

 中国社会科学院创新工程学术出版资助项目



中国社会科学院文库 · 文学语言研究系列  
The Selected Works of CASS · Literature and Linguistics

# 普通话焦点的语音实现和音系分析

PHONETIC REALIZATION AND PHONOLOGICAL ANALYSIS OF  
FOCUS IN STANDARD CHINESE

 中国社会科学院创新



中国社会科学院文库 · 文学语言研究系列

The Selected Works of CASS · Literature and Linguistics

# 普通话焦点的语音实现和音系分析

PHONETIC REALIZATION AND PHONOLOGICAL ANALYSIS OF  
FOCUS IN STANDARD CHINESE

贾 媛 著



中国社会科学出版社

## 图书在版编目(CIP)数据

普通话焦点的语音实现和音系分析/贾媛著. —北京:  
中国社会科学出版社, 2012. 10  
ISBN 978 - 7 - 5161 - 1699 - 9

I. ①普… II. ①贾… III. ①普通话—语音学—  
研究 IV. ①H116

中国版本图书馆 CIP 数据核字(2012)第 263497 号

---

出 版 人 赵剑英  
责任编辑 张 林  
责任校对 张洪波  
责任印制 戴 宽

---

出 版 中国社会科学出版社  
社 址 北京鼓楼西大街甲 158 号(邮编 100720)  
网 址 <http://www.csspw.cn>  
中文域名:中国社科网 010 - 64070619  
发 行 部 010 - 84083685  
门 市 部 010 - 84029450  
经 销 新华书店及其他书店

---

印刷装订 三河市君旺印装厂  
版 次 2012 年 10 月第 1 版  
印 次 2012 年 10 月第 1 次印刷

---

开 本 710 × 1000 1/16  
印 张 16.5  
插 页 2  
字 数 273 千字  
定 价 50.00 元

---

凡购买中国社会科学出版社图书,如有质量问题请与本社联系调换  
电话:010 - 64009791  
版权所有 侵权必究

## 《中国社会科学院文库》出版说明

《中国社会科学院文库》（全称为《中国社会科学院重点研究课题成果文库》）是中国社会科学院组织出版的系列学术丛书。组织出版《中国社会科学院文库》，是我院进一步加强课题成果管理和学术成果出版的规范化、制度化建设的重要举措。

建院以来，我院广大科研人员坚持以马克思主义为指导，在中国特色社会主义理论和实践的双重探索中做出了重要贡献，在推进马克思主义理论创新、为建设中国特色社会主义提供智力支持和各学科基础建设方面，推出了大量的研究成果，其中每年完成的专著类成果就有三四百种之多。从现在起，我们经过一定的鉴定、结项、评审程序，逐年从中选出一批通过各类别课题研究工作而完成的具有较高学术水平和一定代表性的著作，编入《中国社会科学院文库》集中出版。我们希望这能够从一个侧面展示我院整体科研状况和学术成就，同时为优秀学术成果的面世创造更好的条件。

《中国社会科学院文库》分设马克思主义研究、文学语言研究、历史考古研究、哲学宗教研究、经济研究、法学社会学研究、国际问题研究七个系列，选收范围包括专著、研究报告集、学术资料、古籍整理、译著、工具书等。

中国社会科学院科研局

2006年11月

## 前 言

“焦点”(focus)是音系学、句法学、语义学、话语分析等语言学各个研究学科共同感兴趣的问题。“焦点”在不同领域中有不同的解释,在语言的形式理论描述中,它是句子中某个句法成分的一种特征,在意义上是讲话人认为比较重要,需要通过语言手段强调的成分。

“焦点”可以分为常规焦点、对比焦点和话题焦点,焦点可以通过句法、词法、音系和语音等手段来体现(张伯江、方梅 1996,刘丹青、徐烈炯 1998,徐烈炯 2005)。近些年,在语音学和言语工程界,“焦点”问题也成为研究和应用的热点,随着人一机语音交互技术的发展,在语音应用上,需要机器能够听懂自然输入的言语,找到信息的焦点,并进行相应的操作,例如在进行信息检索时,能够做出正确的应答。在这一应用中,主要涉及语音识别、对话理解以及语音合成等方面对焦点问题的研究和应用。焦点是人们交际过程中,交际双方想要传递的凸显信息,对焦点信息的正确识别对言语理解至关重要,而正确设置焦点信息位置,并进行正确的声学表达,则是合成自然口语的重要保证。因此,针对焦点的各种“接口”研究就显得非常重要了。

