

语言的节奏

——贺王均先生八十华诞

曹剑芬

中国社会科学院语言研究所

1 前言

1.1 一般概念

作为语言的一种韵律(prosody)成分,节奏是很难定义的,因为在实际语流中,它还跟其它的韵律成分(例如重音和语调)相互关联、相互影响。而且,不同的研究又往往是从不同的角度来讨论。因此,现存的有关节奏的定义是多种多样的、有争议的。

从功能的角度看,节奏组织是跟话语的产生和感知相关的组块分合策略。现代认知心理学认为,言语理解是对话语的组块作出反应。事实上,无论是言语的产生还是感知,都是以组块的方式实施的。一方面,人在说话的时候,尤其是在说一个较长的句子或短语的时候,既不是一口气毫无停顿地说出来,也不是一个字一个字地蹦出来,或者一词一顿、离散地说出来,而是根据当时语义表达的需要,把若干个词组合成较大的节奏组块,然后一群一簇地说出来。另一方面,从感知的角度来看,人们也是对话语组块反映敏感,而不是逐字逐词地听辨。一些研究已经发现,这种现象是建立在人类认知机制的基础上的。Laver (1994)曾经指出:“言语计划的认知过程表现出典型的片断性(episodic),那就是,在言语的认知上总是以一个个的小片断作为单元式的实体来计划,也是以这样的小片断作为一个个完整的行为来表述和接收的”。这是人类语言里普遍存在的共同现象。因此,从这个角度来看,语言的节奏,本质上就是说话和听话时跟语义表达或理解相关的组词断句策略的语音体现,是由语义的表达和理解需要所决定的一种韵律上的结构模式。现在的问题是,这样的组词断句究竟根据什么样的原则?它们是怎样得以实施的?

1.2 汉语的节奏研究

关于汉语的节奏,近来有许多深入的研究和论述。但是,由于研究的角度不同,因而关于节奏的概念也各不相同。他们中有些从音联的角度看问题(例如,许毅 1986);有些从语文学或美学的角度来研究节奏(例如,吴洁敏, 1992);有的则结合重音和语调来讨论节奏问题(例如,沈炯, 1994);而有些则从句法结构的角看待语言的节奏及其相关的功能(例如,文炼, 1994; 张斌,

1998)。同时,还有一些研究或多或少跟节奏的声学特性相关,譬如说,有关语流中停顿分布的研究(例如,叶军, 1996; 华武, 1998);关于停顿和句法结构关系的探索(如,毛世桢, 1994);还有关于句法边界韵律征兆的研究(如,杨玉芳, 1997)。此外,越来越多的言语工程领域的学者,也通过在语调和短语切分等方面的建模工作,从节奏跟语调或短语切分相结合的角度,加入到对于汉语节奏研究的行列中来。所有这一切,都为我们打下了良好的基础,使得我们有可能对汉语的节奏作进一步的探索。然而,美中不足的是,时至今日,仍然很少关于汉语节奏一般概念的论述。正如有些致力于汉语教学和研究的国外学者抱怨的那样:“既不能在一般的教科书里找到(有关汉语节奏的)详细说明,也不能在通常的词典里查到明确的概念”(Triskova & Lomova, 1999)。

近十年来,本文的作者也进行了一些实验研究,力图从言语时域特性的角度来考察汉语的节奏问题(例如,曹剑芬 1991; 1992; 1994; 1995a; 1998; 1999)。这些考察的结果表明,言语的节奏在客观上跟语言信息的时域分布密切相关。

基于上述情况,本文将通过实验,结合汉语语音在时间域和频率域的相应特性,进一步讨论语言节奏的本质问题。

2 实验材料和实验方法

在目前这个研究中,采用了两组实验材料。第一组选自电视新闻联播,包括男、女各一个播音员的语音(以下简称新闻语音)。这种语音是典型的陈述语气,通常没有特别的强调重音或特殊的语气,因此,更适用于对节奏基本特性的考察。第二组材料选自一个男声的散文朗诵(以下简称朗诵语音),是通过广播录制的。这种类型的语音通常具有相对较慢的语速和较多的节奏间顿(break),通过这种语音资料,可以更细致地考察节奏的本质。

在这里,话语的节奏划分首先是根据听辨试验来确定的。在这样的试验里,要求听音人标注出觉察到的节奏间顿及其强弱等级。同时,通过这些语音的声谱,测

量了跟节奏分合相关的音高和时长变化参量。然后，根据获得的实验结果，集中描写节奏的层级结构及其相关的声学语音学特性。最后，以此为基础，讨论跟节奏的本质相关的一些问题。

