

汉语普通话中双焦点的产出及其感知*

王 蓓^{1†} 刘 璐¹ 张夏夏¹ Caroline Féry²

(1 中央民族大学 北京 100081)

(2 德国法兰克福大学 法兰克福 60323)

2017 年 2 月 22 日收到

2017 年 7 月 13 日定稿

摘要 以往的研究发现双焦点句中第一个焦点后往往缺少音高压缩。这一现象是否与两个焦点间的音节数有关? 实验控制了焦点词的位置和句子长度两个因素, 发现双焦点句中两个焦点既相互独立又相互依存。独立性表现为每个焦点有各自的实现范围, 其音高升高和时长延长都和所对应的单焦点的条件接近; 依存性表现为两个焦点可以在同一个语调短语内共存。第一个焦点后因为有了韵律短语边界而缺少音高压缩, 而与两个焦点间的音节数没有直接关系。感知实验的结果显示: 双焦点的辨认正确率显著低于句首单焦点。这主要因为同时判别两个焦点的难度比较高。如果只要求听音人关注第一个焦点, 其辨认率明显提高, 且与句首单焦点的判断正确率无显著差异。可见, 在双焦点句中, 句首和句末焦点的判断有不对称性。音高和时长的增加可以保证句首焦点的正确感知, 但不能保证句末焦点的正确感知。总体来说, 两个焦点中句首焦点在感知上有更高的突显度。

PACS 数: 43.71

Dual-focus production and perception in Putonghua

WANG Bei¹ LIU Lu¹ ZHANG Xiaxia¹ Caroline Féry²

(1 *Minzu University of China* Beijing 100081)

(2 *Goethe University Frankfurt* Frankfurt 60323)

Received Feb. 22, 2017

Revised Jul. 13, 2017

Abstract Previous studies showed that in dual focus sentences, Post-Focus F0 Compression (PFC) was absent after the first focus. Was it related to the number of syllables between the two foci? In the present experimental study, focus position and sentence length were controlled. The results showed that the two foci were independent as well as dependent on each other. They were independent because each focus had its own domain, and its F0 rising and duration lengthening reached to the same degree as its single focus counterparts. They were dependent because the two foci could coexist in one intonation phrase. The lack of PFC after the first focus was due to the insertion of a prosodic phrase boundary, rather than to the number of syllables between the two foci. The perception experiments showed that the correct perception rate of dual focus was significantly lower than that of initial focus, probably because it was hard to detect two foci in one sentence. When the listeners were asked to only pay attention to the first focus, the correct perception rate was much higher, and it was similar to that of the initial focus counterpart. Therefore, we concluded that the perception of the two foci was asymmetric. The increase of F0 and duration guaranteed correct perception of an initial focus, but not a final focus. In general, the first focus was perceptually more prominent in a dual focus sentence.

* 中央民族大学“研究生培养工作坊”项目和国家社会科学基金项目(18BYY079)资助

† 通讯作者: 王蓓, wangbei@muc.edu.cn

引言

焦点是韵律研究的热点问题之一,主要是因为焦点与语义表达直接相关,有明确的交际功能。焦点是一个语义上的概念,意味着在焦点词的位置还有其它可供选择的信息(alternatives)^[1]。焦点位置不同,句子的意思也不同。如,“妈妈去了超市”中的“妈妈”成为焦点就意味着去超市的人还有可能是爸爸或是其他人,但这里强调的是在这些可能选项中,去超市的人是妈妈。而当“超市”成为焦点就意味着妈妈还可能去其它地方,但是在这些可能去的地方中,妈妈去的是“超市”。这个例子中可以看出焦点有着传达语义的作用。在汉语普通话中,焦点可以通过句法实现,如“是妈妈去了超市”。焦点也可以通过韵律实现,也就是重音。在例句中,重音在“妈妈”或是“超市”时,意义是不同的。有趣的是,有“是”标记的情况下,焦点词仍然有语音上的加重^[2]。也就是说,焦点的韵律形态在传达信息上有着句法所不能替代的作用。

很多语言中都发现,当句中有一个焦点时,焦点词的基频升高,时长延长,音强增强(英语^[3-4];德语^[5];汉语^[6-9])。值得注意的是,这些语言中,焦点后有音高骤降和音域压缩的现象,并且音强也有明显的降低,总称为焦点后压缩(Post-Focus Compression, PFC)。而焦点前音节的音高基本不受影响(汉语^[6-10];英语^[3-4,10])。Xu^[7]将其归结为焦点的三区段模式(tri-zone realization of focus)。另外,在汉语普通话中,句末焦点基频升高幅度较小,与中性焦点^①区别不明显^[7]。时长则不受焦点词位置的影响,无论焦点在什么位置,时长都会明显增加^[6-9]。并且,焦点词对其前后词的时长几乎没有影响。也就是说,汉语普通话中焦点对音高的调节范围是整句,但是对时长的调节基本仅限于焦点词。焦点后音高压缩甚至可以跨过有长停顿的语调短语边界^[11]。

与单焦点相比,双焦点的研究比较有限。如针对问题“谁去了哪儿”,在回答“妈妈去了超市”时就有两个焦点。虽然英语中很早就有关于双焦点的实验研究^[4],但这个问题后来并没有得到足够的重视,直到近几年才有了几项关于英语^[12]、德语^[13]和汉语^[14-18]的双焦点研究。这些研究一致的发现是两个焦点的时长都有延长,且基本相当于单焦点条件。音高的结果却出现了与单焦点不同的表现。

Eady 等人^[4]发现,双焦点条件中两个焦点的基频和时长增加量与相应位置的单焦点条件几乎没有区别。双焦点与单焦点最重要的不同是:双焦点句中第一个焦点后没有音高压缩。他们认为第二个焦点影响了第一个焦点的实现方式,在句子的韵律产出过程中存在逆向作用(anticipatory effect)。Liu^[12]更系统地研究了英语陈述句和疑问句中的双焦点问题,结果进一步支持了 Eady 等人^[4]的发现。Liu 还发现句长并不影响双焦点的韵律编码方式。也就是说,第一个焦点后缺少音高压缩并不是因为两个焦点间的音节数量少,或者是发音时间不够。

Wang 和 Féry^[13]在德语双焦点研究中系统地调节了主语短语(NP)和谓语短语(VP)的长短。结果发现,在主语短语比较长的情况下,有时第一个焦点后会增加一个高的边界调,将两个焦点分开,称为“双短语模式”(two-phrase pattern)。但更多的情况还是两个焦点在一个语调短语内同时实现,每个焦点都是典型的降调,并且第一个焦点后有明显的音高压缩,称之为“双峰模式”(two-peak pattern)。当句子非常短的时候,双焦点句是以“帽子模式”(hat-pattern)实现的,也就是说,第一个焦点是升调,且其后音高一直保持高平,到第二个焦点处出现降调。

