

基于节拍概念的语言节奏时长描述系统

殷治纲

中国社会科学院语言研究所, 北京, 100732

文 摘: 本研究从语言节奏核心问题出发, 参考乐理知识, 对语言的节奏时长描述体系进行了探讨。

语言节奏的典型特征是准周期性, 但是我们发现目前以音节等语言单元时长为基础的研究并不能反映节奏单元的准周期性特点, 所以我们放弃了以“音节”为基础的韵律计时单位, 而代之以独立于语言单元的“拍”的计时单位, 并进而发现拍和节等乐律单位更有利于语言节奏问题的研究。

以节、拍等音乐计时单位为基础, 本研究建立起语言节奏新的时长描述体系, 为深化韵律节奏研究提供了新的思路。

关键词: 节奏; 等时性; 节拍;

中图分类号: H08

不管是江南的吴侬软语, 还是铿锵的三秦高腔, 当人们接触到一种新的语言时, 第一印象往往是它的节奏特性。

那么什么是语言的节奏呢? 按照定义, 语言的节奏, 是指“在言语或诗歌中由突显的要素有规则地间断出现所产生的知觉模式”(《语音学和音系学词典》)。

尽管语言节奏的定义并不复杂, 但在实际研究中, 却往往很难对某种语言的节奏进行直观而量化的描述。这一问题在很大程度上是由于目前缺乏完善的节奏描述系统。

本文将简要梳理当前语言节奏的相关研究, 并尝试对语言节奏的描述系统进行探讨。

1. 语言节奏的相关研究

语言学界对语言节奏问题的研究开展较早, 其研究重心是节奏的组成要素、组织模式和时长规律。

Mark Liberman 1975 年在音系学研究中提出了节律树 (metrical tree) 模式, 首次对语言节律进行了形式化描述。其模式核心思想是英语中轻、重二元对立及其多层次理论, 其具体形式借鉴了语法结构的树形图形式。1977 年 Liberman 和 Price 又推出了节律栅 (metrical grid) 模式, 以多层栅状结构代替树状结构, 改进了前者重音位置和单元界限不清晰的缺点。

Cummins (2002) 研究了影响不同节奏类型产生的因素, 建立了一个层级结构的二元节奏模型, 以二音步为基础, 可以向上以二元结构不断结合; 此外他还关注音步等小单元和更高层单元的交互关系。

李爱军 (1999) 对韵律节奏的组织模式进行了综合研究, 认为: (1) 汉语的韵律结构具有层级结构, 从小到大依次为: 音节、音步、韵律词、次要韵律短语、主要韵律短语和语调组。(2) 每一层韵律单位在停顿级别、音高曲线和重音级别上有各自特点。

郑秋豫 (2005) 提出了阶层式语流韵律架构模型。该架构模型的主要特征是 1) 由上而下的管辖制约; 2) 制约包括: 发声器官; 生理机制 (含呼吸); 认知限制 (语言规划的制约)。

吴洁敏等 (1991, 2001) 认为节奏不仅指音强和音长两个方面, 还包括徐疾、高低、长短、轻重、音色异同等对立因素在一段时间内有规律交替往复的组合形式。言语停顿分为句内停顿和局外停顿。内 3 层停顿: 音节——音步——气群——句子; 外 3 层停顿: 句子——句群——段落——篇章。

叶军 (2001) 认为节奏感知的要素包括 (1) 时间上的重复事件; (2) 轻重; (3) 口语和韵文节奏。他认为节奏的本质是轻重现象的有规律交替出现; 而其深层决定因素是信息结构。

王洪君等 (王洪君 2004; 刘献强 2009) 对汉语韵律节奏类型, 节奏和句法的对应关系以及韵律单元组块规律进行了研究。她认为汉语的节奏典型特征不是轻重关系, 而是松紧关系 (2004)。节奏是语流内松紧交替的回复, 在韵律词内不一定非有轻重之分。

Gibbon (2004): 认为节奏是层级性和线性结构结合的复杂结构。他用由局部时长得到的时间树预测器研究了节奏层次和句法层次的对应关系, 并发现二者具有相关性。

Pike (1945), Abercrombie (1967) 从类型学角度认为语言可以分为音节计时型 (syllable-timed) 和重音计时型 (stress-timed) 两类节奏性语言。其中, syllable-timed 节奏类型是 syllable 时长具有等时性(如英语、俄语等), 而 stress-timed 节奏类型是 stressed syllable 之间的时间间隔具有等时性(如法语等)。Ladefoged (1975) 则认为除了上述 2 种节奏类型外, 还有 mora-timed 语言, 其节奏单元是比 syllable 更小的 mora, 代表语言是日语等。