3 实验结果和讨论

3.1 节奏单元的划分及其相关声学语音学表现

通常，说话人根据主观直觉和常识对词语进行归纳和分组。那么，主观直觉和常识是什么？显然，这是一种看不见、摸不着的抽象关系，它必定要通过某种物质

形式体现出来，才能构成话者与听者之间沟通的物质基础。那么，这种物质形式是什么呢？许多语言的研究表明，一套超音段的韵律特征是话语组块结构的很好的说明物(Swerts,1997)，它是底层语义上的组词断句规则在表层语音上的反映，主要体现为语音时长的规律性停连伸缩和音高的标志性升降起落，通常包括无声间歇、边界前（即单元末尾）音节延长以及音高音阶的下倾走势及其重新设置。在我们的实验结果中，也观察到了这些标志性的语音特性。而且发现，这些参量的变化所表现出的层次结构，都是跟语感上的节奏层级相应的。

表 1 韵律短语开头和末尾音节的时长分布：（1）在待续短语里；（2）在结束短语里

说话人	总体音节时长		短语起首音节时长		短语末尾音节时长	
	平均值（毫秒）	标准差	平均值（毫秒）	标准差	平均值（毫秒）	标准差
女声	179 60		(1) 168 47 (2) 190 56		(1) 298 61 (2) 177 33	
男声	155 59		(1) 154 53 (2) 219 51		(1) 250 106 (2) 142 26	

表 2 中等节奏边界和大节奏边界上的时长调节行为

边界前音节里的韵母时长比(%)			边界上的无声间歇平均时长(毫秒)			
			a.句子间的 b.段落间的			
短语类型	待续短语 平均值 标准差	结束短语 平均值 标准差	待续短语 平均值 标准差	结束短语 平均值 标准差 平均值 标准差		
女声	1.68 0.31	0.93 0.23	158 97	a.538 74 b. 1112 0		
男声	1.46 0.39	0.85 0.0	397 191	a.719 0 b. 2000+ 0		
新闻广播	1.60 0.34	0.90 0.20	278 169	a.583 109 b. 1556 628		
散文朗诵	1.33 0.21	1.05 0.28	52 34	a.480 87 b. 2000+ 0		

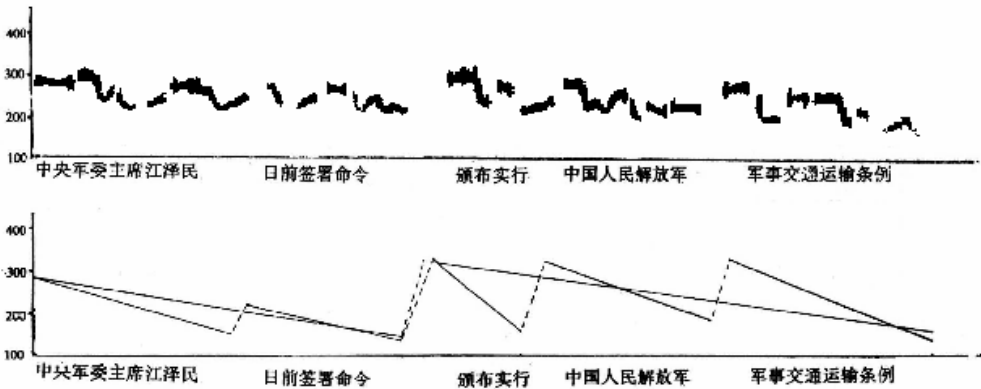


图 1 话语音高变化图例

本实验的主要结果可以用表 1、表 2 和图 1 来概括。表 1 出示了节奏单元开头和末尾音节的时长数据，其中

（1）和（2）分别代表待续短语和结束短语里的情况。所谓待续短语，就是语句内部在语义上尚未完结的短语，

也叫悬念子句；结束短语就是语句末尾在语义上已经完结的短语，也叫结束子句。表2显示了节奏边界上的时长调节行为，包括边界前音节里的韵母时长比(%)和边界上的无声间歇平均时长(毫秒)，其中a项代表句子间的无声间歇时长，b项代表段落间的无声间歇时长。图1出示了新闻话语中两个例句的音高运动变化的例子，图的上部是它们的原始音高变化显示，图的下部是根据原始音高所画的音阶走势图。稍后，我们将讨论这种音高变化跟节奏划分的关系。