汉语普通话关于双焦点的几项研究结果并不一致。Jia 等人^[15]用了一个全是二声的句子(即“刘民凌晨提拔毛兰了”)比较了单焦点、双焦点和多焦点的语调。结果发现,两个焦点(句首+句末)都有基频升高,但第一个焦点后并没有出现音高压缩。黄贤军等^[16]的实验也得到了同样的结果。Yuan 等人^[17]研究了更复杂的双焦点情况,用了双宾语的句式(S+O1+O2),且宾语是由修辞成分加名词构成的。如“张咪扔给窗边的八哥三根香蕉”。该句中第一个焦点为主语(“张咪”)。他们发现,第一个焦点后音高压缩是否出现与第二个焦点的位置有关。如果第二个焦点在 O1 中(“八哥”),动词上就没有焦点后压缩;而当第二个焦点在 O2 中时(“三根”或“香蕉”),动词和 O1 上都有明显的焦点后压缩。

Kabagema-Bilan 等人^[18]的研究却得到了不一样的结果。他们用的是由一声组成的短句(如“猫咪偷鸟窝”),比较了双焦点(句首+句中),单焦点(句首或句中)和中性焦点的语调。在他们的研究中,两个焦点之间的动词只有一个单音节。实验结果发现双焦点句子中的第一个焦点的基频并没有明显的变

① 中性焦点(neutral focus)在很多文章中被称为宽焦点(broad focus),或者是整句焦点(sentence focus)。从另一个角度理解,如果一个句子中所有的词都被强调,那么也就相当于句中没有焦点。焦点的意义在于与句中其它词相比更加突显的词。在 Jia 等人^[15]的研究中就发现,当句中有 3 个词成为焦点时,其语调曲线与中性焦点条件没有区别。本文的中性焦点指句中没有哪个词被特别强调的韵律现象。

化, 而只有时长的延长。第二个焦点则表现出了焦点后基频的压缩和时长延长, 但是焦点词基频升高很有限。他们认为, 双焦点的实现基本相当于只是实现第二个焦点。他们由此提出, 汉语中的重音设置满足重音最大化 (culminativity principle) 的音系规则。即 Hyman^[19] 提出的, 一个韵律成分 (韵律词, 韵律短语和语调短语) 中只能有一个节律上突显的调核 (head)。Jia^[15] 及 Yuan 等人^[17] 也提出, 汉语中的突显是分等级的。靠近句末位置的重音是核心重音 (nuclear accent), 而句首位置的重音是核心前重音 (pre-nuclear accent)。

总体来看, 汉语普通话中的双焦点实现方式在以往的研究中有不一致的结果。另外, 对于第一个焦点后是否有音高压缩还缺少更深入的理论解释。

还有一个以往研究得较少的问题是, 在缺少焦点后压缩的情况下, 双焦点句的这种实现方式是否还能保证两个焦点都被正确感知到, 特别是第一个焦点的感知。

汉语普通话的陈述句^[20] 和疑问句中^[21] 的单焦点 (除了句末焦点) 感知正确率达 90% 以上, 句末焦点的感知正确率却只有 60% 左右。句末焦点句的语调曲线和中性焦点相似, 句末焦点只是基频略有升高和时长延长^[7,20]。他们指出, 在句末焦点条件下没有焦点后压缩, 这可能是导致句末焦点感知正确率较低的主要原因。并且, 声调不影响陈述句中句首和句中焦点的感知, 但是句末焦点在高平调和降调句子中的感知正确率要高于在升调和低平调句子^[21]。

对于焦点的感知, Herment-Dujardin 和 Hirst^[22] 发现时长延长对焦点感知的贡献比较有限。相比之下, 焦点词的基频变化, 比如音域的扩大, 音高的大幅度变化是感知焦点更有效的线索。Rump 和 Hermes^[23] 进一步发现, 音高最大值在感知焦点时的作用要大于音高最小值。关于基频在焦点感知中的作用也有不一致的结果。如, Heldner^[24] 研究了瑞典语中的焦点感知, 结果发现, 虽然在产出实验中焦点是通过基频上升来实现的, 但是基频的升高并不足以使焦点被正确感知到。

对于双焦点句, Rump 和 Collier^[25] 在荷兰语中发现, 如果要感知到一个句子中有句首和句末两个焦点, 第一个焦点需要有明显的音高升高, 而第二个焦点的音高只需要有很小的升高。Yuan 等人^[17] 请听音人听到双焦点句后判断句中主要重音和次要重音的位置。结果发现, 句首词稳定地被判断为主要重音, 而第二个焦点的判断不稳定。由此, 他们提出, 汉语普通话中焦点是有层级的, 有主要焦点和次要

焦点之分。这一结果与 Jia^[15] 研究中发现的双焦点中后一个焦点是核心重音并不一致。我们猜测, 这可能是由于双焦点的感知受注意分配的影响, 对第一个焦点的判断更容易。而第二个焦点在缺少注意分配的情况下感知结果较差。

由以上文献综述可见, 双焦点的音高升高及焦点后压缩在不同的研究中结论并不一致。几项研究中都没有提出一个双焦点韵律形态的理论以解释这些不一致的结果。本文将系统改变两个焦点间的音节数, 但保持第一个焦点后韵律边界不变, 由此考察双焦点的实现方式是否受两个焦点间的距离影响。另外, 以往的研究中很少考虑双焦点的感知问题。这是本文关注的第二个问题。具体来说, 本文关注的两个研究问题以及预期如下:

(1) 汉语普通话中双焦点的韵律形态是怎样的? 是否受两个焦点间音节数的影响?

我们的预期是两个焦点上都会有基频升高和时长延长的现象。在不改变第一个焦点后韵律边界的情况下, 仅增加两个焦点间的音节数不会使第一个焦点后出现音高压缩。

(2) 双焦点的感知正确率如何? 影响双焦点感知的因素是什么?

单焦点的研究结果显示, 焦点后压缩是感知普通话焦点最有效的线索^[20]。因此, 我们预期如果没有焦点后压缩, 双焦点的感知正确率可能比较低。另外一个可能的情况是, 由于处于句末位置的第二个焦点不容易被感知而导致双焦点句感知的正确率下降。

1 双焦点产出实验

1.1 方法

1.1.1 材料

实验程序和 Eady 等人的^[4] 类似。发音人朗读带有不同焦点的句子, 焦点通过问句引出。控制的两个因素分别是焦点条件和句子长度。句子都是 SVO 结构, 有短句 (6 个音节) 和长句 (12 个音节) 两种句长条件。短句中有 3 个双音节词, 长句通过在短句的宾语前添加修饰语, 保持主语和动词不变 (见 (1a) — (1e))。实验用了 5 种不同声调组合的句子, 包括了 1—4 声。句中双音节词的声调按 11、13、31、42 和 24 组合在一起, 和 Wang 等人^[11] 研究中所使用的材料类似。选择这些声调组合的原因是这样可以包括所有的声调, 并且在每个词中便于提取基频的最大值和最小值。全是一声的句子可以排除降阶 (downstep)

的影响, 而直接观察焦点对语调的调节作用。

(1a) 汪英参观车间。

汪英参观山西清乡一街车间。

(1b) 伊猛修理桌椅。

伊猛修理西雅山井鸡厂桌椅。

(1c) 马威拧开顶灯。

马威拧开北京蒋山老屋顶灯。

(1d) 梁娜研制颜料。

梁娜研制重庆房县齐镇颜料。

(1e) 路苗借来链条。

路苗借来大同亮桥庆塘链条。

焦点条件有 4 个: 句首单焦点, 句末单焦点, 双焦点(句首焦点 + 句末焦点) 和中性焦点。以 (1a) 的短句为例, 引出焦点条件的 4 种问句见 (2a)–(2d)。

(2a) 句首单焦点: 谁参观车间?