与语音接口的研究,一般关注焦点位置和非焦点位置的语音表现,探讨窄焦点和宽焦点的声学特征,林茂灿先生在2012出版的《汉语语调实验研究》中对这个问题进行了大量的实例分析;与句法接口的研究,探索各种句法结构与焦点的关系,以及焦点标记等;与音系接口的研究,更多偏重音系的描写,以及焦点与重音实现的关系;与篇章的接口研究,则关注信息结构与焦点的关系,以及焦点的语用功能等。贾媛的这本研究专著的特点是从语音学、音系学和句法等多个层面对普通话的焦点进行接口研究,设计了大量严格的实验对话材料,音系分析建立在大量的语音实验结果基础上。该研究通过分析不同类别和不同数量焦点声学特征,系统地考察了普通话焦点重音的声学特征,揭示了汉语焦点与重音的不对应性特征,以及多焦点句的重音和音系

实现规律，并采用核心重音和核心前重音来解释重音的层级问题。

尽管该研究主要关注“汉语的焦点”问题，但它实质上也是汉语语调研究的一个非常重要的构成部分。在对各种焦点句的语音表现和音系实质研究的基础上，该研究从音系层面对汉语的语调结构进行了系统描写，提出了汉语语调描写的有限状态文法，并指出构成汉语语调的音系事件包括：核心重音、核心前重音和边界调。研究中采用的描写方式，结合了汉语声调特征的音系特征和重音的实现方式，采用原始的 *H*调和 *L*调，描写汉语语调的组构形式，体现了汉语语调与声调的作用关系，可以说，将赵元任先生提出的“橡皮带效应”或者“小波浪加大波浪”关系进行了解读和形式化。

吴宗济先生从 20 世纪 80 年代，就开始面向语音合成，进行了一系列的汉语语调与声调关系的研究，特别是考察了焦点（加强）位置对于语调变化模式的影响，提出了语调的语音和音系描写体系。他在世的时候，经常鼓励贾媛从音系和语音的角度，利用音系学理论，基于汉语音事实对汉语焦点和语调的开展研究。这本专著能在赵元任先生诞辰 120 周年、吴宗济先生诞辰 103 周年的时候出版，是后学们对前辈的最有意义的怀念了。

书中对焦点和重音的概念，焦点作用域，重音实现方式以及焦点和重音的对应关系等基本概念都有详细的介绍。部分内容在人机语音工程通讯会议上发表，获得了优秀论文奖，还有一些章节在国际会议上宣读，得到了很高的评价。是近年来一部难得的从语音、音系和句法等角度对汉语焦点进行系统研究的专著，相信对语言学、语音学以及言语工程领域的研究人员和研究生都将有非常重要的参考价值。

李爱军

2012 年 10 月 15 日

## 摘 要

本研究主要考察普通话语调结构、韵律结构以及焦点的关系，主要研究普通话不同类别以及不同数量焦点的韵律特征。内容主要涉及三个方面：(1) 焦点成分的声学特征，如疑问词引导的(主位和述位)焦点、句法标记的焦点以及两类焦点的交互作用在韵律上的表现；(2) 由不同类别和数量焦点所传达的重音的层级和音系表征；(3) 焦点和重音对应关系的理论解释。根据研究的目的，本研究采用 Pierrehumbert、Ladd、Grice、Gussenhoven 等人提出的语调音系学理论，Selkirk、Nespor 和 Ladd 等人提出的韵律层级理论，以及有关语调的优选论分析模式（见 Gussenhoven 2005, Yip 2002），来解释普通话的焦点的分布问题、焦点和重音的对应关系问题，以及普通话的语调音系表征模式。本研究遵循实验室音系学研究思路，将实验数据与理论分析相结合，在实验数据的基础上概括普通话的音系范畴特征。

