

## 汉语节奏研究综述

殷治纲

**摘要** 本文系统梳理了汉语节奏研究的相关内容，并就汉语的节奏类型问题进行了探讨。

文章首先对节奏概念进行了定义，认为语言节奏是指语言感知上的凸显性要素在时间序列上规律性出现的模式。该概念包含两个方面：第一个特征是由凸显性要素构成的对比性特征；第二个特征是上述凸显性要素在时间序列上的规律性，或叫节奏单元的组织模式。

文章正文部分以上述两方面研究为主线，并结合节奏计算模型和语言节奏类型等内容，对相关内容进行了总结，内容包括：

1. 第一类（对比性特征）研究中，介绍了基于轻重特征的研究、基于松紧（或停延）特征的研究、基于声调特征的研究（见文章第2部分）。(1) 基于轻重特征的研究是语言节奏研究的主流。在英语中，轻重特征是形成节奏的最重要对比性特征，但是汉语除轻声（和轻读）外，是否存在词重音仍存在争议。本部分重点介绍了以往音系学和语音学研究中有关汉语词重音问题的不同观点，同时也简略介绍了语句重音概念。另外，该部分还介绍了影响轻重感知的声学特征，其中最重要的特征是音高和时长，其次是音强特征。(2) 一些基于松紧（或停延）特征的研究关注到了汉语词重音不显著的问题，认为汉语不是像英语那样以重音为节奏支点，而是以松紧（或停延）特征为节奏支点。(3) 一些基于声调特征的研究认为声调是汉语的根本特征，可能是汉语节奏的支点。也有部分研究将声调有无和轻重特征联系起来，为汉语节奏研究提供了新的思路。

2. 在第二类特征（节奏在时间序列的规律性）研究中，梳理了节奏单元的线性模型和层级模型，并介绍了语速等和节奏有关的研究（见文章第3部分）。(1) 在线性模型中，主要介绍了音步、节拍和线性句块模型等内容。(2) 在层级模型中，很多研究基于 Selkirk 韵律层级理论提出了汉语的韵律节奏层级模型，一些典型层级有音节—音步/韵律词—韵律短语—语调短语—句子等。(3) 本部分还介绍了和节奏关系密切的一些研究，如语速和停延、节奏单元的音高时长特征等。

3. 文章第4部分介绍了语言节奏的计算模型，以及 PVI、 $\Delta C$ 、 $\Delta V$ 、%V、节奏共振峰等变量。这些工具为节奏量化研究提供了基础，并为区分言语和语言的节奏类型提供了模型方法。

4. 文章第5部分介绍了不同语言的节奏类型。以往研究把世界语言的节奏类型分为三类，分别是重音计时型 (stress-timed) 语言（如英语）、音节计时型 (syllable-timed) 语言（如汉语、法语）、莫拉计时型 (mora-timed) 语言（如日语）。本文对于节奏类型的一些基本问题进行了讨论，例如语言节奏类型是在什么韵律层级上形成的？区分节奏类型的主要特征是什么？文章认为，一种语言形成其节奏感的韵律层级应是某凸显性要素可形成显著规律性的最低层级；轻重特征仍是语言节奏中最重要的对比性特征。

5. 文章第6部分讨论了汉语的节奏类型问题。以往“汉语是音节计时型语言”和“汉语存在轻重对比”两种观点看似存在矛盾。前者暗示汉语每个音节都是重音，而后者认为汉语音节间存在轻重对立。针对“两难困境”，本文基于“最大公约数”原则提出了一种兼容方案，其核心观点认为汉语（普通话）节奏支点是以“声调有无”为基础的轻重对比特征，即汉语音节轻重不是和韵律位置相关，而是和声调有无相关——正常有调音节为重音，轻声（或轻读）音节为轻音。

**关键词** 节奏，对比性特征，组织模式，声调，重音

## A Review of Chinese Rhythm Research

YIN Zhigang

**Abstract** This article systematically summarized the studies of Chinese rhythm and discussed the rhythm type of Chinese.

Firstly, the article defined the concept of rhythm, which refers to the pattern of regular appearance of language salient elements in the time sequence. The concept contains two parts: the first part is the contrastive feature of the salient elements; the second part is the regularity of the salient elements on the time sequence, or the organizational pattern of the rhythmic units.

The main body of the article presented the studies of two aforementioned parts, as well as rhythmic computational models and language rhythm types.

1. In the introduction of studies related to the first part of rhythm concept (contrastive features), studies based on stress feature, studies based on pausing & delaying features, and studies based on tone feature were introduced (the part 2 of the paper). (1) Studies based on stress feature are the mainstream of language rhythm research. Stress is the most important contrastive feature that forms the rhythm of English. However, it is still controversial whether Chinese has word stress except neutral tone (and lightly pronounced) syllables. This part focused on the different views of Chinese word stress in phonological and phonetic studies, and also briefly introduced the concept of utterance stress. In addition, the acoustic features that influence the perception of stress were discussed, with the most important features being F0 and duration, followed by intensity. (2) Studies based on pausing & delaying features suggest that the core feature of Chinese rhythm is not stress but pausing & delaying feature. (3) Studies based on tone feature considered tone to be the fundamental feature of Chinese, and possibly the core feature of Chinese rhythm. Some studies connected “tonal/ non-tonal” feature with “light/ weight” features, which provided new idea for the study of Chinese rhythm.

2. In the introduction of studies related to the second part (the regularity of the salient elements in the time sequence), the linear models and hierarchical models of rhythmic units were introduced, and some rhythm researches such as speech rate were also presented (the part 3 of the paper). (1) In the section of linear models, foot, beat, and linear block model were introduced. (2) In the section of hierarchical models, some rhythmic hierarchical models of Chinese rhythm were introduced. These studies were mainly based on Selkirk's prosodic hierarchy theory, and some typical hierarchies were: syllable – foot/prosodic word – prosodic phrase – intonational phrase – utterance. (3) This part also introduced some studies that were closely related to language rhythm, such as speech rate, pausing & delaying, pitch and duration characteristics of rhythmic units.

3. The part 4 introduced computational models of linguistic rhythm, as well as variables such as PVI,  $\Delta C$ ,  $\Delta V$ , %V, and rhythm formant. These tools were useful for the quantitative studies of rhythm and the study of language (or speech) rhythm types.

4. The part 5 focused on the rhythmic types of different languages. Previous studies had classified the rhythmic types of world languages into three categories: stress-timed languages (e.g., English), syllable-timed languages (e.g., Chinese, French), and mora-timed languages (e.g., Japanese). This part discussed some basic questions about rhythmic types, such as at what rhythmic level are linguistic rhythmic types formed? What is the main feature that distinguishes rhythmic types? Our view is that the prosodic level at which a language forms its rhythmic sense should be the lowest level at which some salient element could form a significant regularity. Light & heavy feature is the most important contrasting features of language rhythm.

5. The part 6 discussed the rhythmic type of Chinese. In the past, two views of “Chinese is a syllable-timed language” and “Chinese has a light-weight dichotomy” seemed to be contradictory. The former implies that every syllable in Chinese is stressed, while the latter argues that there is a light-heavy opposition between Chinese syllables. In response to the “dilemma” of Chinese rhythm type, the paper proposed a compatible solution based on the principle of “maximum convention”, and its idea was that the core feature of Chinese rhythm is the stress feature based on “tonal/ non-tonal” feature. In other words, tonal syllables are all heavy, and neutral tone (and lightly pronounced) syllables are light. This shows that the Chinese word stress is not related to the position feature, but to category feature of “tonal/ non-tonal”.