以上理论归纳起来可以分为两大类内容, 一类是探讨了节奏的组织模式(层级性和线性), 另一类是节奏的影响要素(轻重、停延、时长等)。但是在更狭义的节奏研究中, 节奏研究的核心领域还是集中在时长问题。

当前一些计量语言节奏的算法也都是从时长角度入手的。例如, Ramus et al. (1999) 提出了一种计算节奏的方法。他用 $V\%$ (元音占句长的比值)、 ΔC (辅音段长度的偏差)、 ΔV (元音段长度的偏差) 来反映节奏特征。但该模型也只是进行了数据的全局平均 (GE), 而没有更多语言学理论意义。Low et al. (2000) 提出了计量语言节奏特征的 GE 和 PVI 指标。PVI 可以计算相邻单元 (元音或者音节等) 规整后的差异性, 计算公式是 $PVI=100 \sum_{k=1}^{m-1} |2*(d_k-d_{k+1})/d_k+d_{k+1}|/(m-1)$ 。该公式仅较适用于严格的二元结构, 不太能反映整体时长的变化规律。

2. 现有节奏描述系统的局限

上述语言节奏的某些研究虽然也尝试给节奏描述定义出适用的特征或者指标 (如 PVI, ΔV 等), 但是这些特征或指标主要是为了从宏观角度来界定语言的节奏特性, 而不是为了描述具体的节奏内容。例如, 它们都不适用于描述具体一句话的节奏内容和特性。

另外, 现在基于“韵素 (莫拉) ——音节——音步/韵律词——韵律短语——语调短语——语句” (Selkirk, 1984; 王洪君, 1999; 刘献强, 2007; 叶军, 2008) 的韵律层级描述系统虽然比较成熟, 但是它对于描述节奏内容同样有不少局限。其中一个主要的局限是: 韵素 (莫拉)、音节和音步等虽然可以称为节奏 (或韵律) 单元, 但却不是严格意义上的节奏计量单位。这是因为计量单位一般应该是脱离语言单元而独立存在的一套客观系统。

这一计量标准问题不妨参考国际长度单位的演变过程。在古代英国等国家, 曾将成年男子单脚的长度公认为长度标准 (foot), 但是由于成

年男子的脚长并不完全相同, 所以上述以“脚”为单位的粗略长度计量法后来慢慢被更加精确而独立的客观计量单位“米”所代替。

在现在的节奏研究中也是如此。虽然现在音节常被作为常用的计时单元 (韵素和音步单元都可以认为和音节有密切的对应关系), 但音节本身由于受音素数目、声调类型、韵律位置、轻重等诸多因素影响, 其时长并不相同。要想对语言单元的节奏时长计时, 最好选择其它独立与语言单元的客观单位。

3. 基于节拍概念的新的时长描述系统

为了寻找语言节奏的客观计时单元, 我们不妨从节拍概念的最初领域——音乐领域中去寻找答案。

我们发现, 音乐体系中的曲式内容和节奏计时有两套不同的描述系统。

音乐的曲式内容系统从低到高由音符——动机——乐节——乐句——乐段构成。这种结构类似语言中的音节——韵律词——韵律/语调短语——语句——语篇 (段落) 的层级结构。

音乐中的计时系统则是独立于上述音乐内容的“拍”和“节”的系统。但是在语言节奏体系中则缺乏对应的计时系统, 而这也是我们认为应该借鉴并引入现代节奏研究的描述系统。

下面对音乐中的节拍概念做一下简单的介绍。音乐中的“拍”是纯计时单位, 其时长采用二进制计算。从大到小可以包括全音符、二分音符、四分音符、八分音符等。对于稳定时长, 还可以用附点 (增长原有音符一半时间)、双附点 (第二个附点增长第一个附点一半的时间)、延长记号等符号表示。以“拍”为基本单位, 音乐中建立起来了以“节”为基本单位的等时性 (周期性) 单元。

乐曲的节奏特点可以从其“拍号”反映。拍号用分数形式表示, 分子表示一小节有几拍, 分母表示以什么样的音符为一拍。例如, 我们常说的四分之二拍就是指以四分音符为一拍, 每小节有二拍。2/4 拍的音乐被认为是稳定的进行曲风格, 而 3/4 拍则有一定摇曳感的圆舞曲风格。

