对地缩短。

3.2.2 音高的升降沉浮

3.2.2.1 音高阈限的沉浮

从图 1 或图 2 上我们发现,在句子层面上,无论是音高高限还是低限,都呈现出自前向后逐渐下沉的趋势,尽管每一句的最高限不一定出现在句首,但总体走势是前高后低、逐渐下倾的。这是许多语言里都存在的作为陈述句边界标记性音高移动的特点。例如,从图 1 可以看到,第一句句首“中”的音高为 285Hz,而句末“令”的高限为 155Hz、低限为 135Hz。其它各句也都存在这样的走势。再看图 2 男声的情况,第一句句首“红”的音高高限为 268Hz、低限为 120Hz,而末尾“子”的音高为 75Hz;第二句句首“他”的音高为 133Hz,末字“夫”的音高为 60Hz,总体走势都是前高后低。为便于观察,我们在每个句子的首字跟末字的音高中值之间画一条连接线,由此不但可以清楚地看出这是一条前高后低的下倾线,而且可以看到句子与句子之间存在着明显的音高阈限错位现象,这就是通常所说的音高下倾线再设置现象,具体表现为,每一个句子起首的音高阈限,不但显著地高于本句句末的,而且也显著地高于边界前句末的。

从图 1 和图 2 还可以看到,在句子内部的各个片段层面上,也都存在着音高阈限的逐渐下倾趋势。例如,女声的第一句由“中央军委主席江泽民”和“日前签署命令”两个片段组成。片段(1)的首字“中”的音高为 285Hz,末字“民”的高限为 185Hz、低限为 125Hz。明显呈前高后低下倾走势;同样,片段(2)的首字“日”的高限为 260Hz、低限为 170Hz,而末字“令”的高限为 155Hz、低限为 135Hz,也是呈下倾走势。男声的情况也一样,例如其中的第二句包含“他们说”、“不能只是冬送寒衣春送粮”和“而要从生活科技和农业基础设施上下功夫”三个片段。从图 2 可以看到,每一个片段的音高阈限,也都是呈现前高后低的下倾趋势。可见这是广播语言里的一种普遍现象。

以上两个层面上的这种整体音高阈限的逐渐下倾现象是不同语言所共有的,它反映出发音生理方面的制约。因为人基本上都是利用呼气时发音,随着时间的推移,作为发音动力源的肺气压逐渐降低,声门下压也随着降低,而音高的变化跟声门下压具有密切的关系(吴宗济 林茂灿 1989),所以随着声门下压的减小,音高也就逐渐下降。

但是,值得注意的是,语音音高阈限的沉浮不仅要受发音生理的制约,更主要的是要服从语义表达和理解的需要。在不同的语言里,音高的作用不一样。因此,音高阈限沉浮的方式也会具有不同的特点。在汉语普通话里,从这里分析的有限资料和数据已经可以看出,其特点主要表现在更低层次语音单元的音高沉浮行为上。例如,无论从图 1 还是图 2,我们都会发现,在各个片段内部的更低层次语音单元之间,也都存在着音高阈限的再设置现象。不过,它跟上述句子和片段层面上的此类现象不完全相同。一个最大的特点是,边界前后音高阈限错位的方向,要依边界后首字的声调音区特征而定。我们都知道,普通话里的阴平具有高音特征,而上声具有低音特征。因此,在边界前后音高错位方面,似乎遵循高者愈高、低者愈低的原则。当边界后首字为上声时,这种错位就表现为边界后首字的音高阈限不是高于、而是低于边界前末字的现象。譬如说,女声片段(1)中“主席”的“主”的高限和低限都低于边界前“委”的高限和低限;片段(4)中“解放军”的“解”的高限和低限都低于边界前“民”的高限和低限。男声片段(4)中“只是”的“只”与边界前的“能”字之间,也表现出这种类型的错位现象。这种现象也许正是反应了汉语的特点。