**Keywords** rhythm, contrasting feature, organizational model, tone, stress

## 1. 绪论

“节奏”是人们在日常生活中经常使用的概念。按照《现代汉语词典》（第6版）定义，节奏有狭义和广义之分。狭义的节奏是“音乐或诗歌中交替出现的有规律的强弱、长短的现象”，而广义的节奏“比喻均匀的、有规律的进程”。

在语音韵律研究中，节奏问题也是一个重要内容。语言节奏的定义比较接近《现代汉语词典》（第6版）中的狭义节奏概念。《现代语言学词典》将节奏定义为“言语中可感知的单位凸显的规律性；此规律性可用重读对非重读，音节长短或音高高低等形式或组合来说明”。《语音学和音系学词典》（2000）认为节奏指“言语或诗歌中由凸显的要素有规律的间断出现所产生的知觉模式。这些要素可能是重音（如英语）、音节（如西班牙语）、重型音节（如古希腊语）或莫拉（如日语）”。

综合以上定义，我们认为语言节奏是指语言感知上的凸显性要素在时间序列上规律性出现的模式。节奏概念包含了两大特征：一是由凸显性要素构成的对比特征，如轻重、长短等对比。二是节奏单元在时间序列上（或组织模式上的）的规律性。

在第一类对比性特征中，最显著的是轻重特征。此外，很多研究认为长短、松紧等也可能是和节奏有关的对比性特征。在第二类时间序列规律性方面，微观层面的规律主要是语速（或时长），宏观层面的规律性则是节奏单元在时间序列上的组织模式，如线性模式和层级模式等。

在本文中，我们将对节奏的相关研究进行系统性梳理，并着重关注如下问题：

- (1) 和节奏特征、理论、方法有关的研究；
- (2) 汉语的节奏类型及节奏支点。

## 2. 和第一类节奏特征（对比性特征）有关的研究

### 2.1 基于轻重特征的节奏研究

轻重音问题是节奏研究的核心内容。有研究认为节奏的本质是轻重现象的有规律交替出现，所以每个节奏单元都由核心重音和其他非重音组成（叶军，2001、2008）。

和轻重有关的语音研究主要包括词重音研究和语句重音研究。由于句重音数量少、间隔远，通常每个句子只会有一个句重音，所以语句重音对形成语言节奏的作用比较有限。真正对语言节奏有重要影响的是词重音。

#### 2.1.1 词重音研究

词重音问题是汉语韵律研究的一个重要课题。一些研究中提到的音步重音也可以归为该类。对汉语普通话词重音研究主要集中于以下几个方面。

##### (1) 有无汉语词重音

该问题关注的是从音系层面看汉语有没有词重音。

英语是一种重音型语言，其至少包括三级词重音音位，即主重音（main stress）、次重音（secondary stress）、轻音（stresslessness），某些英语方言还可以分为四级重音（多了weak stress）（Hayes, 1995）。

汉语是否有词重音则存在争议。有的学者认为汉语有词重音，不过对词重音的定义也不尽相同（徐世荣，1956、1999；殷作炎，1982；陆致极，1984；端木三，1999、2007、2014；殷治纲，2021）。也有学者认为汉语除了轻声音节外，音系学层面的词重音并不存在或不明显（高名凯、石安石，1963；黄伯荣、廖序东，1991；刘现强，2007）。

总的来看，汉语轻声（或轻读）音节和有调音节存在轻重对立基本是学界共识。不过，正常有调音节之间的轻重差异在听感和声学上都不明显，是否区分轻重也多依赖于不同理论的处理方法。这也是

主要争议所在。

### (2) 轻重音的位置和级别

该问题主要包含两类研究，一类是音系学方面的研究——如果有音系学词重音，它的位置和级别如何？另一类是语音学方面的，主要关注词重音的语音学（或声学）表现。

在音系学层面，徐世荣（1956）认为北京话两音节词的重音格式有两个，分别是“重轻式”（前重）和“中重式”（后重）。还有的学者认为是“重中式”和“中重式”（殷作炎，1982；陆致极，1984）。端木三（1999）认为普通话不带轻音两字组重音模式是“左重”，而造成后重说的原因是把末音节的停顿前延长算成了音节时长所致。王志洁、冯胜利（2006）认为北京话词汇重音有三种类型，分别是左重、右重和不分（同重），但最后可以归纳为左重和右重。杨璐（2011）则认为词汇重音有四种类型，分别是重轻、重中、中重和不分（重重）。

在语音学层面，林茂灿等（1984）对普通话不带轻音的两字组考察发现其大多数是后重式，也有的前重，但并非必需。林焘、王理嘉（1992）发现，汉语普通话双音节往往后重；三音节词一般中间音节较轻，前后音节较重，而前后音节这种重的差别并不明显。另一个研究认为普通话两音节词的首音节比末音节时长更长，故其重音模式是前重（王晶、王理嘉，1993）。曹剑芬（2008）认为普通词内虽然存在“重”“中”的相对差异，但是并不构成区别性对立。词重音“重中”和“中重”的差异只是其随机变体，而不是两种词重音类型的差异。

由于英语轻重音和位置关系密切，基本按照重轻交替的音步（foot）模式组织节奏单元，所以以往汉语节奏研究也多受其影响而聚焦于轻重音与位置的关系，但最终研究结果却不尽人意，例如汉语是否存在前重、后重模式一直存在争议。通常来说，存在广泛争议的问题往往意味着该问题本身可能就没有显著规律性。具体到本问题，也许汉语轻重节奏并不和韵律位置存在因果关系，而是有不同于英语的自

身特点。近年来，一些研究从声调特征等视角研究汉语轻重音问题（殷治纲，2021），带来了新的思路。相关研究将在后面声调部分进行介绍。

### (3) 轻声和轻音

轻声（包括轻读）音节虽然在汉语中数量占比不大，但是它和有调音节的轻重对立是学界极少数的共识性结论（叶军，2008；周韧，2018；殷治纲，2021），所以轻声（包括轻读）在汉语节奏研究中的地位非常重要。

曹剑芬（2007）指出了轻音、轻读、轻声的联系与区别，认为轻音分结构轻音和语调轻音。前者主要指韵律边界处的轻读音节，如“的、了、吗”等；后者则是语流中临时轻读的音节。

殷治纲（2021）把声调和轻重音联系起来，认为轻声与其说是个声调概念，不如说是个轻重音（或弱重音）概念。轻声的本质不是一个具体声调，而是一类声调的集合，即无（固定）调集合。它和其他有调（阴平、阳平、上声、去声）音节构成音系层面的轻重对立。

叶军（2001、2008）也分析了轻声问题，认为：A. 轻声是语流中读的又轻又短，且失去本调或不带声调的音节；B. 普通话一些老资格口语词末字，以及一些语法化的单音节词或词缀，在语音上有 A 的特征；C. 短是轻声最重要特征；D. 弱是轻声最本质特征；E. 轻声音节不带声调；F. 轻声字失去本调 G. 轻声是中性调。他认为和轻声表达有关的 3 个因素是时长缩短、失去本调、响度减弱，从重要性上来说前两者更重要。

时长缩短和声调丢失都是形成轻声的重要特征，但是二者谁更重要存在不同的观点：一种观点认为，声调在听辨轻音时的作用没有时长重要（林焘，1985），但也有学者认为声调的作用比时长大（曹剑芬，1986；王韫佳，2004），或者声调的作用至少和时长同样大，因为语流中轻音不短反长的情况很多，这是因为在调型已经给予充分提示后，时长作用可以降低（叶军，2008）。

### 2.1.2 语句重音研究

语句重音是指比词重音更高等级的重音，包括和句法相关的句法常规重音（短语重音、句重音），与语义相关的语义重音，与语用相关的焦点重音等（王洪君，2008）。

语句重音和语言节奏的关系较小。很多学者提出，各种语言的语句重音规律基本上都是相同的（Duanmu, 1990、1999、2007；端木三, 2007、2014）。尤其对于句重音而言，由于出现频率低，间隔大，一般不影响节奏感知。不过，语句重音的很多理论对于了解轻重音的实现形式和规律有借鉴意义，在此也进行简要介绍。

赵元任（1922）的“橡皮筋”理论将字调比喻为橡皮筋，当它承载重音（常规重音、强调重音）时，调域会变宽，就如同把橡皮筋拉长一样。

沈炯（1985、1994、1999）高低音双线理论和调域括敛说指出，语流声调音域的高音线和低音线功能不同，高音线反映重音变化，而低音线反映韵律单元或节奏的顿挫。语句重音（包括常规重音和强调重音）处表现为高音线的突然抬高，同时由于低音线的变化不明显，故综合表现为重音处的音域加宽现象。