在以下四分之二拍示意图中, 虽然每小节的音符数目可能不同, 但是其时长 (拍子数) 都是一样的, 都为 2 拍 (表示停顿的四分休止符 Q 也占 1 拍)。

| CC | CDD | ID | B | CQ ||

上例中音符的拍数可写为 | 1, 1 | 1, 0.5, 0.5 | 1.5, 0.5 | 1 1 ||, 即每个小节合计都是 2 拍。

在此，我们认为“拍”和“节”的单位应该引入语音节奏体系研究。引入这些新的概念，不仅是因为它们便于直观描述语言节奏特点，还因为它们有一系列语言学依据的支持，并有其特有的优势。

1) “拍”是一个相对稳定的时长单位。因为我们假设说话人语速在局部范围内是稳定的，在此标准语速下，一个“拍”长应对应于标准单位音节时长。

2) “拍”是一个独立于音节的时长单位。它具有 Mora 和音节计时不具备地优势是可向下精确切分、向上组合以达到灵活应用。引入“拍”后，音节不再作为计时单位，而成为被计时的观测物。在遇到“非标准”语速音节——如延长的音节（例如强调重音或句末延长音节）时，可以用更多的拍子数来标记它；同理，缩短的音节（如轻音）时长可以用小于一拍的小数（分数）形式的拍子数表示。在此计时体系内，比音节级别更高的节奏单元也直接由“拍”来计时，而不再受音节数量的影响。这使得各级语音单元可以被更准确地计时。

3) “拍”同时又是一个和心理感知相关的时长单位。叶军（2008）认为“本质上节奏不是一种物理（声学）现象，而是一种心理现象，确切地说是一种感知经验”。语感经验告诉我们，人在说话时内心是有个“计拍器”来感知和控制节奏语速的。“拍”可以认为和人说话时的主观时间感关系更密切。

4) “节”是节奏的周期性单位。它反映了语言节奏的周期性特点。

接下来，我们不妨看看语言现象中存在的“节”、“拍”的实例。

首先，以节律感最明显的诗歌（五言绝句和六言绝句）为例来看看节拍的存在：

“白日依山尽，黄河入海流。欲穷千里目，更上一层楼。”（《登鹳雀楼》王之涣，五言绝句）

“心事数茎白发，生涯一片青山。空山有雪相待，古道无人独还。”（《开元乐》李煜，六言绝句）

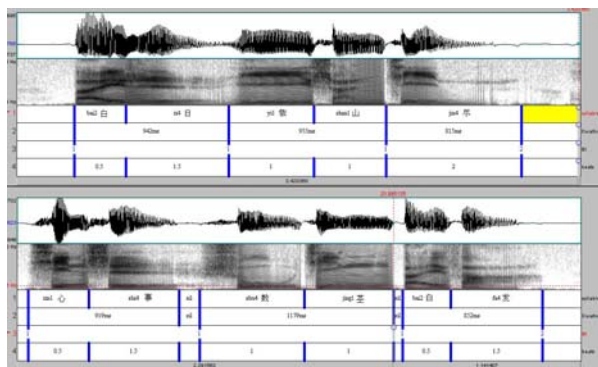


图1 五言绝句和六言绝句节拍图

从传统的音节数目分组看，六言绝句是典型的2+2+2结构，属于字数和时长周期性都均衡的结构。

五言绝句则一般认为是2+3结构或2+2+1结构（吴洁敏 2001），不管是哪种分法，最后的单字从音节数上都破坏了其周期性规律，在局部（一行内）属于一种不平衡结构。

不过，如果抛开以往音节计时的系统，而采用节拍计时系统，则我们很容易发现，六言绝句和五言绝句在节拍系统上一般都属于2拍+2拍+2拍的周期均衡性结构。

在此，我们对五言诗取2+2+1的分法。这样，可以认为五言诗和六言诗各自的前两节具有等时性，而着重研究在韵律位置相同、音节数不同的末节中，它们是否依然做到相对“等时”。在上图中，“白日|依山|尽”中的“尽”字尽管只有1个音节，但是在语言节奏等时性要求下，节拍数从听感上延长到了2拍（时长为815ms），与六言诗末节（双音节）“白发”有同样的听感拍数（时长852ms）。