3.2 汉语节奏的层级结构

根据感知评判结果以及相应的声学语音学参量所显示的内聚性和分界性标志，汉语普通话的节奏可以大致分为三个基本层次。如果借用一般韵律音系学的术语，那就是：韵律词、韵律短语和语调短语。韵律词是最基本的节奏单元，大多数情况下相当于词汇上的词，但也可能大于或小于词汇的词。韵律词通常由两个或三个音节组成，在某些特殊的情况下，也可以由单个音节充当。在那种情况下，这些单音节词都会被显著地拉长，从而在节奏上跟两音节或三音节的韵律词相匹配。作为节奏的基本单元，韵律词是构成汉语节奏组织的基本建筑构件。它可以进一步组成较高层面上更大的节奏单元，也可以直接充当这些层面上的节奏单元。韵律词内部各个音节之间的关系是非常紧密的，所以，它是某些音系过程作用的范围，例如，连读变调行为通常就是在这个范围内发生的。因此，这种基本单元大致相当于节律音系学 (metrical phonology) 里所说的标准音步 (standard foot) 和加强(超)音步 (super foot)。

韵律短语是介乎韵律词和语调短语之间的中等节奏组块。它可能小于句法上的短语，在有些文献里，这种单元也称为音系短语或者语音小句 (phonemic clause)。如前所说，言语的生成和感知的过程两者都是以组块策略为基础的，而在这些过程中，正是语音小句作为语音、句法和意义一揽子的包(块) (package) 发挥作用的 (Boomer, 1978)。它是言语里最常用和最重要的节奏组块。

语调短语是言语里的大节奏群，通常由若干韵律短语构成，一般对应于句法上较短的句子。在具体情况下，也可能直接由一个韵律短语、甚至一个韵律词充当。语调短语是句调信息的基本载体，话语的不同语气特点通常体现在语调短语上。

3.3 节奏组块的内聚特征和分界标志

3.3.1 不同层次节奏单元内部的内聚特征

在韵律词内部，作为一个单元的内聚特征，主要表现为音高变化模式和时长分布模式的相对稳定性。具体说来，就是保持基本的连读变调格局 (吴宗济, 1991) 和以词重音对立为基础的时长结构格局 (曹剑芬, 1989; 1990)。词内各个音节结合紧密，彼此之间不出现韵律上的无声间歇。

在韵律短语内部，主要通过韵律词间连而非断、即延而不停的粘接方式来表示内部的凝聚关系。同时，其内聚性还表现为短语总体音阶的逐渐下倾倾向。

在语调短语内部，也就是句子层面上，节奏上的凝聚性主要也表现为句子总体音阶的不断下倾倾向。同时，从时域来看，还表现为内部各个韵律短语末尾音节韵母的显著延长，这是一种特殊的表示持续关系的粘接方式。

所谓音阶下倾，指语句总体音高音区 (pitch register) 的不断依次下降 (declination) 的走势，即所谓降阶现象。例如，图1的两个例句分别由两个和三个韵律短语组成，每个句子中各个韵律短语的音阶呈现依次下倾的走势，从而形成一个独立的整体。

3.3.2 不同层次节奏单元之间的分界标志

从表1反映的时长伸缩情况来看，不同层次节奏边界上的分界标志不一样。具体说来，语调短语之间的边界总是由较长的无声间歇来标志的，一般没有边界前延长现象，绝大多数情况下反而是缩短的；而韵律短语之间的节奏边界，则总是由明显的边界前延长和相对较短的无声间歇来标志；至于韵律词之间的节奏边界，一般只实现为边界前的韵母延长。而且，韵律词之间的这种边界效应还多半会被消抹掉，因为在实际话语里，尤其在语速较快的新闻广播中，最常见的节奏组块是由若干韵律词组合成的韵律短语，其中某些韵律词间的语义关系相对紧密，因而彼此间的节奏边界就会被淡化。

从图1所显示的音高沉浮涨缩来看，节奏单元边界上普遍存在着音阶重新设置 (resetting) 现象，也就是不同节奏单元之间发生的音阶错位，这是另一个重要的边界标志性特征。具体地说，就是每个节奏单元开头的音阶总是高于前一个单元末尾的音阶，这种错位现象可以从图1 (下部) 清楚地观察到。通常，出现在韵律短语之间或者语调短语之间的音阶重新设置，要比出现在韵律词之间的来得突出。这种音阶重新设置现象，是各个语言里普遍存在的节奏边界效应。不过，在汉语普通话里，还有它自己的特点。主要是，韵律词之间的音阶重新设置的具体方式必须受声调的制约。那就是，如果