吴宗济（1993/2004）音阶移调（音高移块）理论认为，语句重音（常规重音、强调重音）的表现方法是重音所在单元处音阶的整体移动抬高所致。此处音阶是指听感音高（音乐上的半音阶/音阶，或者半音/全音为单位）。它与绝对音高（以赫兹为单位）之间的关系是对数关系，所以越到高频区域，在听感上抬高相同音阶时，其抬升的幅度越大。

有些研究（Duanmu, 1990；Cinque, 1993；端木三, 2007；周韧, 2007）将轻重音和句法、语义结合起来，提出了“辅重原则”“深重原则”“信息量原则”等理论。认为在句法常规重音方面，短语重音遵循“辅重原则”，即句法辅助成分重于中心成分；句重音遵循“深重原则”，指句法上内嵌最深的成分最重。“信息量原则”则是指信息量大的成分得到重音。

Xu 等（1999、2012）的焦点后压缩

理论（Post – focus compression, PFC）研究了焦点重音前后的声学表现，发现句中焦点成分会有基频升高、时长延长现象，而焦点前成分的音高和时长基本保持不变，焦点后成分会有音高下降、音域压缩等现象。

### 2.1.3 影响轻重感知的因素

轻重是一个主观感知量，但是会受到很多客观因素的影响。主要影响因素包括以下方面。

#### （1）时长和音高

目前，语音学界基本公认音高和时长是影响轻重音感知的最重要因素。

音高对重音感知影响很大。很多研究提到重音所在音节的表现是音高上限抬高（上声为压低），音域加宽，重音后音节调域压缩等现象（赵元任, 1979；沈炯, 1985、1994；曹剑芬, 2002；Xu, 1999、2012；王蓓, 2002）。也有很多学者指出时长是重音感知的重要因素（林焘、王理嘉, 1992；林茂灿、颜景助、孙国华, 1984；曹剑芬, 1990；叶军, 2008）。

对于音高和时长对重音感知谁更重要的问题，学界存在不同观点。有的研究认为时长比音高重要（仲晓波等, 2001；仲晓波等, 2002），但也有很多研究认为音高更重要一些。曹文（2010）研究了焦点重音的表现形式，发现韵律信息中最重要的是高音点落差级别。蔡莲红等（2001）提出了一个拟合重音听感的线性公式  $L = 1.5F$ （音高）+ 0.95（时长）+ 0.65R（调域），从系数上暗示音高的重要性更高。还有研究发现双音节韵律词重音差与高音点的差值相关性强于时长差值。（王韫佳等, 2003）。也有研究发现在表现重音时，音高和时长之间呈现以音高为主的互补关系或交换关系（许洁萍等, 2000；仲晓波，杨玉芳, 1999；叶军, 2008）。

#### （2）音强

音强对于轻重音感知的影响力相对弱于音高和时长。早期一些语音研究曾认为音强是重音的主要因素（罗常培、王均, 1981），但后来研究认为音强的作用没有音高和时长那么大（林焘、王理嘉, 1992；林茂灿、颜景助、孙国华, 1984；

叶军, 2008)。

### (3) 其他影响因素

除了音高、时长、音强等因素外, 还有一些其他因素也被认为和轻重感知有关系。

声调特征也被认为是影响轻重感知的另外一个因素。除了轻声音节和轻音的密切关系外, 某些声调被认为会更容易成为重音。贾媛等 (2008) 认为汉语焦点对于音高的作用主要体现在 H 特征 (高音) 上。具有 H 特征多的声调容易承载重音, 所以像阴平调更容易被重读。殷治纲 (2011) 认为除了 H 特征外, “高—低”对比特征 ( $H - L$ ) 也容易形成重音感知, 所以普通话中阴平和去声比其他声调更容易被感知为重音。总的看, 声调特征对轻重感知的影响往往也和音高 (或时长) 等声学特征有关。

音质性质可能对于轻重感知也有一定影响。响度顺序原则 (SSP, sonority sequencing principle) 认为其他条件相同时, 音节的响度顺序原则 (sonority sequencing principle) 从大而小依次是: 元音 (低元音 > 中元音 > 高元音), 半元音, 流音, 鼻音, 擦音 (浊 > 清), 塞音 (浊 > 清) (王洪君, 2008)。但学界一般只把音质看作对响度有制约, 其重要性还不足以将其作为重音感知的变量 (林焘、理嘉, 1992)。

## 2.2 基于松紧 (或停延) 特征的节奏研究

鉴于汉语词汇层面的轻重特征存在争议性, 很多学者尝试从其他方面寻找汉语节奏的对比性特征。有的研究认为松紧 (或停延) 对比是汉语节奏的支点。

王洪君等 (2004) 对汉语节奏类型、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, 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。

刘现强 (2007) 认为目前语言节奏可以分为以重音为支点的节奏 (如英语) 和以停延为支点的节奏 (如汉语) 等几类。

英语以重音为节奏支点, 以两个相邻重读音节之间的音段, 或重读音节及其后面的非重读音节为最小节奏单位 (音步) (转引 Halliday, 1985; liberman, 1977)。汉语则是以停延为支点构建自己的节奏系统。汉语虽然也有重音, 但还达不到节奏“支点”的重要程度。

上述研究意识到了汉语缺乏显著的轻重特征, 因此把节奏问题聚焦于节奏单元的组块模式上, 也就是节奏单元的松紧关系或者停延现象上。但是, 以上理论也存在一些值得商榷之处。一是松紧或停延特征实际只涉及绪论中所说节奏概念两大特征中的第二个特征 (时间序列上的节奏组块模式), 而没有涉及第一个特征 (凸显性要素), 因此是不全面的。二是明显的停延现象多发生在韵律词以上的高层单位, 而语言节奏类型 (如重音计时型、音节计时型等) 的形成一般是发生在韵律词以下的低层单位, 所以停延特征对于确定语言节奏类型作用有限。三是节奏单元的松紧关系或者停延现象, 也是世界语言都有的特征, 很难将其作为区别汉语节奏和英语节奏的本质特征。因此, 要补充上述理论的不足, 仍需要寻找汉语节奏的凸显性要素 (或对比性特征)。

## 2.3 基于声调特征的节奏研究

也有很多研究认为, 声调是汉语节奏类型的“支点”。

Hyman (2007) 认为语言可以分为重音语言 (stress language) 和声调语言 (tone language) 两大类。英语属于重音语言的代表, 而汉语属于声调语言的代表。重音语言带有词层节律结构, 需要重音这种结构属性 (structural property) 来标识音节在节律层级上的相对轻重关系。声调语言带有词层音高特征, 需要声调这种特征属性 (featural property) 来标示音节词汇单元。因此, 作为汉语根本属性的声调特征很可能与其节奏有密切关系。

许希明 (2013) 认为, 从语言类型学的视角看, 重音是英语的节奏支点, 而声调则是汉语的节奏支点。其研究分别研究了汉语声调和重音 (或重征) 的关系、声

调支点节奏与重音支点节奏的区别、单音节性与多音节性差异等问题，指出汉语主要靠声调来调节，变调是汉语节奏的需要；英语主要靠重音来整合，调控重音分布则是英语节奏的需要。但是该文没有明确说明声调形成的是哪种性质的对比特征，例如是轻重对比、调类对比还是综合对比。这使得研究结论具有一定笼统性和含糊性。

还有一些研究把声调和轻重音联系起来。Yip (1980: 57) 指出，“对声调语言来说，是声调决定重音，而不是重音决定声调。”