接下来，我们进一步看一下，在非诗歌的正常朗读语句中，能否发现“节”单元的存在。我们对中国社会科学院语言所语音室制作的ASCCD朗读语料库中的数据按照“拍”单位进行了标注，发现在朗读语言中，“节”的存在依然明显。

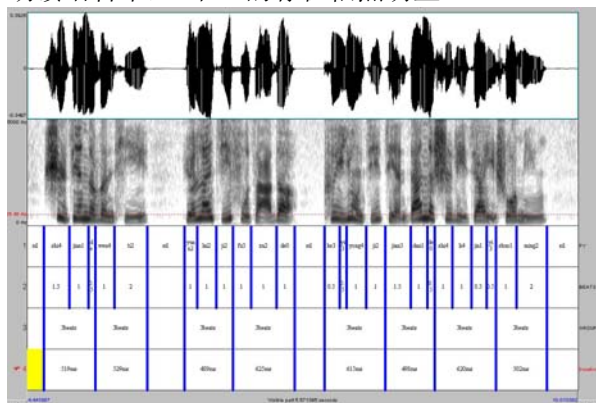


图2 朗读语料中的节拍

如图 2, 语料“世间的|问题, |原来极|复杂的, |可以用极|简单的|事例加以|说明。||”基本是以 3 拍一节的。我们用音乐体系的节奏表示法可以将其定义为四分之三拍节奏, 即以四分音符为一拍, 每小节包含 3 拍:

“|世间的| 问题, |原来极|复杂的, |可以用极|简单的 |事例加以| 说明。||”
|I C D|C B|C C C|C C C |DDCC |I C D |C C DD| C B||

通过以上节、拍单位划分, 我们不难给语言中的“周期性”要素以形式化标记, 从而使语言节奏的“周期性”特征得以突显。

此外, 我们在考察了语料后, 发现汉语普通话节奏和音乐节奏有一些类似的模式和特点:

1) 朗读语体每小节以 2-3 拍为主。该规律也常见于音乐谱式, 并且音系学研究中每个音步/韵律词多以 2-3 音节构成, 也属于类似规律。一个可能原因是一个周期往往指的是某个大单元的最大公约数, 而 2 和 3 正是数学上最常见的最大公约数。例如更大的四分之四拍可以被作为两个四分之二拍的整和, 而四分之五拍这种混合拍子则是四二拍与四三拍的混合。

2) 口语语体中, 随着语速的提高, 每小节拍数有增大的趋势, 可能出现复拍子(即两个相同单拍子的复合, 如 4/4 拍=2/4+2/4)和混合拍子(由不同类的拍子组合起来的拍子, 如 5/8 拍=2/8+3/8)类的节奏。

3) 较小韵律单元(如语句或韵律短语)内节奏往往体现一种拍子类型, 如四分之二拍; 较大的韵律单元内有时不可避免的出现变换拍子的现象(诸如前面是四分之三拍, 后接四分之二拍)。但在变换拍子前后的两个语音单元拍数总量上, 往往也趋向于相同, 例如前一个韵律短语是 3 拍+3 拍构成, 后一个可能是 2 拍+2 拍+2 拍的组合。这样, 在更大层级, 语言节奏也体现出一种等时的周期性特征。这种出现在小单元和大单元不同层级的节拍规律, 说明节奏也存在层级结构。

5. 语言节奏的描述体系

基于前面已经介绍过一些音乐的音符标记, 我们尝试把它们引入语言节奏体系描述每个音节所具有的相对时长(拍)。

| 名称 | 标记 | 拍数(四分音符为一拍) | 说明 |
|----------|-------|-------------|----------------------------------|
| 全音符 | A | 4 | |
| 二分音符 | b B | 2 | 占全音符的 1/2 |
| 四分音符 | c C | 1 | 占全音符的 1/4 |
| 八分音符 | d D | 0.5 | 占全音符的 1/5 |
| 十六分音符 | e E | 0.25 | 占全音符的 1/16 |
| 三十二分音符 | f F | 0.125 | 占全音符的 1/32 |
| 附点音符 | 如 I | 1.5 | 附点延长前面音符时值的一半 |
| 复附点音符 | 如 C.. | 1.75 | 第二个附点增长第一个附点一半的时间 |
| 拍号 | 如 2/4 | | 以四分音符为一拍, 每小节有 2 拍 |
| 小节线 | | | “节”的结束标记 |
| (节奏组)中止线 | | | 节奏组结束标记。一个节奏组后有大的停顿, 一般对应于一个韵律短语 |
| 分组标记 | () | | 用来连接一节内或跨节边界的几个特别紧密的单元 |
| 语速 | C=180 | | 表示当前节奏中, 一分钟的拍子数(本例中为一分 180 拍) |