本研究所采用的实验句子均为陈述句，实验用句根据研究目的的不同，包括单焦点句（疑问词引导的焦点或者句法标记的焦点，如“连”标记或“是”标记的焦点）、双焦点句（包括疑问词引导的主位焦点和述位焦点，双述位焦点以及句法标记焦点和述位焦点）、多焦点句（多个疑问词引导的述位焦点），以考察普通话中不同数量和不同类别的焦点句，对短语层面突显，韵律切分以及语调模式影响，为跨语言的焦点句的研究提供有力证据。

本研究以音高和时长为参数，考察由上述焦点类型所体现的不同类别的重音模式。在分析焦点成分的声学特征时，将每个音节的音高目标值 H 或 L 特征作为分析的基本单元。本研究指出：(1) 在单焦点情况下，无论焦点成分为疑问焦点或者句法标记焦点（“是”或“连”

标记的焦点), 重音与焦点是一一对应的关系。以音高为参数, 句法标记的焦点与疑问词引导的焦点的语音特征类似, 其将焦点成分的音高音域显著抬高而将焦点后成分的音高显著压低。焦点成分的声调组合对重音实现的方式有重要的影响, 如重音可将焦点成分的 H 和 L 调显著抬高, 而以 H 调为声学特征主要体现位置, 但焦点成分的声调组合不影响重音的分布。以时长为参数, 不同类别的焦点都对时长有显著的拉长作用, 但疑问词引导的焦点的作用比句法标记要显著; (2) 在无句法标记结构中, 在双焦点情况下, 由焦点所传达的重音有层级差异。在无句法标记结构中, 主要发现包括以下几个方面: ① 不用层级的焦点可以由不同层级的重音来体现, 如由疑问词引导的焦点主位焦点和述位焦点, 由于述位焦点的层级高于主位焦点, 所传达的重音层级也以述位焦点为主重音; ② 相同层级的疑问焦点引导的述位焦点, 重音的层级也不相同; ③ 双焦点对焦点之间成分的音高无压低作用; ④ 在多个疑问词引导的焦点, 焦点与重音不是一一对应的关系, 只有最右位置的焦点可以实现为重音; ⑤ 在本研究中, 采用核心重音和核心前重音来解释重音的层级问题, 核心重音是重音成分中主要的、非选择性成分, 而核心前重音则是次要的、选择性的成分; ⑥ 疑问词引导的焦点对韵律切分也有影响, 可以将普通话语调切分为中间短语。(3) 在有句法标记的结构中, 双焦点情况下, 根据疑问词和句法标记焦点成分的组合关系主要发现有以下几点: ① 疑问焦点和句法标记焦点分布于同一成分时, 可以实现为核心重音, 对焦点下成分的音高抬高和时长的拉长作用比单焦点更显著, 对后接成分的音高压低也更显著; ② 疑问焦点和句法标记的焦点可以分布于句子的不同成分上, 当句法标记焦点 (“连” 标记焦点) 位于疑问焦点之前, 句法标记焦点和疑问焦点分别实现为核心前重音和核心重音, 前者音高和时长变化比后者幅度小。当句法标记焦点位于疑问焦点之后, 句法标记焦点则失去重读, 只有疑问焦点实现为核心重音。

根据实验结果以及以往语调音系研究, 本研究进一步提出了具有语言学意义的普通话语调音系表征模式。该语调模式主要描述普通话字调层面以上的语调变化模式, 主要包括四类音系事件: 核心重音, 核心前重音, 调头和边界调。其中核心重音和边界调是非选择性成分,