殷治纲 (2021) 认为声调系统和重音系统存在深层联系，汉语词重音关联的是声调而不是韵律位置。他从“质”和“量”的角度提出了判断词汇层面轻重音的两条音系学标准：(1) 具有区别意义的作用；(2) 具有显著听觉区分性。依此标准，其研究发现汉语韵律词所谓“左重”(或前重) 趋势只是一种语音学统计规律，不构成音系规律，而声调类别(有调和无调) 才构成轻重音的音系学规则，具体规则为：(1) 有调音节(包括阴平、阳平、上声、去声) 构成重音；(2) 无调音节(轻声或轻读) 构成轻音。根据该规则，汉语普通话两音节韵律词可以分为两类：(1) 重重式(音节均为有调音节) (2) 重轻式(词尾音节为轻声或轻读)。该分类结论说明在音系学层面，汉语只存在“重、轻”二元对立，不存在“重、中、轻”多元对立。多元对立中的“中”在音系学层面可以归入“重”的类别。该研究虽然是基于声调特征的研究，但因把声调和轻重音联系起来，所以仍然认为汉语节奏是和轻重特征有关的。

叶军 (2008) 从音节和声调的长短方面(莫拉数量) 研究了轻重问题，认为短音节(一个莫拉) 对应轻音，长音节(两个莫拉) 对应重音。汉语音节均为开音节(元音尾或鼻韵尾)，其右边界开放特点使汉语音节占有两个时位(莫拉) 成为可能，也使得普通话不会因为音节结构形成长短对立。汉语有些方言中韵尾为塞音的入声音节为短音节(一个莫拉)，其他舒

声音节为长音节(两个莫拉)，存在长短(和轻重) 对立。普通话目前没有入声短音节，所有带调音节都是舒声长音节，所以有调音节都是重音。普通话目前只有轻声音节是短音节(一个莫拉)，属于轻音。这一研究虽然是从声调长短视角切入，但是结论和上面研究(殷治纲, 2021) 的结论是一致的，并提出了汉语轻音除了包括轻声外，还可能包括入声调(普通话已消失，某些方言中还有)。

以上研究关注到声调特征对于汉语节奏的重要意义，并把以往认为是两大不同系统的声调系统和重音系统作了关联，对于人类语言研究具有类型学意义。

### 3. 和第二类节奏特征(节奏单元组织模式)有关的研究

如本文绪论所说，节奏概念的两大特征除了有对比性特征外，还有凸显性要素在时间序列上的规律性特征，即节奏单元的组织模式问题。

王洪君 (2008) 认为节奏单元有如下几种组块模式：(1) 层次模型：认为节奏组织的形成模式是像法树一样按照套合层次纵向组织的，每增加一个新成分就要增加一层。(2) 层次—连续复合模型：认为节奏可以表现为有限层次上、同一层面线性组块的组合。句子长度和复杂度的增加只能导致各层级上并列的节奏单元数目的增加，而非深度的增加，即节律层级倾向于横向生长。(3) 线性句块模型：需要将句块内成分在保持层级结构前提下将其投射到单线上来。然后在线性层面上处理黏合、组合、等立等三种关系。等立结构是指成分之间彼此独立且地位平等，呈现并列关系。黏合结构是指成分之间关系紧密，有很强的内聚性，构成结构紧密的一个复合单元。组合结构是指成分间关系较松散，按照一定规律组合成更大单元。

我们借鉴了这种分类方法，并从线性模型和层级模型两个方面来介绍节奏单元的组织模式。线性模型关注的是线性时间序列上的微观节奏单元，而层级模型关注的是不同层级、规模的节奏单元如何构成

宏观的节奏体系。

### 3.1 节奏的线性模型

线性模型主要关注节奏单元（例如音步、节拍、停延段等）在线性序列上的组织模式。

#### 3.1.1 音步

音步是英语节奏研究中的关键概念。一个音步可以定义为一次轻重交替。由此概念可见音步和重音的关系非常密切。重音就是音步中的重拍（Hayes, 1980/1995）

对于汉语是否存在音步单位，以及汉语音步和英语音步是否一样，目前存在不同意见。有的学者认为汉语有音步单位（冯胜利，1996；端木三，2000；Chen, 2000；王洪君，2004），也有很多学者觉得汉语没有明显的音步单位，可以用韵律词等单位代替。

在认为汉语有音步的理论中，对于汉语音步是否和英语音步一样有重音也存在不同观点，有的研究认为汉语音步不一定有重音（Chen, 2000；冯胜利，1996；王洪君，2004），有的则认为汉语音步和英语音步一样都有重音（端木三，2000）。

关于组成音步的基本单位也存在不同观点，通常认为是莫拉或音节，也有的认为是拍。Hayes (1995) 认为一种语言只能以莫拉或音节中的一种为基本单位。Duanmu (2007) 则认为一种语言可以同时有莫拉拍和音步拍。

关于音步长短，一般两拍步是最常见的。有的观点认为也有三拍步（如诗歌），但很少见其他长度（端木三，2014）。

此外，很多研究中也有超音步和蜕化音步的概念。（1）一般认为超音步是三音节构成的音步，汉语中主要是三音节词的情况（Selkirk, 1980a；Shih, 1986；冯胜利，1996、1998；Chen, 2000 等）。蜕化音步则是单音节构成的音步，汉语一般是单音节词的情况（冯胜利，1997）。但是也有学者认为超音步和蜕化音步概念都是多余的，只用基本音步（双音节或两拍）概念就够了——超音步可以看作基本音步和其他音节复合而成，多出的音节可以不

入音步（端木三，2018），而单音节词中的落单音节可以后接空拍构成双拍音步（Duanmu, 2000）。

总的来看，汉语音步概念并不如英语音步那么明显，相关理论也存在很多争议。

#### 3.1.2 节拍

有一些研究借用音乐节拍概念来描述语言的节奏。

安英姬（2002）利用音乐的节拍理论分析汉语的节奏，认为汉语中有不同的句间停歇：“停顿”与“间歇”。“间歇”之间的“节”除了具有单位的共性外，也具有节奏的周期性和普遍性，是汉语的节奏单位。节奏单位既不是语法单位，也不是语义单位。

殷治纲（2013）参考乐理知识，对语言的节奏时长描述体系进行了研究。认为节奏的典型特征是准周期性，但是以音节等语言单元时长为基础的韵律计时单位缺乏客观性和准确性，也不能反映节奏单元的准周期性特点，而独立于语言单元的“拍”和“节”的计时单位更有利于语言节奏计时问题研究。该研究以节、拍等音乐计时单位为基础，建立了语言节奏新的时长描述体系。

上述基于节拍概念的语言节奏研究把语言和音乐联系起来，使人们在更高、更宏观的层面看待节奏问题。同时，节拍计时单位也比音节等单位更具有客观性、精确性和跨领域适用性。不过，由于节拍概念更偏重音乐概念，而不具有语言学含义，所以在传统语音研究中被较少使用。

#### 3.1.3 线性句块模型

王洪君（2008）基于韵律单元间松紧关系，根据线性句块模型，认为音步、停延段按照“二常规、一三可容、四受限”规则组成韵律结构（单位是音节）。她利用松紧节奏探讨了韵律结构与语法结构的对应关系，提出了满足韵律长度条件下（韵律词遵循“二常规、三可容、一四受限”的音节约束条件；韵律短语遵循“二常规、一三可容、四受限”的韵律词约束条件）的韵律结构预测方法：首先把语法结构与韵律结构分“一致型”和“不一致

型”，然后进行：（1）韵律结构预分，确定文本中“一致型”语法结构如何构成韵律结构；（2）韵律结构调整及确定，在第一阶段所得到的半成品韵律结构基础上，针对“不一致型”，根据相关规则，调整并确定语句最终的韵律结构。其预测模型又包括韵律词预测和韵律短语预测2个部分。

### 3.2 节奏的层级系统

节奏层级模型比线性模型更适合描述语句及以上的宏观结构。

Selkirk (1980a、1980b) 提出了韵律层级理论，认为韵律单元遵循层级结构，每个韵律层级单位由其下一层的单位组成。韵律单位从小到大可以分为音节 (syllable) — 音步 (stress foot) — 韵律词 (prosodic word) — 韵律短语 (phonological phrase) — 语调短语 (intonational phrase) — 话语 (utterance)。