(2b) 句末单焦点: 汪英参观什么?

(2c) 双焦点: 谁参观什么?

(2d) 中性焦点: 什么事?

每个发音人总共说 80 个句子, $2(\text{句长}) \times 4(\text{焦点条件}) \times 5(\text{声调组合}) \times 2(\text{重复}) = 80$ 个句子。总共分析了 400 个句子。这样的实验设计保证了句长并不改变第一个焦点后的句法边界。由此可以考察只是增加两个焦点间的音节数是否会使第一个焦点后出现音高压缩。

1.1.2 发音人

5 位普通话母语发音人(3 女 2 男) 参与了实验, 年龄在 22~30 岁之间, 没有任何言语和听力上的障

碍。参与实验获得少量报酬。

1.1.3 录音程序

录音在德国波茨坦大学语音实验室完成。发音人听到问句之后用恰当的语调自然朗读目标句, 正确强调相应的词, 并且尽可能在句中避免明显的停顿。采样率为 48 kHz。目标句通过一个自主编写的 Java 程序随机展现在电脑屏幕上^[26]。问句由本文第一作者录制。实验时还有 160 个填充句。每个发音人的录音时间大约是 1.5 小时。实验前有一个简短的练习。

1.1.4 声学测量

每个目标句以 wav 的形式存储。用 Praat 脚本(ProsodyPro)^[27] 测量目标句中各音节的基频和时长。为了提取可靠的基频值, 对 Praat 标记的声带振动周期进行了人工校正, 同时借助共振峰和能量曲线标记每个音节的边界。每句话中的声带振动周期标记, 边界标记和基频值都会自动保存到文本档里。Praat 脚本还可以计算出每个音节的基频最大值、基频最小值以及时长。基频用式 (1) 以 50 Hz 为参考值, 将 Hz 转化成半音(st), 这样可以基本平衡不同发音人间的差异。

$$F0(\text{st}) = 12 \times \log_2 \left(\frac{F0(\text{Hz})}{50} \right). \quad (1)$$

1.2 结果

1.2.1 语调曲线分析

借助 Xu^[27] 的 Praat 脚本文件, 在每个音节中提取 10 个等距离的基频点用于做语调曲线图。图 1

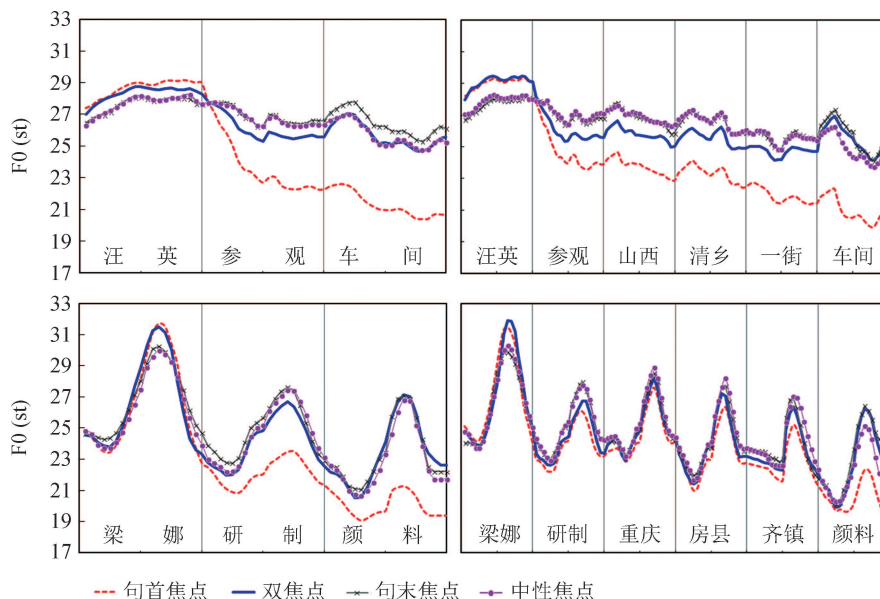


图 1 4 种焦点条件下 11 和 24 调短句及长句的语调曲线

是 4 种焦点条件下由 11 和 24 两种声调的词组成的短句和长句的基频曲线。图中每个点的基频值都是 5 个发音人重复两次的平均值，竖线表示词边界。其他实验句呈现出相同的语调模式，由于篇幅限制，不再逐一呈现。从图 1 中双焦点句的语调曲线可见：

(1) 双焦点条件下第一个焦点的基频几乎和句首单焦点的基频一样，比中性焦点句对应词语的基频值高；

(2) 第一个焦点后的基频没有降到和句首单焦点条件同样的水平；

(3) 第二个焦点的基频几乎达到了和句末单焦点相同的高度。

这些结果在短句和长句以及不同声调组合中都是一致的。长句双焦点条件下，第一个焦点后似乎有一定程度的音高压缩，还需要做进一步的统计检验。

1.2.2 基频分析

如上所述，实验句是 SVO 结构的，其中第一、第二以及最后一个词为目标词。图 2 是短句和长句中目标词的基频最大值。由图中可见：当目标词是焦点时，其基频最大值最高。单焦点和双焦点条件，以及焦点在首和末时焦点词都有音高升高。第二个词是焦点后的部分，它的基频值在句首单焦点条件下最低，双焦点条件下的基频值介于句首单焦点和中性焦点条件之间。

以焦点条件、句长和声调为自变量，对每个目标词的基频最大值做三因素重复测量的方差分析。第一个目标词的基频最大值的三因素重复测量方差分析结果表明，焦点条件 ($F(3, 12) = 10.85^*$) 和声调 ($F(4, 16) = 7.19^*$) 都有主效应，而句长 ($F(1, 16) = 0.791, n.s.$) 没有效应，三者之间没有交互作用。后效检验 (S-N-K) 表明，第一个词在句首单焦点 (30.2 st) 和双焦点条件下 (29.8 st) 基频最大值比在中性焦点 (28.8 st) 和句末单焦点 (28.7 st) 条件下更高 ($p < 0.05$)。这在短句和长句中都是一样的。然而，句首单焦点和双焦点条件下第一个词的基频最大值没有明