3.2.2.2 音高音域的涨缩

首先,在每一个句子的末尾,音节的音域明显向内收敛。这可以从图 1 的片段(2)、(5)、(7)和(10)以及图 2 的片段(2)和(5)末尾音节的高限与低限间的动态范围观察到。例如,女声第一句末“令”的音高变化范围大约只有 20Hz,而该句内的“日”字同样也是去声声调,其音域就大得多,大约为 100Hz。又如男声第一句末“子”的音高变化范围几乎为零,而同一句中的“走”字,也是上声声调,其音域大约为 75Hz。可见句末的音域收缩是一种

普遍现象。

其次，通过跟句末音节音域的收缩相比较，我们发现，凡是句内片段末尾的音节，其音域常有相对扩大的倾向。例如，女声片段(1)末尾“民”的音域大约为 60Hz，这不但比该句末“令”的音域大，而且也是该句中音高动态变化范围最大的音节。男声也存在类似的情况，例如第一句中片段(1)末尾的“贫”的音域大约为 70Hz，比该句末“子”的音域要大得多。第二句中片段(4)末尾的“粮”的音域为 60Hz，比该句末“夫”的音域要大得多。这说明，边界前音节音域的相对扩大还是缩小，同样具有标志不同层次边界的作用。

3.2.3 音高沉浮涨缩与时长伸缩流变的配合关系

3.2.3.1 音域涨缩与韵母时长伸缩

联系 3.2.1 节中的时长流变情况，我们发现，上述句末与非句末音高动态范围的差异看来并不是偶然的，而是一种系统的差异，看来是出于语义表达的需要。我们看到，凡是句内片段末尾音节，不但音高动态范围相对扩大，而且总是伴随韵母的延长。例如，上述几例中的“民”、“贫”和“粮”的韵母归一化时长分别为 2.15、1.90 和 1.66；相反，凡是句末的音节，不但音高音域缩小，其韵母也没有延长行为。譬如说，上述几例中的句末音节“令”、“子”和“夫”的韵母时长分别为 0.96、0.00 和 0.85。在我看来，这正是不同层次边界效应的反映。因为在非句末的片段末尾，由于意犹未尽，于是，就用音域的相对扩大和韵母的额外延长来起提示下文的作用；而在句末，音域的缩小和韵母的缩短就给人以告一段落的感觉。

3.2.3.2 音高阈限沉浮与停顿时长流变和韵母时长伸缩

综观表 1、表 2 显示的停顿时长变化行为和图 1、图 2 显示的音高阈限沉浮行为，可以看到三个层面上的音高和时长变化的不同配合关系。这种不同的配合关系客观上说明了语音的节奏层次。

第一，在整句层面上，句末音高阈限的绝对下沉伴随着较长的停顿间歇，边界标志性音高错位、即边界后首音节的音高阈限显著高于边界前末尾音节的音高阈限，伴随着边界前音节韵母的相对缩短；

第二，在句内片段层面上，片段末尾音高阈限的相对下沉伴随着较短的停顿间歇，边界前后的音高错位现象伴随着边界前音节韵母的绝对延长；

第三，在片段内各个更低层次的语音单元末尾，一般不存在停顿间歇，但往往存在边界前的音节韵母延长。同时，边界前后也存在音高的错位，但错位行为跟句子层面和片段层面的不完全相同，如 3.2.2.1 节所述，边界后首音节音高阈限究竟是高于还是低于边界前末尾音节的，要依它的声调类型而定。