Halliday (1985) 提出了单位层级的划分标准：（1）要保证划分单位的同质性，如都按语音根据，或语法依据。（2）要保证每一级单位有自己不同于上、下级其他单位的特性。（3）每级单位要有自己特有的组合模式。他在其音系理论中建立了四个层面的音系单位，从高到低分别是：调群 (tone group) — 音步 (foot) — 音节 (syllable) — 音位 (phoneme)。

目前，汉语韵律节奏研究者也大都采用了层级模型，其中较有代表性的系统和研究如下。

李爱军 (1998、1999) 对汉语韵律节奏的层级模型和音高、时长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：音步、韵律词、次要韵律短语、主要韵律短语和语调组。（2）韵律词后感知到的间断 (break) 很小，声学上一般没有无声停顿出现。词内音节之间没有停顿，具有一个词重音 (韵律上的凸显成分)。（3）次要韵律短语 (minor phrase) 是由一个或多个韵律词构成的，具有一个短语重音。其后感知到的间断比韵律词大，但比主要韵律短语小，可以是较短的无声停顿 (silent

pause) 或是有声停顿 (filled pause) (Lin, 1998)。（4）主要韵律短语 (major phrase) 由一个或多个次要韵律短语构成，具有一个短语重音。其后可以感知到明显的停顿，有明显的无声段，与后接短语之间有明显的音高重置 (pitch resetting)，并且短语的音高曲线有明显的递降 (declination)。（5）语调组指由一个或多个主要韵律短语构成且有一个语句重音的语句。随着韵律单位的增大，其后接的停顿也越长。在以上研究基础上，中国社会科学院语言研究所还提出了 C - ToBI (Chinese - Tones and Break Indices) 汉语韵律标注系统 (李爱军, 2021)。

王洪君 (2008) 提出了纯韵律层级单位：语调段一大停延段一小停延段—词重音段—音步—音节—莫拉。另外，还提出了具体汉语普通话的韵律层级单位：语调段一大停延段一小停延段—音步—音节—莫拉。她结合郑秋豫有关研究结论 (Tseng, 2005)，详细说明了从低到高几个韵律层级单位的概念：（1）音步：一到三个正常音节，还可以加一到两个轻声音节。两音节为常规形式。（2）停延段：一到三个正常音步，两音步为常规形式。此外还有“复二步”说法，一般认为它是介于音步和停延段之间的单位。（3）语调

① 郑秋豫 (2008、2010)、Tseng (2005) 提出了阶层式语流韵律架构模型 (HPG, Hierarchical Prosodic Phrase Grouping)，并在其基础上提出了另外一种汉语韵律标注系统 M - ToBI (Mandarin - Tones and Break Indices)，其中将韵律层次划分为 6 级：退化的音节、正常的音节、较小的短语停顿、较大的短语停顿、呼吸群、韵律组 (Peng et al., 2005)。HPG 模型的主要特征是：（1）由上而下的管辖制约。（2）多种制约机制：发声器官；生理机制 (含呼吸)；认知限制 (语言规划的制约)。此外，该模型还注重工程应用，其声学模型包括 4 个组成部分：基型 1 – 多短语韵律句群之基频模板；基型 2 – 多短语韵律句群之音节时长模板；基型 3 – 多短语的音强分布组型；基型 4 – 韵律句群之阶层性边界 / 停延组型。该模型不同于其他韵律层级模型的地方是出现了一级叫做“呼吸群”单位。该级单位是以呼吸生理现象命名而非通常的语言现象，其节奏单元规模应该大致对应于语调短语或小句规模。

段：由一个或多个停延段组成。语调段特点是最后一个停延段一定载有语调，其出现位置取决于句法的句重音（必有）和取决于语用的强调重音、对比重音。在以上单位中，音步和停延段内部有组成单元的数量要求。它们倾向于等长结构或等距离出现，是构成节奏的主要单位，叫做“节奏单元”。语调段没有等距离出现的特点，故不是构成节奏的主要单位。它主要表现语调、重音等韵律信息，是韵律单元。

曹剑芬（1999、2003）认为语言节奏本质上是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。汉语普通话的节奏可以大致分为三个基本层次：韵律词、韵律短语和语调短语。（1）各层次节奏单元有其内聚特征：韵律词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；韵律短语内部，主要表现为韵律词间连而非断的连接方式和总体音阶的逐渐下倾倾向；语调短语内部的凝聚性主要表现为总体音阶的不断下倾倾向，以及内部各个韵律短语末尾音节韵母的显著延长（一种特殊的表示待续关系的连接方式）。（2）各节奏单元的边界特征也不相同：韵律词边界多表现为边界前的韵母延长，且该效应还多半会被消抹掉；韵律短语边界，多表现为边界前延长和较短的无声间歇；语调短语边界表现为较长的无声间歇，边界前一般无延长现象（多会缩短）。另外，该研究认为，虽然语言节奏中不存在等长结构，但是节奏单元的长度变化范围大致有定，并且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。

吴洁敏等（1991、2001）提出言语停延分为句内停延和句外停延。其中，内3层停延包括“音节—音步—气群—句子”，而外3层停延包括“句子—句群—段落—篇章”，并且认为节奏特征不仅指音强和音长两个方面，还包括徐疾、高低、长短、轻重、音质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。该研究提及的对比特征涉及多种对立因素的组合；其层级概念也从句子以内扩展到了篇

章等大语言单元；其面向领域也从语言学扩展到了语言教学、艺术朗诵等领域，属于广义节奏研究。

殷治纲（2011）沿用了主流的“音节—韵律词—韵律短语—语调短语—语句”层级结构模式，系统研究了各韵律层级内的时长规律、音高规律和重音规律等内容。前人基于层级模型做的类似研究虽然不少，但往往无法“彻底”贯彻层级观，主要表现为在研究高级别韵律层级（如短语、语句）时，仍然以最底层的音节为研究对象。这实际上不符合层级模型中“每一层级内的研究对象都是其直接下辖单元”的要求。为了改变上述问题，该研究的一大特点是较彻底地贯彻了层级思想，在研究每一层级时都以该层直接下辖单元为对象，例如在研究“韵律短语”层特点时，会以其直辖的“韵律词”（而不是底层的音节）为研究对象，这样就可观测到大单元在宏观层面的变化规律。

### 3.3 与汉语节奏有关的一些语音研究

和节奏有关的语音研究很多，除了前面已经提及的轻重音、声调、节奏单元等内容外，比较重要的还有关于语速（时长）的研究。因为狭义的节奏特征主要体现在语速（或）时长规律上。

#### 3.3.1 语速（时长）研究

语速是语言韵律、节奏研究的重要内容之一。它一般用单位时间内的音节个数来衡量（曹剑芬，2003）。计算公式为：

$$\text{rate} = \text{syllable number} / \text{duration} \quad (1)$$

公式（1）中，rate 表示一段话的平均语速，syllable number 表示该段话所含音节个数，duration 表示话语时长。

语速和时长关系紧密，存在倒数关系。相对于时长研究，以“语速”为观察对象的优点是可以更方便地说明发音运动快慢规律，还可以把“延长”和“停顿”现象统一纳入语速变化的范畴—音段“延长”可视为发音降速，而“停顿”就是语速降为“零”。

语速概念有狭义和广义之分。狭义的“语速”（speaking rate, speed of speech）一般和“发音速率”（articulation rate）相

对应，指一段话的整体速度（R. L. 特拉斯克，2000；戴维·克里斯特尔，2000；Koreman, 2006）。计算狭义语速时，话语时长（公式 1 中 duration）要包括无声停顿长度。与其相关的“发音速率”概念则是指“可听语音的表达速度，包括填声停顿，但不包括无声停顿”（R. L. 特拉斯克，2000；戴维·克里斯特尔，2000）。广义的“语速”则不严格区分二者区别，如曹剑芬（2003）认为语速（广义）“通常是指发音速度，也可以指听觉上对话语速度的感知印象”。