显差异。与图 1 的第一个观察结果一致，双焦点句中第一个焦点和句首单焦点的基频升高程度相当。

用相同的方法检验第二个目标词的基频最大值，结果表明焦点条件 ($F(3, 12) = 10.289^*$)，声调 ($F(4, 16) = 33.742^{***}$) 和句长 ($F(1, 4) = 21.2^*$) 都有主效应。后效检验 (S-N-K) 表明，句首单焦点条件下第二个词的基频最大值明显小于其它 3 种焦点条件；而在双焦点、句末单焦点和中性焦点条件之间并没有明显差异。也就是说，双焦点句中第一个焦点后没有明显的音高压缩。句长和焦点条件 ($F(3, 12) = 16.47^{**}$) 以及声调和焦点 ($F(12, 48) = 9.125^{**}$) 之间还存在交互作用。交互作用的来源是，短句中句首单焦点后的基频下降程度要大于长句中的。图 1 的第二个观察结果也得到了统计的支持，即双焦点条件下，第一个焦点后没有明显的音高压缩，且与句长无关。

对最后一个目标词的基频最大值做同样的统计分析，发现焦点条件 ($F(3, 12) = 38.275^{***}$)，句长 ($F(1, 4) = 8.428^*$) 和声调 ($F(4, 16) = 38.644^{***}$) 都有主效应，但三者没有交互作用。后效检验 (S-N-K) 结果表明，句首单焦点条件下的最后一个词比其它 3 种焦点条件下的基频更低，而中性焦点条件下的基频又比双焦点和句末单焦点条件下更低。另外，双焦点和句末单焦点句中最后一个词的基频并没有显著差异。总之，双焦点句中的第二个焦点的实现方式和句末单焦点基本一致。这在短句和长句中也都成立，支持了图 1 的第三个观察结果。

图 3 是短句和长句中目标词的基频最小值 (不包括都是一声的句子)。以焦点条件、句长和声调为自变量，对基频最小值做三因素重复测量的方差分析。结果表明，焦点条件在第一个词 ($F(3, 12) = 3.826, n.s.$) 和最后一个词上 ($F(3, 12) = 2.404, n.s.$) 都没有主效应。焦点条件 ($F(3, 12) = 4.506^*$)、句长 ($F(1, 4) = 11.9^*$) 和声调 ($F(3, 12) = 4.43^*$) 在第二个词上都有主效应，并且句长和焦点条件这两个因素在第二个词上存在交互作用 ($F(3, 12) = 7.489^*$)。从图 3 可见，

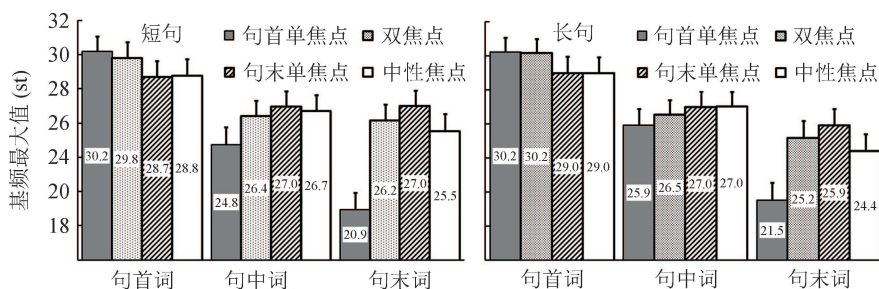


图 2 短句 (左) 和长句 (右) 中 3 个目标词的基频最大值

② * 表示 $p < 0.05$, ** 表示 $p < 0.01$, *** 表示 $p < 0.001$

句首单焦点条件下，第二个词的基频最小值比其它 3 种焦点条件下低。另外，短句中句首单焦点后的基频比长句中下降得多，这与音高最大值的结果是一致的。

对基频的定量分析表明，双焦点句的两个焦点的基频升高程度和相应的单焦点没有显著区别。重要的是，第一个焦点没有焦点后音高压缩，在短句和长句中都是如此。另外，焦点词的基频升高主要是在基频最大值上，而单焦点后的基频下降在基频最大值和最小值上都有体现。

1.2.3 时长分析

图 4 是短句和长句中 3 个目标词在四种焦点条件下的时长。

以焦点条件、句长和声调为自变量，对 3 个目标词的时长做三因素重复测量的方差分析，并进行了后效检验 (S-N-K)。结果表明，在句首单焦点和双焦点条件下，第一个词的时长要比在句末单焦点和中性焦点下更长 ($F(3, 12)=0.565^*$)。第二个词的时长在 4 种焦点条件下没有显著差异 ($F(3, 12)=0.611, n.s.$)。这说明焦点对时长的作用只局限在焦点词上，对焦点后的词的时长没有影响；双焦点条件下第一个焦点词后没有增加韵律边界，也就是说两个焦点可以在同一个语调短语中实现。双焦点和句末单焦点条件下，最后一个词的时长要比句首单焦点和中性焦点条件下更长 ($F(3, 12)=5.554^*$)。

总之，双焦点句中两个焦点的时长延长量与对应的单焦点条件没有明显区别，且焦点前和焦点后的词

没有表现出任何系统的时长变化。双焦点句中，第一个焦点后也没有增加新的韵律边界。

1.3 小结

在汉语普通话双焦点句中，两个焦点在基频升高值和时长延长量上和对应的单焦点条件基本相同。与句首单焦点条件不同的是，双焦点句中的第一个焦点没有焦点后压缩。这和 Jia 等人^[15]及黄贤军等^[16]的结果是一致的。此外，短句和长句中双焦点的韵律实现方式也是相似的，也就是说，在不改变第一个焦点后韵律边界的情况下，第一个焦点后缺少音高压缩不受两个焦点之间音节数量的影响。这也就意味着韵律短语边界很可能限定了第一个焦点的范围。因此，并不是第一个焦点后缺少音高压缩，而是第二个焦点前没有明显的音高变化。在本实验中，第一个焦点的范围就是焦点词自身，其后的词都应该归于第二个焦点前的词。对这一问题还有必要做更系统的实验研究。

2 感知实验 1

由上述分析可知，双焦点和句首单焦点在语调上的区别主要表现为第一个焦点后是否有音高压缩，而双焦点和句末单焦点之间的区别在于第一个词是不是有音高升高。第一个感知实验就是为了调查听者能否很好地感知这些区别。这里最关心的是，双焦点条件下第一个焦点仅有音高升高，是否还可以有效地被感知。另外，还将考查句长和声调是否影响焦点的感知。

