

# 汉语韵律结构和功能语调

林茂灿

中国社会科学院语言所

[linmaocan@263.net](mailto:linmaocan@263.net)

## 摘要

本研究认为，汉语普通话功能语调有两个特征，一是重音，二是疑问信息。我们从确定语句的韵律结构入手研究重音和疑问信息。韵律结构在这里指韵律词和韵律短语。韵律词是一个  $F_0$  变化组，韵律词内部各音节  $F_0$  有特定表现，其中尖锐有力的一两个音节就是韵律词的重音，韵律词与韵律词之间还会发生  $F_0$  重设，这些是韵律词在口语韵律方面形成的十分重要特点，对口语的自然悦耳起了很大作用，而且也是语调研究不可缺少的切入点。韵律短语有以无声停顿为边界的，也有以有声停顿为边界的。韵律短语有大小之分。重音具有层次性。我们从韵律词重音出发构建  $F_0$  高低线， $F_0$  高低线也是有层次性的。

对一般是非问句（不带“吗”是非问句）来说，疑问信息（question information）由语调短语（韵律大短语）末了韵律词中最后一（两）个音节所携带，疑问信息有时也可由其开头一（两）个音节携带。因而，是非问句的疑问信息称为边界调（boundary tone）。我们用实验说明韵律大短语是语调短语，因为它的边界处携带功能句调信息。本实验结果表明，陈述句末了音节不论什么声调，如果它不被强调的话，其  $F_0$  曲线大多是下降的，而携带疑问信息音节的  $F_0$  曲拱象单念时的那样；如果其时长短于 200 毫秒（平均），其  $F_0$  曲拱位置要比陈述的抬高约一度半（声调五度值）；疑问信息音节的时长如果较长，阴平，阳平和去声  $F_0$  曲拱位置不抬高但其终点抬高，上声  $F_0$  曲拱整个抬高；携带疑问信息音节如果  $F_0$  曲拱位置抬高，而且阴平，阳平和去声的终点位置也抬高，时长也长，其疑问意味就很浓。至于最后第二音节的  $F_0$  曲线位置的抬高，如果它跟疑问信息无关的话，它是由  $F_0$  协同发音引起。

本研究看到，语句重音是以韵律词为基本节奏单元按层次分布的。由于语句重音的层次性，据此所构建的  $F_0$  高低线也有三个层次，它们是韵律词的  $F_0$  高点和低点（或  $F_0$  高低线），韵律短语的  $F_0$  高低线，以及全句的  $F_0$  高低线。韵律短语和全句的  $F_0$  高低线呈现出三种类型：前重（SW），后重（WS）和中重（WSW）。在三种类型的韵律短语  $F_0$  高低线中， $F_0$  底线总是下降的，而  $F_0$  高线跟重音密切相关：前重型的  $F_0$  高线从第一个韵律词逐步下降，后重型的  $F_0$  高

线从第一个韵律词逐步上升，而中重型的  $F_0$  高线从第一个韵律词上升到第二个韵律词重读音节处为最大，然后向最后韵律词逐步下降。在各个韵律层次中， $F_0$  位置越高（上声的  $F_0$  位置越低）和/或  $F_0$  音域越大的音节或音节组，给人以更加突显、更重（正常重音）的感觉。语调短语末了韵律词有时出现这样的情况：一个音节  $F_0$  值大但时长很短，而邻近音节  $F_0$  值虽小但时长较长， $F_0$  值虽小但时长较长的音节却给人以较重的感觉。总的说来，音节  $F_0$  曲拱（包括时长）越接近单字调的，就给人以越重的感觉。

关于句调与声调之间关系问题，我们从句调的边界调和重音两个特征来讨论。从是非问的边界调声学表现看到，句调有个特征加到语调短语携带疑问信息的音节上，这个特征是携带疑问信息音节的  $F_0$  曲拱不仅象单念时的那样，而且阴平，阳平和去声的终点位置（音阶）和 / 或  $F_0$  曲拱整个位置（音阶）比陈述的约抬高，和 / 或时长也长。再从语调短语重音的声学表现看，音节念得越重，其音节  $F_0$  曲拱（包括时长）就越接近单字调，而在  $F_0$  五度值空间，普通话非重读阴平，阳平和去声的  $F_0$  曲线位置有处于‘4’，‘3’，也有‘2’甚至‘1’的，上声  $F_0$  曲线位置都很低，非重读的多处于‘2’左右。至于非重音节  $F_0$  曲线方向，阴平的可以上升也可以下降，阳平也可以下降，上声有略升的也有略降的，去声可以略升。语句中重读音节发音时用力足，声带足够紧张，使得  $F_0$  曲拱饱满，而非重读音节发音时用力相对不足，声带相对不够紧张，加上受连续变调和协同发音的影响和发音生理上的限制，使得  $F_0$  曲拱偏离了‘预定’（本调）位置和方向。语调短语主要受重音的作用，加上连续变调和协同发音的影响和发音生理上的限制，使得各个音节  $F_0$  有了不同表现。