遇到后一单元开头的音节为上声声调时,由于其区别性特征为低,因此,它的音阶并不总是高于、而往往是低于前一单元末尾音节的音阶。例如,比较图 1(上部)中“军委”和“主席”的音阶,就可以看到“主”的音阶低于前一单元末尾音节“委”的音阶。在汉语里,音阶重置的总的趋势是遵循高者愈高和低者愈低的原则,这也许是汉语作为声调语言不同于其它语言的又一个特点。

4 一般讨论

4.1 节奏跟言语时域行为的关系

在传统的语音学理论中,节奏被看作某种语音单元在时间上的等间隔出现现象。例如,人们熟知的英语是典型的重音节节奏语言,意思是英语中重音的出现是周期性的;而法语是典型的音节节奏语言,意思是法语的音节是等长的(Pike, 1946; Abcrombie, 1967)。这种观点已经扎根语言学界好几个世纪了,而且至今仍然相当流行。

近来,许多关于节奏的语音学研究已经把注意力集中到寻找这种等时性的声学证据,以便支持这种传统观念。可是,遗憾的是,这些研究却表明,节奏“跟重音之间的间隔时长毫无关系”(Dower, 1983),“无论在重音节节奏语言里还是在音节节奏语言里,都不存在等时性的可靠证据”(Arvaniti, 1994)。这些研究结果说明,所谓节奏的“等时”特性即使在英语和法语里也是不存在的。

尽管如此,作为一种韵律现象,言语的节奏的确是所有口头语言中普遍存在着的。人们注意到这种现象已经很久很久了,而且,在语感印象上,人们强烈地觉得,节奏是跟言语产生和言语感知的时域变化行为密切相关的。问题是,究竟什么是节奏的本质?怎样看待节奏的组块分合跟言语信息的时域分布之间的关系?

在一些早期的研究中,节奏被定义为气呼的节落(breath group)(Passy, 1930),也就是现在一些文献里常说的“气群”或“呼吸群”。显然,这是从生理机制的角度看问题,也许这正是把节奏跟所谓的“等时性”联系起来的一个原因。因为人的呼吸是大致等间隔的。可是,Passy 同时也指出过,这种组块行为还跟意群(sense group)有关,那就是,由于语义表达转折的需要,也为了能够给听觉感知提供足够的时间,因而每一个气呼的节落大致对应于一个简单的意义单元。于是,言语中声音的组块就或多或少要受逻辑原则的制约,而

这种意义上的转折是多变的,或者长一点,或者短一点,因此,各个组块并不一定是等长的。此外,他还解释说,气呼的节落还可以进一步划分为若干力的节落(force group)。通常,每一个这样的力的节落都是由两到三个音节组成的。这些音节在意义关系上非常紧密,而其中必有一个音节在意义上是更加重要的,因而每一个力的节落中必有一个音节是重读的。当语速相对较慢的时候,这种力的节落也可以直接充当气呼的节落。因此,根据 Passy 的论述,言语的节奏组块分合不仅不必是时间上等长的,而且在大多数情况下,看来也不可能是等长的。

4.2 汉语节奏特点的启示

许多人认为,汉语的音节是等长的,因此,汉语是音节节奏语言。但是,根据作者先前的实验研究以及其他人的相关研究,汉语的音节并不是等长的。在实际语流中,由于音节所处的位置和轻重地位的不同,其实际时长具有极大的变化范围(曹剑芬, 1990-1991; 1995a)。目前这个实验研究的结果,也没有提供可以证明汉语是以音节的“等时性”为基础的音节节奏语言的证据。在这里,由于对节奏的本质尚未达成一致的看法,我们暂且不讨论汉语究竟是不是音节节奏语言的问题。不过,我们的确发现了一些有趣的现象,也许有助于了解节奏的本质。首先一点是,节奏单元的长度变化范围大致有定,也就是说,每个节奏单元所包含的音节数目虽然多寡不等,但是接近于一个常数;另一个有趣的现象是,每个节奏单元内部以及不同的节奏单元之间,都存在着相对不变的时长分布关系。

4.2.1 节奏单元内的音节数目跨度大致有定

在 3.2 节中我们曾经说过,汉语的韵律词多半由两到三个音节组成,这说明,作为基本节奏单元,其中的音节跨度是大致有定的。而作为中等节奏单元的韵律短语,它所包含的音节数目也总是在一定的范围内变化。