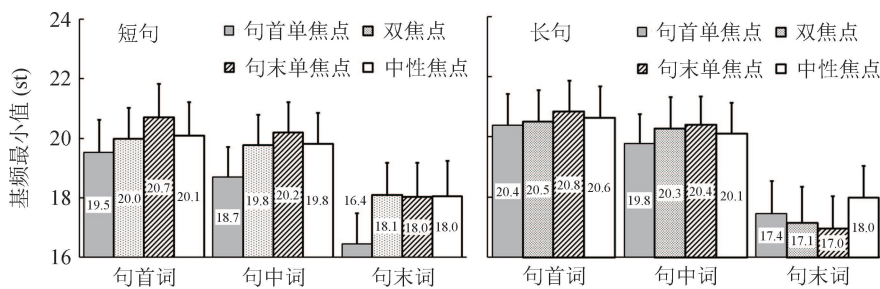


图 3 短句 (左) 和长句 (右) 中 3 个目标词的基频最小值

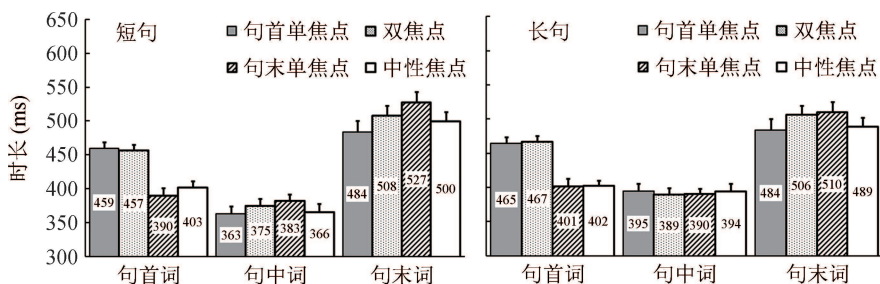


图 4 短句 (左) 和长句 (右) 中 3 个目标词的时长

2.1 研究方法

2.1.1 材料

听辨材料都是来自产出实验。去掉了其中一个男性发音人, 因为相比其他人, 他的句子中有较多明显停顿。他所有声音文件的 10% 都有明显的无声停顿 (平均 84 ms)。进行感知实验时, 所有句子都是随机呈现的。共有 4 (焦点条件) × 2 (句长) × 5 (声调) × 4 (发音人) = 160 个句子。

2.1.2 被试

年龄在 19 到 22 岁之间的 16 名大学生 (8 男 8 女) 参加了感知实验, 没有任何听力和言语障碍。参与实验得到少量报酬。所有听音人均没有参加过双焦点产出实验。

2.1.3 实验程序

实验在中央民族大学语音实验室完成。请被试听到句子后判断句中哪个词被强调了。有 4 个选项: 句首词, 句末词, 句首和句末词以及没有哪个词被强调。感知实验在 Praat 软件中借助 Experiment MFC 脚本文件^[21]进行, 句子播放顺序随机。

在正式实验之前, 听音人先做一个练习。练习材料是从 4 个发音人的录音中选择的 16 个句子。这些句子包含 4 种焦点条件, 并有五种不同的声调组合。练习部分中的声音不再在感知实验中使用。正式实验时, 每个句子只能听一次。

2.2 实验结果

表 1 是焦点感知混淆矩阵。由表 1 可知, 句首单焦点的感知正确率最高 (77.5%), 接下来分别是中性焦点 (49.8%), 句末单焦点 (36.1%) 和双焦点 (30.5%)。尽管双焦点句的感知正确率相对较低, 但仍然大于随机概率水平 (25%)。

表 1 焦点感知混淆矩阵 (%)

| 判断结果 焦点条件 | 句首单焦点 | 句末单焦点 | 双焦点 | 中性焦点 |
|--------------|-------------|-------------|-------------|-------------|
| 句首单焦点 | 77.5 | 2.5 | 4.8 | 15.2 |
| 句末单焦点 | 8.0 | 36.1 | 27.2 | 28.8 |
| 双焦点 | 30.8 | 15.9 | 30.5 | 22.8 |
| 中性焦点 | 16.6 | 15.3 | 18.3 | 49.8 |

以原始焦点条件和感知到的焦点条件为自变量, 做两因素重复测量的方差分析。结果表明, 感知到的焦点条件上有主效应 ($F(3, 45) = 12.042^{***}$)。原始焦点条件和感知到的焦点间有交互作用 ($F(9, 135) = 83.778^{***}$)。简单效应检验结果表明, 句首单焦点被正确感知 (77.5%) 的比例明显高于感知成其它 3 种焦点

条件的概率; 句末单焦点被正确感知的比例 (36.1%) 也显著高于被感知成其它焦点的概率; 双焦点条件被感知为句首单焦点 (30.8%) 和中性焦点 (22.8%) 的比例显著高于句末单焦点 (15.9%); 中性焦点被正确感知的比例 (49.8%) 显著高于被感知成其他焦点的概率。

为了进一步研究影响焦点感知正确率的因素, 将从句长和声调两个方面来分析, 见表 2 和表 3。以焦点条件、声调和句长为自变量, 对各焦点条件的感知正确率做三因素重复测量的方差分析。结果显示, 焦点条件 ($F(3, 45) = 35.456^{***}$) 和声调 ($F(4, 60) = 3.406^*$) 都有主效应, 但是句长没有主效应 ($F(1, 15) = 0.076, n.s.$)。焦点条件和声调 ($F(12, 180) = 5.443^{***}$) 以及焦点条件和句长 ($F(3, 45) = 9.185^{**}$) 有交互作用。简单效应检验表明, 句首单焦点条件下, 短句焦点的感知正确率会高于长句的 (分别是 82.8% 和 72.5%) (见表 2)。其它 3 种焦点条件下都没有发现句长对焦点感知有影响。声调在句首和句末单焦点条件下对焦点感知有影响, 但在双焦点和中性焦点条件下不同声调句间没有显著差异。具体来说, 句首单焦点在 24 调的句中感知正确率最高, 句末单焦点在 31 调的句中感知正确率最低 (见表 3)。

表 2 短句和长句中焦点判断正确率 (%)

| 焦点条件 句长 | 句首单焦点 | 句末单焦点 | 双焦点 | 中性焦点 |
|------------|-------|-------|------|------|
| 短句 | 82.8 | 33.1 | 28.1 | 50.9 |
| 长句 | 72.5 | 39.1 | 32.5 | 48.8 |

表 3 不同声调句中焦点判断正确率 (%)

| 焦点条件 声调 | 句首单焦点 | 句末单焦点 | 双焦点 | 中性焦点 |
|------------|-------|-------|------|------|
| 11 | 75.0 | 43.8 | 25.8 | 56.3 |
| 13 | 75.0 | 39.1 | 22.7 | 46.9 |
| 31 | 81.3 | 25.8 | 18.0 | 56.3 |
| 24 | 84.4 | 39.8 | 48.4 | 42.2 |
| 42 | 72.7 | 32.0 | 36.7 | 47.7 |

2.3 小结

上述分析表明, 双焦点的感知正确率远不如句首单焦点。句长和声调对双焦点的感知没有显著影响。双焦点容易被混淆成句首单焦点和中性焦点。双焦点感知正确率较低, 这可能是由于 4 个选项中有两个单焦点选项 (句首或句末词), 但只有一个双焦点选项, 因此听音人判断为其它焦点条件的概率更高。为了进一步排除这种可能性, 第二个感知实验中

将只用句首单焦点、双焦点和中性焦点条件 3 个选项。这样可以更直接地考察缺少焦点后压缩是否影响对第一个焦点的判断。换句话说, 仅有音高升高和时长延长是否可以保证焦点的有效感知?