而核心前重音和调头则是选择性成分。音系事件的语音实现是以 H 和 L 调的描述为基础的, 语调的模式是以局部的音系事件为基础, 按照线性的序列组合而成。核心重音和核心前重音的内在特征不同, 核心重音具有唯一性特征, 核心重音实现是无标记的, 而核心前重音的出现却是有标记的, 在普通话中只出现在双焦点情况下。在上述焦点语音实现的描述中, 多焦点情况下只有一个重音出现, 而在有句法标记情况下, 也可以失去重音。因此, 在普通话中, 有焦点的位置未必由重音来表征。根据以上分析, 本研究提出制约表层核心重音实现的底层原因, 即节律上相对的“轻-重”关系, 节律上相对“重”的成分获得核心重音, 节律上相对“轻”的成分则根据焦点环境失去重音或者实现为核心前重音, 如在单焦点环境下, 节律结构为“重-轻”, 因此, 焦点位置实现为核心重音, 而焦点后成分则失去重音; 在双焦点以及多焦点环境下, 节律结构为“轻-重”, 在双焦点情况下, 实现为核心前重音和核心重音, 而在多焦点情况下, 只有最右位置焦点实现为核心重音。

在优选论的框架下, 采用普遍性的制约条件等级排列, 来解释从焦点到重音的实现过程, 共分为三个步骤: 表层重音类型的生成包括:  $H^*$ ,  $L^*$ ,  $LH^*$  或者  $H^*L$ ; 重音的连接位置以及韵律切分位置。

**关键词:** 焦点, 核心重音, 核心前重音, 优选论

## Abstract

The subject matter of this research is the relation between intonational structure, prosodic structure, and focus in Standard Chinese (hereinafter SC). It is an investigation of prosodic effects of various kinds and different numbers of focuses in SC, on the one hand, and of the accent patterns conveyed by the focus, on the other. The research examines three aspects: (i) the acoustic manifestations of the entities bearing focus, that is, *wh*-elicited focus, *syntax*-marked focus (*lian*-marked or *shi*-marked focus), and interaction of these two kinds of focuses; (ii) the phonological representation of the accents pattern and hierarchical levels of accents induced by the focus in discussion; and (iii) a theoretical explanation of the corresponding relation between focus and accent in SC. To accomplish this aim, this research employs (a version of) the theory of *Intonation*, as developed by Pierrehumbert and colleagues, Ladd, Grice, Gussenhoven, and others, (a version of) the theory of the *Prosodic Hierarchy*, as developed in the work of Selkirk, Nespor, and Ladd and (a version of) the theory of *Optimality Theory*, as developed by Gussenhoven and Yip, among others. The chosen approach here is simultaneously theoretically and empirically based, much along the lines of laboratory phonology research, in which experimentally collected speech data is adopted to investigate questions about the abstract categories of phonological structure.

The sample sentences employed in this research are confined to declarative sentences with one or more (single, double or multiple) focuses. The inclusion of focus in the domain of research is motivated by the fact

that the focus affects phrasal prominence, prosodic phrasing, and other aspects in intonation in various languages. It is expected that the present examination of focus may provide important evidences for the prosodic organization of SC.

Fundamental frequency ( $F_0$ ), together with duration, is an acoustic measure that is adopted to signal the contrast in focus conditions. The underlying tonal target *H* (*High*) or *L* (*Low*) is employed in dealing with the phonetic realization of focus phenomenon. It is mainly concerned with the following aspects: (i) in the single-focus condition, regardless of the kind of focus; i.e., whether it is *wh*-operators elicited information focus, *lian*-marked focus, or *shi*-marked focus, the accent is correlated with the focus. Taking  $F_0$  into consideration, the *syntax*-marked focus resembles the phonetic nature of the *wh*-elicited information induced focus.

Specifically, in the under-focus domain, the  $F_0$  ranges of the focused words are expanded as the *H* and *L* tones of the focused syllables are raised. The accent that results from a *syntax*-marked focus also exerts a compressive effect on the following constituents that can extend to the every end of a sentence. The tonal combinations of the focused constituents contribute greatly to the specific manner in which an accent is manifested, that is, the accent is realized by modifying the original tonal targets *H* or *L*.