和语速相关的研究主要包括以下方面：

- (1) 语速在不同韵律单元中的变化规律（曹剑芬，2003；李爱军等 2007；Li & Zu, 2008；Tseng et al., 2004、2005；殷治纲, 2019）；
- (2) 语速变化对不同语音单元（辅音、元音）伸缩效果的影响（Li & Zu, 2008；梅晓, 2010）；
- (3) 语速对韵律单元组块的影响（李爱军等, 2007）；
- (4) 语速对音高的影响（曹剑芬, 2003；Li & Zu, 2008）；
- (5) 语速对重音等韵律特征的影响（张家騤等, 2002）。

### 3.3.2 语调研究

音高与节奏单元的关系相对较小，但是节奏单元的音高变化会涉及语调特征。

汉语语调问题的一些重要理论包括赵元任（1968/1979）的大波浪与小波浪理论。该模型将语调比喻为大波浪，字调比喻为小波浪，认为语调和字调的关系是小波浪叠加在大波浪上，具体音高值表现为二者的“代数和”。沈炯（1985）在借鉴英国传统语调理论基础上提出了语调结构模型，语调结构内部按照出现次序包括调冠（可无）、调头、调核、调尾等四部分。林茂灿（2004、2005）的边界调理论认为语气一般由句末边界处的声调段来承担。

### 3.3.3 和节奏有关的语音研究小结

总的来说，目前对节奏研究主要体现在轻重、快慢和节奏单元大小等方面。

- (1) 重音处的发音增强现象在音高和

时长都有体现。音高方面：会强化声调本身的音高特性，使高者愈高（如阴平），低者愈低（如上声）；使负载它的韵律单元音域扩大；使其后音域变窄。时长方面：时长一般延长。时长和音高的相关性存在争议。

(2) 韵律节奏单元呈现层级关系，大致分为音节——音步/韵律词——韵律短语/停延段——语调短语——句子/气群/语调群等。

(3) 韵律边界停延和韵律单元大小有关。节奏边界体现多层次性，可以通过多种手段实现。

(4) 韵律单元内基频一般呈下倾趋势，其基频（F0）和所含音节数目的变化曲线是类似于一种指数函数的下降曲线形式。

(5) 韵律单元边界后（尤其是较大的）的新单元开头存在音高重置现象。

## 4. 语言节奏的计算模型

节奏问题一般认为是个主观感受量，但是为了进行定性和定量的分析，必须找到合理的声学物理量对其进行测量。此外，语言学家一直希望能够对世界不同语言的节奏类型进行量化分析和分类，这也需要统一的计算模型。本部分主要介绍节奏量化研究中采用的一些模型和方法。

早期语言学家把世界语言的节奏类型归纳为重音计时型（stress – timed）和音节计时型（syllable – timed）语言（Pike, 1945），之后又增加了莫拉计时型（mora – timed）语言（Ladefoged, 1975）

有的研究猜测不同节奏型语言中的相应特征存在周期性或准周期性规律，如Abercrombie (1967) 假设：(1) 音节计时型语言中的音节时长趋向于相同，其平均音节时长偏差要小于重音计时型语言；(2) 重音计时型语言的重音间隔时长趋向于相同，其重音间隔时长偏差要小于音节计时型语言。

为验证以上假设，Roach (1982) 提出了定量分析节奏的方法，其方法关注全局节奏单元时长的偏差，以及局部单元的

时长比例和辅、元音时长比例。具体做法是用语调单元的时长除以音节个数，得到“理想”音步的平均时长，并且用该均值对单元时长的差异进行规整。但该方法实际测得的是时长特征，其研究也发现上述等时性假设并不成立。

Ramus et al. (1999) 提出了一种计算节奏的方法，他用以下变量对时长类别进行了分类。(1)% V—元音占整个句长的比值。该数值可以反映 CV、CVC 等不同音节结构复杂程度。音节越复杂则% V 值通常越小；(2)  $\Delta C$ —辅音段长度的偏差，可以反映辅音时长的一致性（数值越小意味辅音等时性越好）；(3)  $\Delta V$ —元音段长度的偏差，可以反映元音时长的一致性。该研究以% V 和  $\Delta C$  为横坐标和纵坐标建立直角坐标系，可以较好地把三类节奏型语言区分开，而  $\Delta V$  的作用不明显。

Cummins (2002) 研究了影响节奏类型的因素，并进行了严格受控、激发状态以及自由状态的各种试验；建立了一个层级结构的二元节奏模型，以两音步为基础，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。

Low et al. (2000) 设计了新的指数 PVI (Pairwise Variability Index, 成对差异指数) 用来计算相邻音段（元音或者辅音段）平均时长差异性。进一步对语速进行规整后，还可以得到语速归一后的成对差异指数 (nPVI)，其计算公式是：

$$nPVI = \frac{100}{m-1} \left| \frac{2 * (d_k - d_{k+1})}{d_k + d_{k+1}} \right| / (m-1) \quad (2)$$

该公式中的 m 表示音段数量， $d_k$  为第 k 个音段的时长。nPVI 可以得到的最小数值为 0 (单元时长相等)，最大数值为 200。研究发现 nPVI – V (语速归一后的元音段成对差异指数) 可以有效区分存在不同元音缩减特性的语言，如新加坡英语和英式英语。

Grabe & Low (2002) 在前人研究基础上，使用 5 个参数：nPVI – V、rPVI – C (未经语速归一的辅音段成对差异指数)、 $\Delta C$ 、% V、 $\Delta V$  对 18 种语言进行了分析，发现 nPVI – V 和% V 组合能较准确区分重

音节奏型和音节节奏型语言。

Fuchs (2014a、2014b) 认为以往主要基于时长信息的研究存在不足，把响度、基频等更多特征纳入 PVI 方法，基于 PVI 设计了多个新指标，拓展了研究视野。

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，具体包括两步：(1) 由标注数据的时长得到时间树；(2) 分析时间树和句法树的相似度。研究发现二者具有相关性。

Gibbon (2020) 基于节奏共振峰理论 (Rhythm Formant Theory)，利用节奏共振峰分析方法 (Rhythm Formant analysis, RFA) 来研究语言的节奏特征。该方法把语言节奏看作一种低频波动现象，通过分析言语数据的（波形）振幅调制 (amplitude modulation, AM) 和频率调制 (frequency modulation, FM) 等低频 (low frequency, LF) 变量，可以得到节奏频谱图 (rhythm spectrogram) 和节奏共振峰 (rhythm formant, RF)。该方法的特点是不需要预先标注语料，研究语料也不限于单词等受限小单元，可以是文本级别的大单元。该模型不但可以区分英语和汉语的节奏类型差异，还可以区分男性和女性言语风格的节奏差异。

以上计算方法和模型为语言节奏的量化研究和语言对比分析提供了基础。计算模型的丰富度和复杂度也在不断提高，从最初只研究时长特征，扩展到多元特征，从最初的线性时长特征分析，发展到三维频谱特征研究。不过，现有特征和方法仍远未到完善程度，未来节奏测量应进一步增加测量的语言数量，多维度增加有效测量指标，并实现自动信号提取和测量 (杨茂霞、陈美华，2021)，努力提高研究的广度和深度。

## 5. 语言的节奏类型分析

### 5.1 节奏感觉和语言节奏分类

不管是音乐还是语言，不同的节奏类型都会带来不同的整体感受。

在音乐中，不同节奏类型会有不同的风格特色。例如：(1) 四二拍（以四分音符为一拍，每小节含有两拍。以下拍号含义类推）节奏模式为“强弱”交替，具有一种稳定的行进感。(2) 四四拍节奏模式为“强、弱、次强、弱”。它实际是两个四二拍的复合，但是复合后把第二个强拍弱化成“次强”了，从而使音乐更富有抒情性。(3) 四三拍节奏模式为“强弱弱”，其特点是有动荡、摇摆的感觉，该特征在华尔兹乐曲里很明显。