根据本研究所获得的数据,汉语里的韵律短语一般由两个或两个以上的韵律词构成,所包含的音节数目也多寡不等。可是,变动的范围大致有定,其跨度一般为 7 个音节左右,最多不超过 9 (7+2) 个音节,看来这是一个常量。一旦超过这个限制,如果没有采用韵律上的停延手段,就必然会采用“的”或“和”之类的功能词来分节,从而使得韵律短语的长度仍然保持在那个常量之内。此外,我们还发现,这种中等节奏组块的具体长度,还跟总体语速快慢以及语体的不同有关。例如,从

新闻话语测得的数据来看,极大部分韵律短语的长度都在 7 个音节左右,而在散文朗诵话语中,极大部分在 5 个音节左右。

以上观察到的韵律短语长度的限制并不奇怪,因为类似的现象也出现在其它语言里。例如,早在 1978 年,Boomer 就报道过:“在自然话语里,存在着可辨别的‘组块’,即若干音节的音节串,通常为 1 至 7 个或 8 个音节作为一个单元一起说”。而且,一些心里语言学家(如 Dittman 等, 1967; 1968) 也指出,这样大小的小句正是言语编码、同时也是言语解码的合适的心里语言学单元。此外,根据其它领域的一些相关研究,对每一次整体说出的音节系列的记忆跨度大约为 7 ± 2 (Miller, G, 1956); 在幼儿呀呀学语时、或者在单词句中,连续说出的音节数目从来不超过 7 个 (Kohn, M. & Tsu Shina, 1989)。因此,有理由认为,这里在汉语中发现的韵律短语长度的限制应该不是一种偶然现象,而是进一步证实了上述 Dittman 等的发现。它可能揭示了人类言语产生和言语感知过程中,在时域控制方面的一种相对不变的普遍规律。正因为如此,才在语感上导致了节奏单元似乎是“等时”的知觉印象。

4.2.2 节奏单元内部和节奏单元之间相对不变的时长分布关系

(1) 节奏单元首尾音节的时长调节

表 1 以韵律短语为例,显示了节奏单元开头和末尾音节的时长分布情况。从此表所列的数据可以看到,第一,短语首音节跟尾音节的时长存在着系统的差异;第二,这种差异的具体情况还以短语类型的不同为条件。具体地说,在待续短语里,首音节的时长接近或略短于话语音节总体平均时长,而尾音节则显著地比总体平均时长要长。可是,在结束短语里,情况则刚好相反,首音节长而尾音节短。而且,不管在男声还是女声话语里,这种系统差异都是一致的。看来,它代表了汉语普通话里的一般规律。短语开头和结尾所表现出的这种语速调节现象再次说明,言语的节奏并不是以某种语音单元在时间上的等长性为基础,而是以单元内部规律性的时长伸缩为基础的。

(2) 节奏边界上停与延的时长分布

从表 2 反映的时长伸缩行为来看,不同层次节奏边界上的停与延的规律性分布,是实现言语节奏的基本语音手段。它们相互配合、相互补充,构成相应的层级:在节奏的最低层面上,韵律词之间可能出现中等的边界前音节延长,但没有无声停顿;在节奏的最高层面上,

语调短语之间总是出现显著的无声停顿,但一般没有边界前延长;而在中等节奏层面上,韵律短语之间总是出现特别突出的边界前延长,并多半伴有较短的无声停顿。很清楚,正是这种在一定位置上出现的停与延的规律性分布模式,构成了言语节奏的物质基础。

5 小结

基于这个实研究的初步结果以及其它相关研究结果,起码可以得出下列两点看法。

(1) 语普通话的节奏结构包含三个基本层级:韵律词、韵律短语和语调短语。每一个节奏层面都有独特的、韵律上的内聚特征和分界标志。这些特征和标志都体现为一定的时域和频域的声学语音学参量变化模式,而且以一定的层级方式起作用。

(2) 节奏上的凝聚性和分界性,反映语言学单元之间在语义关系上的松紧程度。作为跟语义表达或理解相关的组词断句策略的语音体现,节奏是由语义的表达和理解需要所决定的一种韵律上的结构模式,主要通过规律性的音节时长伸缩和停延分布来实施。因此,节奏本质上是语言信息时域分布的规律性模式或规则,这种规则通常体现为某种韵律现象(例如,标志性的音高和时长变化)在一定位置上的规律性出现,而不是任何特定语音现象(例如重音)或语音单元(例如音节)在时间上的等间隔出现。这是存在于不同语言里的普遍现象,但具体特性可能随语言而异。