3 感知实验 2

第二个感知实验和前一个实验唯一的区别就在于材料中排除了句末单焦点句。由感知实验 1 可知, 双焦点和句末单焦点之间的混淆率 (15.9%) 远低于双焦点和句首单焦点之间的混淆率 (30.8%)。另外, 句末单焦点被判断成双焦点的比例达到 27.7%。实验的预期是当句末单焦点句被排除之后, 双焦点句的焦点感知正确率会提高, 但仍低于句首单焦点。

3.1 实验方法

3.1.1 材料

听辨材料和感知实验 1 相同, 只是去掉了句末单焦点句。总共有 3 (焦点条件) \times 2 (句长) \times 5 (声调) \times 4 (发音人) = 120 个句子。

3.1.2 被试

21 名大学生 (11 男 10 女) 参与了本实验, 年龄在 19~25 岁之间, 来自中国北方, 讲普通话并且没有明显口音。他们都没有参加过感知实验 1 和双焦点产出实验。参与实验得到少量报酬。

3.1.3 实验程序

同感知实验 1。

3.2 结果

由表 4 的感知混淆矩阵可见, 句首单焦点的感知正确率仍明显高于双焦点, 分别为 69.9% 和 41.5%, 双焦点句的感知正确率比随机概率高 8.2%。双焦点混淆成句首单焦点与中性焦点的概率几乎相同 (分别是 30.1% 和 28.3%)。也就是说, 在简化了感知任务后, 双焦点的判断仍然是相对困难的。或者说, 双焦点的实现方式并不能保证其被有效感知。

表 4 焦点感知混淆矩阵 (%)

| 判断结果 焦点条件 | 句首单焦点 | 双焦点 | 中性焦点 |
|--------------|-------|------|------|
| 句首单焦点 | 69.9 | 11.4 | 18.7 |
| 双焦点 | 30.1 | 41.5 | 28.3 |
| 中性焦点 | 14.5 | 28.9 | 56.5 |

为了进一步研究句长和声调对焦点感知的作用, 同感知实验 1 一样, 以句长、声调和焦点条件为自变量, 对焦点的感知正确率做三因素重

复测量的方差分析, 结果显示焦点条件有主效应 ($F(2, 40) = 13.736^*$), 而声调 ($F(4, 80) = 1.546, n.s.$) 和句长 ($F(1, 20) = 1.029, n.s.$) 却没有。另外, 3 种条件两两之间有交互作用 (句长 \times 声调: $F(4, 80) = 4.577^*$; 句长 \times 焦点: $F(2, 40) = 12.765^*$; 声调 \times 焦点: $F(8, 160) = 4.862^*$), 3 种条件间的交互作用也显著 ($F(8, 160) = 1.986^*$)。简单效应检验表明, 句首单焦点和双焦点条件下, 句长的作用是相反的。句首单焦点在短句中的感知正确率要高于长句 (分别为 80.2% 和 59.8%), 但是双焦点在长句中的感知正确率高于短句 (分别为 49.0% 和 34.0%)。另外, 双焦点在声调组合为 42 和 24 句中 (二者的平均正确率为 51.5%) 的感知正确率要高于 11、13 和 31 句 (三者的平均正确率为 34.8%)。由此可见, 动态声调 (dynamic tone) 相比于静态声调 (static tone) 更有利于双焦点的感知。第一个感知实验中也有同样的趋势 (见表 3)。

3.3 小结

感知实验 2 得到了与感知实验 1 相似的结果。在这两个实验中, 双焦点的感知正确率与中性焦点大致相当, 明显比句首单焦点低。另外, 双焦点被混淆成句首单焦点和中性焦点的概率基本相同。感知实验 2 的新发现是, 在简化了实验任务后, 句长和声调都会影响双焦点的感知。双焦点在长句中的感知正确率要高于短句, 动态调 (2 声和 4 声) 的双焦点要比静态调 (1 声和 3 声) 更容易感知。另外, 句首单焦点在短句中的感知正确率要高于长句。这两个感知实验中, 双焦点感知较差还有一个可能的原因是, 第二个焦点的感知很差, 所以整体上正确率偏低。针对这种可能性, 我们做了第三个感知实验。

4 感知实验 3

实验 3 中主要考察缺少焦点后压缩是否会影响第一个焦点的感知。实验任务为判断句首词是否被强调。也就是让听音人只关注句中一个焦点。如果句首单焦点和双焦点句的判断有明显差别, 则表明双焦点感知差是第一个焦点后缺少音高压缩造成的。如果句首单焦点和双焦点句的判断结果没有明显区别, 则表明缺少焦点后压缩并不影响第一个焦点的感知。由此可以推断双焦点句感知差的原因应该是由于处于句末位置的第二个焦点比较难判断。

4.1 实验方法

4.1.1 材料

将实验 2 的听辨材料分成两部分, 一部分是句首单焦点和中性焦点句, 总共有 2 (焦点条件) \times 2 (句

长)×5(声调)×4(发音人)=80个句子;另一部分是双焦点和中性焦点句,同样也是80个句子。

4.1.2 被试

30名大学生(7男23女)参与了本实验,年龄在19~25岁之间,均来自中国北方,讲普通话并且没有明显口音,都没有参加过前面的实验。他们被随机分为两组,每组被试听其中一组材料,即句首单焦点和中性焦点句,或是双焦点和中性焦点句。参与实验得到少量报酬。

4.1.3 实验程序

除了要求被试判断句首词是否被强调外,其它同实验2。

4.2 结果

表5为各焦点条件下,听音人判断为句首词被强调的比例。因为中性焦点条件在两组实验材料中都有,且判断为句首有强调的比例大致相当(在句首单焦点和中性焦点组为36%,在双焦点和中性焦点组为33.8%),这里将两组中的结果进行平均。由表5可见,句首词被判断为强调的情况在句首单焦点条件下比例最高,其次是双焦点,中性焦点最低。但是,以焦点条件为变量的单因素方差分析结果显示,句首单焦点和双焦点条件下句首词被判断为强调的比例并没有显著差异($F(1,28)=2.732, n.s.$)。对第一组材料的重复测量方差分析结果显示,句首单焦点条件比中性焦点下将第一个词判断为强调的比例差异显著($F(1,14)=91.249^{***}$)。第二组材料得到了相同的统计结果($F(1,14)=61.067^{***}$)。