对于不同语言和方言而言，不同的节奏特点也会带来不同的听觉感受。例如英语轻重长短明显，如敲摩尔斯电码；法语节奏分明，如机关炮。很多学者分别用重音计时型(stress-timed)和音节计时型(syllable-timed)来描述上述两类语言(Pike, 1945; Abercrombie, 1967)。早期研究认为，音节计时型语言(如法语等)的音节时长具有等时性特征，而重音计时型语言(如英语、俄语等)的重读音节(stressed syllable)的间隔具有等时性特征。后来，Ladefoged (1975) 认为除了上述两种节奏类型外，还有莫拉计时型(mora-timed)语言，其节奏单元是比音节更小的莫拉，代表语言如日语等。日语节奏比法语等更紧密快速，节奏可比喻为机关枪。我们把以上特点称作语言的节奏类型。它因为具有语言类型学意义，可以算世界语言节奏研究中的一个核心问题。

除了语言的节奏类型外，日常话语也会因不同的节奏特征组合(如“轻重”、“速度”)而形成不同的话语风格：强烈的轻重对比，加上较快的语速，往往表示激昂、果决的风格，而较弱的声音加上较慢的语速，则给人温和、柔弱的感觉。这属于更加宽泛的话语节奏风格问题，可以和很多语用学、社会语言学研究相结合，目前研究还较少。

## 5.2 节奏类型是在什么韵律层级上形成的

要分析语言的节奏类型，首先需要明确语言节奏感觉是在什么节奏层级产生的。

Selkirk (1980b) 最先提出了六个韵律层级单位，McCarthy 和 Prince (1986) 又增加了莫拉，这样就形成了七级单位：莫拉(mora) — 音节(syllable) — 音步(stress foot) — 韵律词(prosodic word) — 韵律短语(phonological phrase) — 语调短语(intonational phrase) — 话语(utterance)。

很多研究都以韵律词单位为界，认为其前(下)、其后(上)两部分的语言学取向有所不同(王洪君, 2008; 端木三, 2014)，而节奏特征主要体现在韵律词以下单位。我们同意该观点，认为语言节奏感是个基础性语音/音系特征模式，主要和韵律词及其以下单位(音节、音步、韵律词)特征有关。韵律词以上层面(短语、语句)虽然也有节奏单元和相关规律，但主要面向语义、语用等高层信息，和语言节奏类型关系不大。

对于韵律词以下单位，因为不同语言所含节奏层级也不尽相同(例如音步层)，也很难统一说节奏感形成于哪层单位。但总的来说，一种语言形成节奏感的层级应是某凸显性要素可以形成显著规律性的最低层级。

## 5.3 区分语言节奏类型的主要特征是什么

按照本文绪论中的定义，语言节奏是指语言感知上的凸显性要素在时间序列上规律性出现的模式。所以，要区分不同语言的节奏类型，首先要确定作为区分标准的凸显性要素是什么。

从前面的介绍可知，目前学界具有广泛认可度的凸显性对比特征是轻重特征。在实践中，由于轻重音不好计算，节奏型计算使用较广泛的特征是时长特征(如PVI、%V、ΔC、ΔV)。另外，因为时长也是影响轻重感知的关键因素之一，所以现有节奏计算模型仍然可以认为计算的是轻重特征。需要说明的是，影响轻重感知的另外一个关键特征——音高特征，因其受语调、声调、发音人等很多因素影响，很难进行计算比较，所以在实际节奏计算模型中很少被使用。从以上方面综合来看，

目前区分语言节奏类型的主要特征实质上还是轻重特征。

对于重音而言，因为存在音步重音、词重音，甚至短语重音和句重音等不同级别，那么哪级重音是计算节奏的关键特征呢？首先，很多学者都提到要区分重音的性质，如节奏重音和语义重音。（王韫佳等，2003；冉启斌等，2013）很显然，在此需要考虑的是节奏重音而非语义重音，而节奏重音是指可以通过音系规则获得的、能形成节律轻重循环规律的重音。

具体联系到前面分析的产生节奏感觉的韵律层级问题，我们认为形成一种语言节奏感的轻重特征应该是在韵律词以下的节奏重音，并且是可形成轻重规律的最低级别的重音。

下面，我们以英语和汉语为例，分析一下各自的节奏特点和类型。

#### 5.4 以英语为代表的重音计时节奏型

英语是一种重音计时型语言。它以重音为节奏“支点”，其节奏单元包括语流中前一个重音（含）到后一个重音之前的所有音节。英语底层的音步重音和词重音都具有规则性，但能够作为节奏支点作用的应该是最小级别的重音，即音步重音。更高一级的词重音的作用则相对次之。

#### 5.5 以汉语为代表的音节计时型节奏

在以往研究中，汉语被归入音节计时型语言，即汉语的凸显性要素是音节，在音节层面上即已经形成规律性。由于汉语音节不存在长短对立，所以我们认为其凸显性还是体现在轻重上，即相对于英语而言，汉语的每个音节几乎都是重音。

这一结论存在几个值得关注之处：

(1) 从跨语言对比视角，汉语（有调）音节相对于英语都是重音。和英语单词中音节间存在的“强弱”交替感相比，汉语词汇中每个正常带调音节的确都更清晰饱满，可谓是字字重音。这种跨语言的音节轻重差异性在语音声学分析和主观听感上都是成立的。

(2) 汉语系统内部的重音问题存在争议。一些研究认为重音是相对的，是组合

关系上的一种主观感觉，是在同前后音节强度的比较中得到的。不同话语之间的音节轻重比较是没有意义的（沈炯，2001）。从这个角度看，“每个音节都是重音”的观点似乎是有问题的，因为如所有音节都重，也就失去了凸显性要素和对比性特征，那么汉语的节奏类型也就无从谈起了。

正是基于“轻重只有在前后对比中才能体现”的观念，以往很多研究（见前面汉语词重音部分）都希望证明汉语词重音和位置有关系，并提出了“前重”“后重”等模式。但是上述结论除了前面已经谈及的争议外，也和目前世界语言节奏分类的结论矛盾，例如：(1) 现有节奏计算模型证明汉语是音节节奏型语言而非重音节奏型语言。(2) 如果汉语也存在音步，且遵循“前重”（或“后重”）模式，那么汉语就和英语一样都成为重音节奏型语言了，明显和目前汉语是音节计时型语言的观点不符；(3) 另外，现有的声学分析和听感实验不能很好支持汉语“前重”（或“后重”）的结论。如果脱离了证据支持，那么相应的音系学假设也就难以令人信服了。

#### 5.6 汉语的节奏支点

上述“两难”困境既是对汉语节奏研究的挑战，也是揭示汉语节奏本质的契机。

我们认为该问题最好的解决方案是找到各种理论的“最大公约数”和兼容性观点。所谓“最大公约数”，就是对汉语节奏问题可以达成的共识，目前主要有如下几条：

(1) 汉语的轻声（或轻读）音节是轻音，与之相对的有调音节是（正常）重音；

(2) 汉语是音节计时型语言。汉语语音学界尚未发现有对此观点有不同意见的观点。

目前，能够较好兼容上述共识的是殷治纲（2021）的观点，即(1) 汉语普通话语词重音不是和韵律位置关联的，而是和声调类别关联的。有调音节（包括阴平、

阳平、上声、去声)是重音,而无调音节(轻声或轻读)是轻音(某些方言中保留的入声调可是轻音,但普通话中已不存在)。这条说明汉语至少在声调(有无)类别上是存在轻重对立的,不存在无对比性特征之说。(2)根据第(1)条规则,汉语普通话两音节韵律词可以分为“重重”式(音节均为有调音节)和“重轻”式(词尾音节为轻声或轻读)。由于汉语轻声(或轻读)音节很少,所以“重轻”式较少,而“重重”式占比很大,其特点就是各音节都是重音。从统计频率上可以支持汉语是音节计时型语言——几乎各音节都重的观点。