在语言的节奏描述系统中, 可以直接标记“拍”子数值来表示语言单元的节奏长度。需要说明的是, 语音的节奏远没有音乐严格, 所以用不到全部上面的标记。例如从拍子数量上来说, 一般我们可以分辨的拍子长度短到 0.25 拍可能就够了, 再小的拍子已没有太多意义。






拍号是表示语言节奏感的重要标记, 由于说话的随意性远大于音乐, 所以语言中的变拍现象会更多, 拍子的种类也会较多。总的来说, 拍号和语言的两种特性相关, 一是语言的速度感, 二是语言节奏的分组。因为拍号是分数形式构成(如 3/4 拍), 其中: 分母表示以几分音符为一拍, 对于速度感快的语言(如日语), 我们可以用更短的音符(如八分音符)计拍, 中等速度语言可以用中等时长的音符(如四分音符)计拍, 而速度感慢的语言可以用更长的音符(如二分音符)计拍; 分子表示每小节的拍数, 如汉语朗读体一般每节 2-3 拍较常见, 口语体每节的拍数可能更多些。

表 1 节奏的时长标记符

此外,语言的语速还可以用拍速相对精确表示,由于人说话时某些局部的语速会发生变化,所以语言的拍数也可以根据变化给出一个范围(如拍速=180-250拍),或者直接在变速时给出特别标记。

和时长、组块相关的另一个概念是停顿。节奏中的停顿现象可以用休止符表示,具体符号见下表。值得注意的是,停顿在节中也占相应拍数。

表2 休止符(停顿)标记

| 五线谱记谱 | 简谱记法 | 休止符名称 |
|--|-------|--------|
|  | 0000 | 全休止符 |
|  | 00 | 二分休止符 |
|  | 0 | 四分休止符 |
|  | 0 | 八分休止符 |
|  | 0 | 十六分休止符 |
| | | |

6. 语言节奏中“拍”的计算

以上虽然列出了语言节奏的拍子描述系统,但是在实际计算中,如何定义一拍的长度呢。我们认为可以基于以下方法计算。

首先,由于语句中的标准拍(类似于2/4或3/4拍音乐中的四分音符)反映了平均语速,所以,可以将语句中的平均音节时长作为标准拍的参考长度。通过将语句中每个音节的实际时长和标准拍时长进行比值计算,就可以得到每个音节的实际拍长。用公式可以表示为:

$$BEAT_{ref} = DUR_{ut} / NUM_{syl} \quad (1)$$

$$BEATNUM_{syl} = DUR_{syl} / BEAT_{ref} \quad (2)$$

上式中, $BEAT_{ref}$ 表示标准拍的时长, DUR_{ut} 表示一个句子的总时长, NUM_{syl} 表示该句子所含音节数, $BEATNUM_{syl}$ 表示句中某个音节 syl 对应的拍子数, DUR_{syl} 表示音节 syl 的时长。

7. 不同语言的节奏特征描述

采用节拍概念的语言节奏描述系统不仅能描写语句内微观层面的节奏特性,还可以以拍号为特征从语言的宏观层面描述其节奏特性。

例如,按照目前语言节奏的划分,世界语言大致可以分为三种节奏类别,即:以英语为代表的重音计时型(stress-timed)语言;以法语、汉语为代表的音节计时型(syllable-timed)语言,如法语;以日语为代表的莫拉计时型(mora-timed)语言等。那么,如何用节拍概念来表示不同语种的节奏特性呢?本部分将以汉语普通话和英语为例,简要比较说明不同语言的节奏描述与类型对比。

对比研究采用了我们收录的多语言语音语料库,主要对比了各种语言朗读语体的语句和语篇(《北风与太阳》)的节奏差异,发现它们具有以下各自特点。

以汉语普通话朗读语体语速作参照标准,我们可以用四分音符为一拍给普通话进行计时;从语速来说,普通话每分钟拍数大约在180-280拍间;对于音节时长和拍子的对应关系来说,大多数音节占1拍,句末延长的音节一般占1.5拍到2拍,轻声音节或者部分语速快的音节占0.5拍;从小节组块规律看,普通话每2-3拍组成一节,故普通话最常见的节拍型是2/4拍和3/4拍。