也就是说,双焦点句中的第一个焦点在缺少焦点后压缩的情况下仍能被较好地感知。由此推测,前两个感知实验中双焦点句感知较差很可能是由于第二个焦点不容易识别而被误判为句首单焦点或是中性焦点。这也解释了为什么第一个感知实验中,双焦点句较少被误判为句末单焦点,因为双焦点句的第一个焦点还是可以很好地被感知到的。

表5 各焦点条件下句首词被判断为强调的百分比(%)

| | 句首词强调 |
|-------|-------|
| 句首单焦点 | 75.0 |
| 双焦点 | 68.1 |
| 中性焦点 | 34.9 |

5 讨论

本文研究了汉语普通话中双焦点的产出以及感知。双焦点句的韵律形态是两个焦点都有音高升高

和时长延长,且基本达到了所对应的单焦点水平。新的发现是,第一个焦点后韵律短语边界的设置限定了两个焦点各自的实现范围。在不改变第一个焦点后韵律边界的情况下,增加两个焦点间的音节数不改变两个焦点的实现方式。另外,第一个焦点后也没有增加语调短语边界,可见汉语普通话中允许一个语调短语里有两个重音。最后,双焦点的感知正确率整体较低,这主要是由于句末位置的第二个焦点不容易被感知,第一个焦点相对更突显。下面将分别对双焦点的产出和感知进行讨论,以回答引言中提出的两个问题。

(1) 汉语普通话中双焦点的韵律形态是怎样的?是否受两个焦点间音节数的影响?

焦点产出实验表明,汉语普通话双焦点句中的两个焦点词基频和时长的增加量都与对应的单焦点条件没有明显区别。重点是,双焦点句中第一个焦点后没有音高压缩。这些结果与Jia等人^[15]及黄贤军等^[16]的结果一致,也与Eady等人^[4]和Liu^[12]对英语的研究结果一致。Eady等人^[4]指出,第二个焦点对第一个焦点的实现有逆向的影响。但是并没有解释清楚产生影响的根本原因是什么。其它几项研究也缺少对这一现象的解释。通过本实验,我们认为,韵律短语边界限定了第一个焦点的实现范围,这与黄贤军等^[16]的发现是一致的。并且,韵律短语边界的设定可能是基于句法的。

本实验中用到的是主谓结构句(“汪英参观车间”),其中主语就是一个名词。在这样的句子中,主语后是一个韵律短语边界。虽然长句中(“汪英参观山西清乡一街车间”),动词短语非常长,但是第一个韵律短语是主语,而不会是“汪英参观”。可见,句法结构对韵律边界还是有很强的限制作用的。Selkirk^[28]提出的韵律—句法匹配理论(Match Theory)有更详细的讨论。从双焦点的语调曲线(图1—图4)看起来,主语成为焦点时,其范围就是它本身,后面的动词短语构成另外一个韵律短语。Wang等人^[11]的研究发现,韵律边界对边界前音节时长的延长作用和对焦点的延长作用是独立的。由此可以判断,第一个焦点后并没有增加语调短语边界,两个焦点是在一个语调短语中实现的。主语后没有焦点后音高压缩的现象,这是因为主语后的动词应该是第二个焦点的焦点前成分。而焦点前的音高一般没有明显变化^[7,11]。也就是说,并非是第一个焦点后缺少音高压缩,而是第二个焦点前音高没有明显变化。引言部分介绍的Yuan等^[17]的双焦点实验发现,第一个焦点后也是可以出现音高压缩的。在他们的实验句中,“张咪扔给窗

边的八哥三根香蕉”，第二个焦点如果在第二个宾语上，动词上就有焦点后音高压缩。我们对该结果的解释如下：这个句子的韵律短语边界与句法短语对齐：“(张咪) φ ^③(扔给) φ (窗边的八哥) φ (三根香蕉) φ ”，而韵律—句法匹配理论^[35]允许韵律成分有递归，也就是韵律短语可以递归到更高一级的韵律短语。该句在双焦点条件下的韵律结构为：“((张咪)(扔给) φ (窗边的八哥) φ) φ (三根香蕉) φ ”。也就是说第一个焦点的范围可以是递归后的最高一级的韵律短语。第一个焦点范围的确定与句法边界的关系还需要更系统的实验研究。

在 Kabagema-Bilan 等^[18]的研究中，双焦点句(句首单焦点 + 句中焦点)的两个焦点上都没有表现出明显的基频升高。这是否因为句子太短(6~8个音节)，且两个焦点之间只有一个音节？与他们的句子相比，本研究中的短句有6个音节，两个焦点之间有两个音节，但两个焦点上都有基频的升高。这两项研究在实验控制上有很多不同，因此很难判断产生差异的原因。至少可以认为，句长可能并不直接影响双焦点的实现方式。

最后，在时长变化上，第一个焦点的时长延长和句首单焦点没有明显区别，且在短句和长句中结果一致；然而，第二个焦点的时长只在短句中有延长，在长句中则没有。这可能是由于句末音节延长的效应是有上限的^[11,29]。也就是说，句末音节成为焦点时，其时长的延长就变得比较有限了。

(2) 双焦点的感知正确率如何？影响双焦点感知的因素是什么？

前两个感知实验的结果都表明双焦点的感知正确率远低于句首单焦点，仅略高于随机水平，而句首单焦点的感知正确率在70%以上。这有两种可能性，第一种可能性是双焦点句中的第一个焦点不容易判断。双焦点句中第一个焦点后没有音高压缩，可能使得第一个焦点不容易确定。Xu 等人^[20]已经注意到焦点后的压缩对焦点感知的重要作用。这里有两个相关发现。第一，句末焦点远没有非句末焦点感知正确率高，这很可能是由于缺少焦点后压缩，因为句末焦点后已经没有让基频再下降的空间了，正如 Xu 等人^[20]以及 Liu 和 Xu^[21]的研究所发现的，非句末位置的单焦点感知正确率很高(大于90%)，而句末焦点容易和中性焦点混淆。第二，没有焦点后压缩的语言比有焦点后压缩的语言表现出更低的焦点感知正确率。尽管北京话和台湾闽南话同属一种语言，但焦点在韵律上也会有不同的编码方式，北京话

有焦点后压缩，而台湾闽南话中却没有；台湾闽南话中句末和非句末位置的焦点感知正确率都很低(大约60%)^[20]。

第二种可能性是，双焦点句中的第二个焦点不容易判断。这可能是由于句末焦点本身更难被正确感知。Yuan 等人^[17]的双焦点感知实验同样发现，第一个焦点被感知为句重音的比例明显高于第二个焦点。他们认为重音是分层级的，句首焦点为主要重音，而第二个焦点为次要重音。

本文的第三个感知实验支持了第二种可能，与 Yuan 等人^[17]的结果相似。在第三个感知实验中要求听音人单独判断双焦点句的第一个焦点，其正确率与单焦点条件没有显著区别。也就是说，仅有音高升高和时长延长也能保证焦点的正确感知。焦点后压缩并不是判断焦点的一个必要条件。正如前面提出的，焦点后压缩的主要功能很可能是预示焦点后是否还有重要的信息出现。