另外,在语言类型学方面,江荻(2011)比较了声调、重调和重音的关系,分析了它们之间的内在联系性,证明声调系统和重音系统是有关的。也支持了上述研究的结论。

综上可知,汉语的节奏类型是音节计时型语言,可以认为其节奏支点是以“声调有无”为基础的轻重对比特征。

## 6. 结语

本文较系统地梳理了汉语节奏研究的相关内容,并就汉语的节奏类型问题进行了探讨,其内容主要包括以下几部分。

首先,文章认为语言节奏是指语言感知上的凸显性要素在时间序列上规律性出现的模式。该概念包含两大特征:第一类特征是由凸显性要素构成的对比性特征;第二类特征是节奏单元在时间序列上(或组织模式上的)的规律性。

在第一类对比性特征相关研究中,文章介绍了基于轻重特征的研究、基于松紧(或停延)特征的研究和基于声调特征的研究。

在第二类特征相关研究中,文章梳理了节奏单元的线性模型和层级模型,并介绍了相关研究内容。

之后,文章介绍了语言节奏的计算模型。

文章结尾部分介绍了不同语言的节奏类型,分析了语言节奏类型是在什么韵律

层级上形成的,区分节奏类型的主要特征是什么,并分析了重音计时节奏型语言(英语)和音节计时节奏型语言(汉语)的特点。

最后,文章着重讨论了汉语的节奏类型问题。针对以往汉语节奏类型研究中面临的“两难困境”,本文基于“最大公约数”原则提出了一种兼容方案,其核心观点认为汉语节奏支点是以“声调有无”为基础的轻重对比特征,为汉语节奏研究提供了一种新观点和新思路。

总之,本文希望能够在有限篇幅内对汉语节奏研究的脉络和框架给予系统性梳理,并对汉语节奏支点等重点问题给出相对明确的观点。但由于语言节奏研究的复杂性以及作者自身水平的限制,文章内容难免存在纰漏、偏颇之处,还希望能得到学界同仁的批评指正,使研究内容趋于客观、准确和完善。

## 参考文献

- Ahn, S. (2018). The role of tongue position in laryngeal contrasts: An ultrasound study of English and Brazilian Portuguese. *Journal of Phonetics*, 71, 451–467.
- Cao, J., & Maddieson, I. (1992). An exploration of phonation types in Wu dialects of Chinese. *Journal of Phonetics*, 20, 77–92.
- Chen, Z. (2015). Breathy voice and low tone. *Journal of Chinese Linguistics*, 43 (1A), 90–117.
- Chiba, T., & Kajiyama, M. (1941). *The vowel, its nature and structure*. Tokyo: Kaisekian.
- Chomsky, N., & Halle, M. (1968). *Sound pattern of English*. New York: Harper & Row.
- Fant, G. (1970). *Acoustic theory of speech production* (2nd ed.). The Hague: Mouton and Co.
- Guthrie, M. (1967 – 1970). *Comparative Bantu*. Farnborough: Gregg.
- Jakobson, R., Fant, G., & Halle, M. (1952). *Preliminaries to speech analysis*. Cambridge: MIT Press.
- Johnson, K. (1997). The auditory/perceptual basis for speech segmentation. *Ohio State University Working Papers in Linguistics*, 50, 101–113.
- Johnson, K. (2006). Resonance in an exemplar-based lexicon: The emergence of social identity and phonology. *Journal of Phonetics*, 34, 485

- 499.
- Labov, W. (1994). *Principles of linguistic change. Vol. 1: Internal factors*. Oxford: Blackwell.
- Ladefoged, P., & Maddieson, I. (1996). *The sounds of the world's languages*. Oxford: Blackwell.
- Liberman, A. M., & Mattingly, I. G. (1985). The motor theory of speech perception revised. *Cognition*, 21, 1–36.
- Liberman, A. M., Cooper, F. S., Shankweiler, D. S., & Studdert-Kennedy, M. (1967). Perception of the speech code. *Psychological Review*, 74, 431–461.
- Liberman, A. M., Harris, K. S., Hoffman, H. S., & Griffith, B. C. (1957). The discrimination of speech sounds within and across phoneme boundaries. *Journal of Experimental Psychology*, 54, 358–368.
- Lieberman, P. (1984). *The biology and evolution of language*. Cambridge: Harvard University Press.
- Mathangwane, J. (1996). *Phonetics and phonology of Ikalanga: A diachronic and synchronic study*. PhD dissertation, University of California, Berkeley.
- Möttönen, R., & Watkins, K. E. (2009). Motor representations of articulators contribute to categorical perception of speech sounds. *Journal of Neuroscience*, 29 (31), 9819–9825.
- Odden, D. (2005). *Introducing phonology*. Cambridge: Cambridge University Press.
- Ohala, J. J. (1974). Experimental historical phonology. In J. M. Anderson & C. Jones (Eds.) *Historical linguistics* (pp. 353–389). Amsterdam: North-Holland.
- Ohala, J. J. (1983). The origin of sound patterns in vocal tract constraints. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 189–216). New York: Springer-Verlag.
- Ohala, J. J. (1987). Experimental phonology. *Proceedings of the 13th Annual Meeting of the Berkeley Linguistics Society*, 207–222.
- Ohala, J. J. (1990). There is no interface between phonology and phonetics: A personal view. *Journal of Phonetics*, 18, 153–171.
- Ohala, J. J. (1993). The phonetics of sound change. In C. Jones (Ed.), *Historical linguistics: Problems and perspectives* (pp. 237–278). London: Longman.
- Ohala, J. J. (1997). Emergent stops. *Proceedings of the 4th Seoul International Conference on Linguistics*, 84–91.
- Ohala, J. J., & Jaeger, J. J. (Eds.). (1986). *Experimental phonology*. Orlando: Academic Press.
- Pierrehumbert, J. B. (2001). Exemplar dynamics: Word frequency, lenition, and contrast. In J. Bybee & P. Hopper (Eds.) *Frequency effects and the emergence of lexical structure* (pp. 137–157). Amsterdam: John Benjamins.
- Stanton, J. (2018). Environmental shielding is contrast preservation. *Phonology*, 35, 39–78.
- Stevens, K. N. (1972). The quantal nature of speech: Evidence from articulatory-acoustic data. In P. B. Denes & E. E. David Jr. (Eds.) *Human communication: A unified view* (pp. 51–66). New York: McGraw Hill.
- Stevens, K. N. (1989). On the quantal nature of speech. *Journal of Phonetics*, 17, 3–46.
- Stevens, K. N. (2003). Acoustic and perceptual evidence for universal phonological features. *Proceedings of the 15th International Congress of Phonetic Sciences*, 33–38.
- Stevens, K. N., & Keyser, S. J. (2010). Quantal theory, enhancement and overlap. *Journal of Phonetics*, 38, 10–19.
- Storto, L. R. (1999). *Aspects of a Karitiana grammar*. PhD dissertation, MIT.
- Watkins, K. E., & Paus, T. (2004). Modulation of motor excitability during speech perception: The role of Broca's area. *Journal of Cognitive Neuroscience*, 16 (6), 978–987.
- 鲍怀翹：《再谈语音量子理论》，《中国语音学报》2015年第5辑，第1—10页。
- 陈忠敏：《吴语清音浊流的声学特征及鉴定标志——以上海话为例》，《语言研究》2010年第3期，第20—34页。
- 陈忠敏：《重论浊阻塞音、气嗓音与低调的关系》，待刊。
- 李智强：《区别特征理论的语音学基础》，《中国语音学报》2014年第4辑，第27—37页。
- 石锋：《苏州话浊塞音的声学特征》，《语言研究》1983年第1期，第49—83页。
- 周颖异、朱晓农：《江西修水渣津赣语的十六调系统》，《中国语文》2020年第5期，第570—590页。
- 朱晓农：《语言语音学和音法学：理论新框架》，《语言研究》2011年第1期，第64—87页。

**殷治纲**

博士，中国社会科学院语言研究所、语言研究所语料库暨计算语言学研究中心副研究员。主要研

究领域为语音学、语料库语言学、法律语言学，以及语音和语法、语义、语用的接口研究。

E - mail: yinzhg@cass.org.cn

[本文原载《中国语音学报》第18辑，2022年]