参考文献

- Boomer, D. S., 1978. The phonemic clause: speech unit in human Communication, in A. W. Sigman etc., (eds) Nonverba Behavior and Communication, Hillsdale NJ: Lawrence Erlbaum Associates.
- 曹剑芬, 1991 年,“汉语普通话音节的时长变化模式”英文,《第十二届国际语音学会议论文集》,法国。
- 曹剑芬, 1992 年,“汉语两音节词框架内的时长结构”英文,《第二届国际言语处理学术会议论文集》,加拿大。
- 曹剑芬, 1994 年,“句重音和词重音对语句内时长分布的影响”英文,《第三届国际言语处理学术会议论文集》,日本。
- 曹剑芬, 1995 年 a,“汉语普通话语句时长结构的基本格局”《中国语言学报》第 7 卷。
- 曹剑芬, 1995 年 b,“连读变调和轻重对立”《中国语文》1995 年第 4 期。

- 曹剑芬, 1998 年, “汉语普通话语音节奏的初步研究”, 中国社会科学院语言研究所语音研究室 1998 年《语音研究报告》。
- Dauver, R. M., 1983. Stress-timing and syllable-timing re-analyzed. *Journal of Phonetics*, 11.
- Dittmann, A. T. & L. G. Llewellyn, 1967. The phonemic clause as a unit of speech decoding. *Journal of Personality and Social Psychology* 6 (quoted from Laver, 1994).
- KOHNO, M & TSU, Shima. 1989. Rhythmic phenomenon in a child's babbling and one-word sentence. *The Bulletin* No. 191, The Phonetic Society of Japan.
- LAVER, John. 1994. *Principles Of Phonetics*, Cambridge Press, New York.
- MILLER, G. 1956, The marginal seven, plus or minus two? *Psychological Review*.
- PASSY, Paul. 1930. *Outlines of Comparative Phonetics*, Chinese version translated by LIU Fo, The Commercial Press, LTD., Shanghai.
- PIKE, K. L. 1946. *The intonation of American English*, 2nd edn, Ann Arbor, University of Michigan Press.
- TRISKOVA, H. & LOMOVA, O. 1999. Preliminary announcement of International Workshop on Tone, Stress and Rhythm in Spoken Chinese.
- 文炼, 1994 年, “汉语话语的节奏方面”《中国语文》第 1 期。
- 吴洁敏, 1992 年, “汉语节奏的周期和层次”《中国语文》, 第 2 期。
- 许毅, 1986 年, “普通话音联的声学语音学特性”《中国语文》第 5 期。
- 杨玉芳, 1997 年, “句法边界的韵律征兆”《声学学报》第 22 卷第 5 期。
- 张斌, 1998 年, 《汉语语法》, 上海教育出版社。

THE RHYTHM OF LANGUAGE

Cao Jianfen

Institute of Linguistics, CASS

ABSTRACT

In order to study the characteristics of the rhythm in Chinese, the author of the present paper have conducted several related investigations in the past decade. This paper try to discuss the essence of Chinese rhythm based on these investigations. According to the preliminary results obtained so far, we suggest that: (1) the rhythmic structure of Mandarin Chinese contains three basic layers, i.e., prosodic word, prosodic phrase and intonation phrase. Each of them is characterized by certain prosodic coherency and demarcation markers. These features and markers are all embodied as certain pattern of acoustic parameters both in spatial and temporal dimensions. (2) Both the coherency and demarcation of the rhythm are the reflection of semantic distance between different linguistic units. As the phonetic embodiment of the strategies in prosodic word formation or sentence segmentation, which is strongly

related to semantic expression or comprehension, rhythm is a kind of prosodic organization pattern that determined by the requirement in semantic expression and comprehension, and is mainly manifested through the lengthening and shortening of syllables and the distribution of pausing. Consequently, in this sense, rhythm essentially is the regular pattern or rule on temporal distribution of the linguistic information in a language. Generally, this pattern or rule is embodied as the regular occurrence of certain prosodic phenomenon, such as the marked variations of pitch and /or duration, taken place in certain position; in stead of the so-called “isochronic” occurrence of any particular prosodic event (such as accent) or speech unit (such as syllable) in time dimension. Rhythm is an universal phenomenon commonly existed in different languages, but the manifestation is language-specific.