英语则以单词为基本语义单位,每单词包含一到多个音节,故在以音节为单位的发音语速方面看,英语的音节语速要远快于汉语,所以我们可以用八分音符为英语音节计时;在小节分组方面,英语是以多个非重读单词围绕一个重音形成一小节,故每分钟小节数目一般不多,通常在20-40个,但每小节包含的拍数则很多,多在2-6个,故其节奏型可表示为4/8、6/8拍等。

8. 总结

本文初步探讨了语言节奏的描述系统问题。

文章的基本观点是,语言的节奏描述系统应该独立于语言单元(音节、韵律词等)本身。本文借助乐理知识建立的节、拍节奏描写系统符合上述原则。

该节奏描述系统不仅可以较好的描述语言微观和宏观的节奏特性,还便于进行工程计算,对语言节奏的定性描述和定量分析有一定理论意义和实用价值

参考文献

- [1] Aniruddh D.Patel, Hidden connections between linguistic and musical melody, *Speech Prosody* 2010, 2010
- [2] Cornelius, R.R., *The Science of Emotion. Research and Tradition in the Psychology of Emotion.* Prentice-Hall, Upper Saddle River, NJ, 1996.
- [3] Dafydd Gibbon, Tone and timing: two problems and two methods for prosodic typology, *TAL* 2004, 2004
- [4] Ee Ling Low et al., Quantitative characterizations of speech rhythm:

- Syllable timing in Singapore English. *Language and speech*, 43, 2000
- [5] Frank Ramus et al., Correlates of linguistic rhythm in the speech signal. *Cognition*, 1999
- [6] McCarthy and Prince, *Prosodic Morphology I*. ms., 1993
- [7] Fred Cummins, 2002, Speech rhythm and rhythmic taxonomy. In *proceedings of Speech Prosody 2002*.
- [8] Peter Roach, On the distinction between 'stress-timed' and 'syllable-timed' languages. , 1982
- [9] Petra Wagner, Rhythmic alternations in German read speech. In *proceedings of Prosody 2000*, 2000
- [10] Pike, K.N., *The Intonation of American English*. Ann Arbor, University of Michigan Press., 1945
- [11] Rulan Chao Pian, Tone and tone: applying musical elements to Chinese words, *Journal of Chinese Linguistics*, Vol.28 No.2(2000) , 2000
- [12] Selkirk, E.O.. *Phonology and Syntax: The relation between sound and structure*, The MIT Press, 1984
- [13] Shaver, P., Schwartz, J., Kirson, D., O'Conner, C., Emotion knowledge: Further exploration of a prototype approach. *Journal of Personality and Social Psychology* 52, 1987.1061-1086
- [14] 李爱军, 《普通话对话中韵律特性的声学表现》, 1999 年全国语音学会, 1999
- [15] 李重光, 《五线谱基本乐理视唱练耳基础教程》, 湖南文艺出版社, 2009 年
- [16] 刘献强, 《现代汉语节奏研究》, 北京语言大学出版社, 2007 年王洪君, 《试论汉语的节奏类型_松紧型》, *语言科学* 2004 年
- [17] 王洪君, 《汉语非线性音系学》, 北京大学出版社, 2008 年
- [18] 郑秋豫, AN ACOUSTIC PHONETIC STUDY ON TONES IN MANDARIN CHINESE, *Institute of history & philology academia sinica, Special Publications No.94*, 2005
- [19] 吴洁敏, 《什么是语言的节奏》, *语文建设* 1991 年第 5 期
- [20] 吴洁敏, 朱宏达, 《汉语节律学》, 语文出版社, 2001
- [21] 叶军, 《节奏说略》, 2001 年全国语音学会, 2001
- [22] 叶军, 《现代汉语节奏研究》, 上海世纪出版集团, 2008 年

An Annotation System of Speech Rhythm Based on Tempo Concepts

Abstract: This research tries to design a new annotation system of speech rhythm based on the music theory.

Pseudo-periodicity is the substitutive characteristics of speech rhythms. However, it is very difficult to be found in current prosodic researches. In order to research the isochronous rules of speech units, a new annotation system of speech rhythm based on the tempo concepts of music theory was founded, and this annotation system provides a new approach to characterize the speech rhythm.

Key words: rhythm; isochronous; tempo

(本文原载 NCMMS2013 中国贵阳 2013 年 8 月)