As for the durational adjustment, the focus can trigger a significant effect on lengthening; (ii) in the syntax unmarked sentence, in double focus condition, what is crucially involved in the acoustic reflection of focus is the hierarchical level of accents. Concerning the  $F_0$  and durational pattern, the observations are threefold: (a) different levels of focus can lead to a different magnitude of pitch register raising and durational lengthening in a sentence; i.e., as between *wh*-elicited theme focus and rheme focus, the primary role is due to the rheme focus; (b) double rheme focus exhibits similar acoustic manifestations, namely, each of the focused units are characterized by  $F_0$  lifting and durational lengthening; (c)

consistent non-lowering of  $F_0$  on the second constituent reveals the fact that the accents have a primary status; (d) the accent of multiple rheme focus displays distinct properties in comparison with the single and double focus condition in that the distribution of the accent exhibits no corresponding relation with the focus. Only the last focused constituent of a sentence serves as the anchor in realizing accent in the surface form; (e) the hierarchical level differences of accents are adequately accounted for by the *nuclear accent* and *pre-nuclear accent* distinction in SC. The *nuclear accent* bears the obligatory and primary characteristics while the pre-nuclear tone is optional and secondary; (f) focus is also assigned the crucial role of determining the phrasing at the intermediate phrase boundary in SC; (iii) with regard to a syntax-marked sentence, the acoustic evidence for the double focus is based on the relation between the syntax-marked focus and the *wh*-elicited rheme focus: (a) they can combine together to induced a *nuclear accent* correlated with the larger  $F_0$  excursion and durational lengthening, and a more obvious  $F_0$  drop of the post focus constituents; (b) they may co-exist on different components. When the syntax-marked focus is located in the proceeding position of the rheme focus, these two focuses can be marked by *pre-nuclear accent* and *nuclear accent*. These two accents also observe a phonetically degree-based difference, in that the latter is marked by more obvious phonetic features than the former. Crucially, when the *syntax*-marked focus is preceded by the rheme focus, it is de-accented by the compressive effect of the preceding rheme focus.

Based on empirical investigations, as well as the existing literature on the phonological study of intonation, this study further proposes a phonological representation of the linguistically significant non-lexical configurations in SC. The grammar states that the intonation contour consists of four types of categorical phonological events; i.e., *nuclear accent*, *pre-nuclear accent*, *onset*, and *boundary tone*. Among these four entities, the *nuclear accent* and the boundary tone bear an obligatory

nature, whereas, the appearance of the *pre-nuclear accent* and the onset are optional. The intonation patterns are constructed by phonological events in a linear sequence, and are represented by the primitive tonal targets *H* and *L*. The phonetic realization of any given *H* or *L* tone depends on a variety of factors (e.g., status and number of focuses, position in an utterance) that are of essential importance to the identity of those tonal targets. Overall trends of contours (e.g., local range expansion or gradual compression of overall range) mostly reflect the operation of localized linguistic factors. The essential properties of *nuclear accent* and *pre-nuclear accent* are distinct from each other; the former is unmarked and bears a unique feature, while the latter is restricted by the focus condition in that it only appears in a double-focus utterance. Thereafter, there is a status difference between these two accents; they exhibit a primary and secondary status distinction. Evidence of ‘focus without accent’ and ‘accent without focus’ in SC further demonstrates that sentence-level accent is not only a matter of where the accent is located, but also involves a ‘*strong* or *weak*’ relation in the prosodic structure that in turn determines the distribution of the accent in the surface form. Under a single focus condition, the metrical relation is ‘*s-w*’ that determines the accenting and de-accenting phenomenon in the surface form. Under a double or multiple focus environment, the metrical structure is captured as ‘*w-s*’ that determines the distribution of the *nuclear accent* in the rightmost position elicited by the focus in SC sentences.

Within the framework of Optimality Theory, a ranking of constraints in the generation of accent in the surface form is conducted in three steps: (i) the analysis of accent pattern generation, i.e., *H\**, *L\**, *LH\** or *H\*L*; (ii) the association of the accents; (iii) the location of the prosodic boundary.