3.2.4 汉语节奏的层次和基本节奏单元

3.2.4.1 汉语节奏的层次

根据 3.2.1 至 3.2.3 的观察分析，一方面可以看出，在普通话里，不同语音单元之间的韵律分界信息存在着三个不同的层次，具体表现为不同层面边界上音高和时长变化作用的不同以及音高和时长变化之间的配合关系不同。总的说来，层次越高，无声间歇和音高错位的作用就越大；层次越低，韵母延长的作用就相对扩大。这就客观地标志出话语中各语音单元间关系疏密程度的不同：层次越高，语音单元间的结合关系就相对疏松；层次越低，各单元间的结合关系就相对紧密。因此，上述三个层面实际上代表着语音节奏的三个主要层次，这里把它们叫做基本节奏单元、节奏群和句子。前者对应于本文所说的片段内部的各个较低层次语音单元，后者对应于本文所说的片段和句子。另一方面，从音高和时长变化的基本方式来看，这三个层次又具有同构性，句子或节奏群的边界一定也是基本节奏单元的边界。

3.2.4.2 基本节奏单元

作为基本节奏单元，即本文所说的片段内各个较低层次的语音单元，其内部各音节间的结合关系，与节奏群内部各节奏单元之间或者句子内部各节奏群之间的情况显著不同。图 1、

图2和表1、表2的资料都表明，音高阈限的错位现象和边界前韵母延长现象不发生在这种低层次的语音单元内部。从这里分析的材料来看，这些低层次的单元主要为双音节和三音节组块。这类结构具有这样一些特点：第一，组块内部不可能出现停顿间歇；第二，其中的各个音节或音节结构之间不存在较高层面上的那种音高错位现象；第三，其中的各个音节或音节结构之间不存在较高层面上的那种边界前韵母延长。这表明，这种双音节和三音节结构内部结合关系非常紧密，不存在上述种种节奏边界效应，因而在韵律上构成一个独立的基本单元。

此外，基本节奏单元还包括一部分多音节组块，如女声的“颁布实行”、“中国人民”和男声的“冬送寒衣”等，多半为两个双音节结构的联合体，图中用虚线连接。这些组块跟上述双音节和三音节基本单元具有同样的特性。例如，在女声的片段(4)内包含着“中国人民”和“解放军”两个较低层次的语音单元，彼此之间既存在着音高错位，又存在着边界前韵母延长，因而分属不同的节奏单元；而在“中国人民”内部的情况则不一样，虽然它是由两个双音节词组成，但是，彼此间既没有无声间歇，也不存在音高错位，同时，词界前“国”的韵母只有0.89，相对于“中”字的1.03，显然没有延长，反而缩短了。所以，“中国人民”和“解放军”是同一个节奏群里的两个不同的基本节奏单元，而“中国”与“人民”在这里却属于同一个基本节奏单元。

以上的讨论仅就语音单元间的分界信息来界定节奏单元。事实上，节奏单元的界定势离不开轻重音的配置，这就要涉及汉语究竟是重音节奏型还是音节节奏型或者其它什么节奏型的问题，关于这一点，还须更进一步的考察。

4. 结论

本课题对于有限新闻广播语料的初步分析表明，在汉语普通话里，韵律节奏可大致分为基本节奏单元、节奏群和句子三个主要层次，基本句子单元主要是双音节和三音节组块，节奏群和句子可以是一个或者几个基本节奏单元的组合。同许多其它语言相仿，不同层次的节奏边界也是由若干超音段特征的特殊模式来表征的。在时间域，它们或者实现为无声间歇，或者实现为边界前的韵母延长，或者两者兼备；在频率域，主要实现为边界标志性音高移动，包括节奏单元边界两边的音高阈限沉浮以及音域涨缩的特殊模式。

从时长伸缩流变来看，不同层次节奏边界应用时长手段的方式不一样。具体说来，相当于整个段落或句子的节奏群边界总是由停顿来标志的，一般没有明显的边界前延长现象；而相当于短语的片段节奏群边界，总是由明显的边界前延长和相对较短的停顿来标志；至于短语内各个相当于韵律词的基本节奏单元的边界，一般只实现为边界前的韵母延长。