双焦点的判断也可能与注意的分配有关。单焦点会引起人们对句中某一信息的特别关注，在言语理解中对注意有调节作用。但是，双焦点引起的可能是对整个句子信息的关注，而不仅是两处信息。关于双焦点感知的基本认知机制还需要更多的实验研究。

6 结论

本文通过控制焦点词的位置和句子长度两个因素，发现句中两个焦点既相互独立又相互依存。独立性表现为每个焦点有各自的实现范围，其音高升高和时长延长都和单焦点的情形十分接近；依存性表现为两个焦点可以在同一个语调短语内共存。第一个焦点后缺少音高压缩主要与句法边界设置有关，而与两个焦点间的音节数没有直接关系。为了进一步考察两个焦点在感知上是否有相同的突显程度，进行了3个感知实验。结果表明：相比于句首单焦点，双焦点的辨认真确率非常低(30.5%)，主要因为同时判别两个焦点难度比较高。如果只要求听音人关注第一个焦点，其辨认率(61.3%)有明显提高，且与句首单焦点的判断无显著差异。可见，在双焦点句中，句首和句末两个位置的焦点判断有不对称性。对于句首位置的焦点而言，焦点后音高压缩虽然在一定程度上对焦点的判断有帮助，但音高和时长的增加仍是焦点感知的主要声学线索。而对于句末位置的焦点，仅有音高上升和时长延长并不能保证正确的焦点感知。

③ φ 表示韵律短语边界

致谢

感谢许毅和李晓庆以及两位审稿专家为本文提供的宝贵建议。

参 考 文 献

- 1 Krifka M. Basic notions of information structure. *Acta Linguistica Hungarica*, 2008; **55**: 243—276
- 2 贾媛, 李爱军, 马秋武, 熊子瑜. 具有焦点标记作用的乞字句重音分布研究. *中文信息学报*, 2009; **23**: 103—109
- 3 Cooper W E, Eady S J, Mueller P R. Acoustical aspects of contrastive stress in question-answer contexts. *J. Acoust. Soc. Am.*, 1985; **77**: 2142—2156
- 4 Eady S J, Cooper W E, Klouda G V, Mueller P R, Lotts D W. Acoustical characterization of sentential focus: narrow vs. broad and single vs dual focus environments. *Language and Speech*, 1986; **29**: 233—250
- 5 Féry C, Kügler F. Pitch accent scaling on given, new and focused constituents in German. *Journal of Phonetics*, 2008; **36**: 680—703
- 6 Wang B, Xu Y. Differential prosodic encoding of topic and focus in sentence-initial position in Mandarin Chinese. *Journal of Phonetics*, 2011; **39**: 595—611
- 7 Xu Y. Effects of tone and focus on the formation and alignment of f0 contours. *Journal of Phonetics*, 1999; **27**: 55—105
- 8 沈炯. 汉语语调构造和语调类型. *方言*, 1994; **3**: 221—228
- 9 Chen Y Y, Gussenhoven C. Emphasis and tonal implementation in Standard Chinese. *Journal of Phonetics*, 2008; **36**: 724—746
- 10 Xu Y, Xu C X. Phonetic realization of focus in English declarative intonation. *Journal of Phonetics*, 2005; **33**: 159—197
- 11 Wang B, Xu Y, Ding Q F. Interactive prosodic marking of focus, boundary and newness in Mandarin. *Phonetica*, 2017; **75**: 24—56
- 12 Liu F. Single vs. double focus in English statements and yes/no questions. In: Proceedings of the Speech Prosody 2010, Chicago (USA), 2010: **100975**: 1—4
- 13 Wang B, Féry C. Prosody of dual focus in German: Interaction between focus and phrasing. *Language and Speech*, 2017; **61**(2): 303—333
- 14 Wang B, Féry C. Dual-focus intonation in Standard Chinese. In: Proceedings of the 18th Oriental COCODA/CASLRE Conference, Shanghai, 2015: 101—106
- 15 Jia Y, Li A J, Xiong Z Y. A phonetic and phonological analysis of dual and multiple focuses. In: Proceedings of the Speech Prosody 2010, Chicago, USA, 2010: 75—78
- 16 黄贤军, 郑海洋, 吕士楠, 杨锦陈. 韵律短语边界对降阶和焦点后音高骤降的影响. *声学学报*, 2016; **41**(4): 529—536
- 17 Yuan Y, Li A J, Yuan J, Hu J, Surány B Z. Phonetic realizations of post-nuclear accent under dual-focus conditions in Standard Chinese. In: Proceedings of the Speech Prosody 2016, Boston, USA, 2016: 941—945
- 18 Kabagema-Bilan E, López-Jiménez B, Truckenbrodt H. Multiple focus in Mandarin Chinese. *Lingua*, 2011; **121**: 1890—1913
- 19 Hyman L M. Word-prosodic typology. *Phonology*, 2006; **23**: 225—257
- 20 Xu Y, Chen S W, Wang B. Prosodic focus with and without post-focus compression (PFC): A typological divide within the same language family? *The Linguistic Review*, 2012; **29**: 131—147
- 21 Liu F, Xu Y. Parallel encoding of focus and interrogative meaning in Mandarin intonation. *Phonetica*, 2005; **62**: 70—87
- 22 Herment-Dujardin S, Hirst D. Emphasis in English: A perceptual study based on modified synthetic speech. In: Proceedings of the Speech Prosody 2002, Aix-en-Provence, France, 2002: 379—382
- 23 Rump H, Hermes D J. Prominence lent by rising and falling pitch movements: Testing two models. *J. Acoust. Soc. Am.*, 1996; **100**: 1122—1131
- 24 Heldner M. Is an F0-rise a necessary or a sufficient cue to perceived focus in Swedish? In: Werner S eds. *Nordic Prosody: Proceedings of the VII Conference*, Peter Lang: 1996: 109—125
- 25 Rump H H, Collier R. Focus conditions and the prominence of pitch-accented syllables. *Language and Speech*, 1996; **39**: 1—17
- 26 White S. <http://qaparadigm.sourceforge.net/2007>, 2007
- 27 Xu Y. ProsodyPro — A tool for large-scale systematic prosody analysis. In: Proceedings of tools and resources for the analysis of speech prosody (TRASP 2013). Aix-en-Provence, France, 2013: 7—10
- 28 Selkirk E. The syntax-phonology interface. In: Goldsmith J A, Riggle J, Yu A C L eds. *The handbook of phonological theory* (2nd ed.), Oxford: Wiley-Blackwell: 2011: 435—484
- 29 王蓓, 杨玉芳, 吕士楠. 汉语语句韵律层级结构边界的声学分析. *声学学报*, 2004; **29**(1): 29—36