**Key words:** focus, nuclear accent, pre-nuclear accent, Optimality Theory

# Abbreviations

**Ad:** Adverb

**AM Theory:** Autosegmental-Metrical Theory

**BF:** Broad focus

**BrE:** British English

**Con:** Constraint

**C-ToBI:** Chinese Tone and Break Index

**DTE:** Designated Terminal Element

**DF:** Double focus

**Eval:** Evaluator

**F<sub>0</sub>:** Fundamental frequency

**FTA:** Focus-to-Accent

**Gen:** Generator

**IPO:** Instituut voor Perceptie Onderzoek in Holland (Institute for Perception Research)

**IViE:** Intonational Variation in English

**NA:** Nuclear accent

**NF:** Narrow focus

**Obj:** Object

**OT:** Optimality Theory

**PRA:** Pre-nuclear accent

**RB:** Rheme background

**RF:** Rheme focus

**Sub:** Subject

**SC:** Standard Chinese

**SVO:** Subject, Verb, Object

**TB:** Theme background

**TF:** Theme focus

**ToBI:** Tone Break Index

**USFP:** Unique Strong Focus Principle

**V:** Verb

***Wh*-elicited focus:** *Wh*-question elicited focus

# Symbols

**%**: Boundary tone

**L%**: Low boundary tone

**H%**: Hing boundary tone

**H**: High target of tone

**L**: Low target of tone

**H\***: Accented high target

**L\***: Accented low target

**\***: Violation of the constraints in Optimality Theory

**\*!**: Fatal violation in ranking constraints

**☞**: Optimal candidate in ranking constraints

**[+LianF]**: *Lian* marked focus

**[+LianF & +RF]**: *Lian* marked focus and rheme focus

**P**: Prominence

**[+ShiF]**: *Shi* marked focus

**[+ShiF & +RF]**: *Shi* marked focus and rheme focus

**S-W**: Strong-Weak

**W-S**: Weak-Strong

**T**: Tonal events

# Abbreviations

**Ad:** Adverb

**AM Theory:** Autosegmental-Metrical Theory

**BF:** Broad focus

**BrE:** British English

**Con:** Constraint

**C-ToBI:** Chinese Tone and Break Index

**DTE:** Designated Terminal Element

**DF:** Double focus

**Eval:** Evaluator

**F<sub>0</sub>:** Fundamental frequency

**FTA:** Focus-to-Accent

**Gen:** Generator

**IPO:** Instituut voor Perceptie Onderzoek in Holland (Institute for Perception Research)

**IViE:** Intonational Variation in English

**NA:** Nuclear accent

**NF:** Narrow focus

**Obj:** Object

**OT:** Optimality Theory

**PRA:** Pre-nuclear accent

**RB:** Rheme background

**RF:** Rheme focus

**Sub:** Subject

**SC:** Standard Chinese

**SVO:** Subject, Verb, Object

**TB:** Theme background

**TF:** Theme focus

**ToBI:** Tone Break Index

**USFP:** Unique Strong Focus Principle

**V:** Verb

***Wh*-elicited focus:** *Wh*-question elicited focus



# Table of Contents

## **Chapter One Introduction / 1**

### **1.1 Introduction / 1**

### **1.2 Literature review / 1**

#### **1.2.1 Intonational phonology / 1**

#### **1.2.2 Prosodic phonology / 3**

##### **1.2.2.1 Syntax-phonology mapping / 3**

##### **1.2.2.2 Prosodic domains / 4**

##### **1.2.2.3 Properties of prosodic structure / 5**

#### **1.2.3 Focus / 6**

##### **1.2.3.1 Focus from three different angles / 7**

##### **1.2.3.2 Focus in phonetic approach / 8**

##### **1.2.3.3 Focus in phonological approach / 10**

### **1.3 Significance of the research / 14**

### **1.4 Basic concepts in this study / 16**

#### **1.4.1 Definition of focus / 16**

#### **1.4.2 Nuclear accent and pre-nuclear accent / 19**

#### **1.4.3 Relative prominence / 20**

### **1.5 Outline of this research / 20**

## **Chapter Two Methodology / 23**

### **2.1 Background and claims / 23**

### **2.2 Design of acoustic experiment / 24**

#### **2.2.1 Materials selection / 24**

#### **2.2.2 Syntactic structure of target sentence / 25**

##### **2.2.2.1 Unmarked structure / 25**

##### **2.2.2.2 *Lian...dou* structure / 27**