从音高的沉浮涨缩来看，每一个节奏单元的开始都由音高阈限的再设置来标志，因而各个层次的边界上都存在着边界前后音高阈限的错位现象，只是具体表现有所不同。在节奏群之间，边界前的音高阈限，包括高限和低限，总是低于边界后的；而在节奏群内各个基本节奏单元之间，虽然也存在着音高错位现象，但是具体情况则取决于边界后起首音节声调的音高特征，总的说，遵循高者愈高和低者愈低的原则，这是汉语不同于其它语言的主要特点。

以上这些信息，也许可以为语音的自动切分和标记提供较为客观的依据。不过，由于本研究所用的语料非常有限，所涉及的也仅限于陈述语句，由此得出的还只是一些初步的看法。如若引入不同的语调和重音，情况必定更加复杂。粗略的观察已经发现，汉语节奏的内涵十分丰富，有待于更加深入的探讨。

参考文献

- 保尔·巴西 著，刘复译 1933 《比较语音学概要》，北京，商务印书馆。
- 吴宗济 林茂灿(主编) 1989 《实验语音学概要》，北京，高等教育出版社。
- 许毅 1986 普通话音联的声学语音学特性，《中国语文》第5期。
- 叶军 1996 停顿的声学征兆 《第三届全国语音学研讨会论文集》。
- Blaauw, Eleonora (1994) The contribution of prosodic boundary markers to the perceptual difference between read and spontaneous speech. *Speech Communication* 14.
- de Rooij, J. J. (1979) *Speech punctuation: An acoustic and perceptual study of some aspects of prosody in Dutch*, Ph.D. Dissertation, Utrecht University.
- Kohno M. & Tomoko Tanioka (1990) The nature of timing control in language. In *Proceedings of ICSLP' 90*, Kobe, Japan.
- Kohno M. & Tsu Shima (1989) Rhythmic phenomena in a child's babbling and one-word sentence. *Bulletin No.191*, The Phonetic Society of Japan.
- Swerts, Marc & Mari Ostendorf (1997) Prosodic and lexical indications of discourse structure in human-machine interactions. *Speech Communication* 22.
- Trask, R. L. (1996) *A Dictionary of Phonetics and Phonology*. Routledge, London and New York.

A preliminary study on the rhythm in Mandarin Chinese

Cao, Jianfen

Rhythm in a language is closely related to the intonation of speech. Consequently, it is important for language learning and speech signal processing to clarify the rhythmic structure of natural speech. This study is concerned with the rhythm of Mandarin Chinese. As the basis of this study, a set materials selected from news broadcasting were used, pitch and duration measurements were made on the spectrograms, and perception tests on pausing and lengthening were conducted as well. This paper is devoted to the search in the rhythmic structure and the related boundary markers at different rhythmic levels. Three points emerge from the experimental results. (1) Structurally, the rhythm in Mandarin Chinese consists of three layers, namely, basic rhythmic unit, rhythmic group and sentence. The main basic unit of rhythm is disyllabic and trisyllabic chunk which is equivalent to a standard foot and heavy foot respectively in metrical phonology. Other polysyllabic structures are the extended form of these basic chunks, they form the other two layers of the rhythm. (2) Functionally, these chunks can mark different relational strength between syllables in running speech, and may work roughly as the prosodic substantiation at three levels, namely, prosodic word, phrase and sentence. (3) Acoustically, there is no boundary information related to rhythm within the basic rhythmic unit, while there do exist regular informations that show the rhythmic boundary between the basic units in different levels. As the boundary marker, the effects contain pre-boundary lengthening, pausing and their combination, as well as the special way of boundary pitch movement that including the variation of pitch threshold and pitch range. As a typical tone language, the boundary information in Mandarin Chinese not only show the universals existed commonly in other languages, but also characterized by its own property. The main points is that the basic rhythmic unit is disyllabic

and trisyllabic chunk, and the specific form of pitch movement must be restricted by its tonal features. The results summarized above should be particularly of benefit to the practice of segmentation in speech recognition and the connection in speech synthesis.