通过合成试验证明，以上关于边界调和重音声学表现研究结果是符合语音实际的。

我们认为，句调对声调的作用是非线性的，这是因为重音是层次性的，边界调与  $F_0$  曲拱及时长关系也不是一一对一的。

以上是我们对普通话句调的初步看法，写出来跟大家讨论。对 973 电话对话语料和自编的对话语料，还在继续分析和观察，期望这个研究能更深入，能得到可靠定量的结果。

## 1. 汉语韵律结构

这儿先谈为什么和如何研究韵律结构，并介绍本研究所用的语料和方法。

### 1.1 研究目的和所用语料

Ladd (1996)指出，语调的调子(tune)和相对突显(relative prominence)特征在语句中的分布受其韵律结构制约。这是根据非线性音系学理论，总结 Janet Pierrehumbert 1980 年发表《英语语调的音系学和语音学》(The Phonology and Phonetics of English Intonation)以来语调研究取得的进展，而提出来的语调理论观点。研究韵律结构必须涉及话语不同等级单元的边界信号，这儿研究其组成成分---韵律词和韵律短语的边界信号。确定话语不同等级单元的边界信号，我们从研究话语的音联或间断入手。

Trager 和 Bloch (1941) 首先提出音联(juncture)这个概念，并对英语音联作了分析和研究。Lehiste 和 Ann Arbor (1960)用语图仪专门研究了英语内部开音联(internal open juncture)的声学语音学表现。用声学语音学方法研究普通话语音联是从许毅(1986)开始的。Hockett (1947)也谈到了北京话的几种音联。赵元任(1980)把 juncture 翻译为“断续度”，他说“juncture 意思是‘连’，不过说 juncture 这个字，也可叫 disjuncture，因为它是说一段跟一段连的程度是松是紧”。他举例“‘a nice man’跟‘an ice man’平常说话的时候，虽然第一个 n 跟后头的 i 总是连来说，可是还有分别。这种差别怎么分呐？现在一般语言学者的分法就是说：有一个 juncture，叫做 plus juncture（因为用一个加号‘+’），‘a nice man’就是<e+nais+mæn>，‘an ice man’就是<en+ais+mæn>。那个里头可以算是一个分号，也可以算是个连号，就是在那个地方比较断一断。”赵元任指出，“这不是文字拼写的问题，是说起来的确实有的分别，”。从此看到，人们听到话语中不同大小的间断(break)，对应于话语中的音联。

人们说话时使用停顿和音高等韵律参数，使话语产生可以感觉到的间断；人们在听别人说话时则通过这些韵律特征来决定语流的间断(break, disjuncture)程度(Sanderman and Collier, 1996)。间断使话语具有韵律结构(prosodic structure)。话语的韵律结构知识，在言语合成上很有用。合成的言语声清晰度可以满足当前的要求了，可是其自然度不很理想，人们不愿意多听合成的言语声，多听了就厌烦。Boogaart 和 Silverman (1992)指出，在言语合成系统中，如果在韵律上运用得当，其合成的言语声就自然，听起来悦耳。Pijper 和 Sanderman (1994)指出，自然悦耳的合成言语声会减少听音人的认知负担，从而达到更好地理解话语的目的。

Pijper 和 Sanderman (1994)又指出，在言语识别系统中，恰当地运用话语的韵律信息，有助于提高语言学知识在系统中的使用效果。话语的韵律信息在汉语言语工程中得到

了运用，汉语韵律结构的研究也愈来愈深入。初敏和吕士楠(1996)合成普通话时，在短语和句子边界作适当延长，以逼近自然话语的时长分布；她们还考虑了实际语音中  $F_0$  表现，其合成音质提高了。Lee, Tseng and Ouh-Yong (1989)合成汉语国语时，在话语中加入四个等级的间断，使其音质大有改善。扬玉芳(Yang, 1997)看到，词、词组、短语、小句和句子的边界之前音节在时域和频域上有系统差别。李爱军(Li, 1998)从知觉实验出发看到，自然念的普通话语句有大短语(major phrase)和小短语(minor phrase)之分；大短语末音节的时长，跟英语一样比小短语末音节的长，不过如果考虑声调，其情况就复杂了。

Pijper & Sanderman (1994)提出，做韵律研究用的语料至少应符合以下两个要求：1) 发音人事先应不知道他/她说的话是用于研究韵律的，这样，他/她在念话语时不会故意强调某种韵律现象（如，轻重，停顿，速度，等等），使得所念的话语在韵律方面十分丰富而且自然。2) 所用句子涵盖所研究语言的主要句型，使之尽可能出现不同类型和不同强度的韵律短语（以及各种韵律词）。语言学家(如 Price et al, 1991)指出，韵律短语跟句法短语以及韵律词跟句法词并不总是相同的。

为了准确地得到韵律短语边界（以及韵律词边界）的位置，试验用的话语音段信息一定要去除掉（Lehiste and Wang, 1977）。去除话语中的音段内容有各种方法。Lehiste and Wang 运用频谱倒置技术去除话语中音段内容，只保留其韵律信息；还有用低通滤波去除话语共振峰，只保留其中一两个谐波。Pijper and Sanderman (1994)的试验结果表明，由毫无语言学知识且没受过语言学试验训练的人所判断的正常话语的韵律边界，没有受到词汇，句法及语义等的影响。

“普通话自然口语的韵律标注研究”（李爱军，2002）一文谈到，为了观察受过韵律标注训练人员能否正确地标注韵律，她作了两个试验。一个将所研究语段进行低通滤波处理（截止频率 500HZ），将高频成分滤出，只保留语音的基频部分。经这种滤波处理以后的声音人们不容易听不出来语段本来的意义，而声调、语调和时长、重音等信息会完整保留。这种经滤波的无意义语段的间断听辩试验，叫做无意义间断试验。另一个试验叫做有意义语段的间断听辩，它是听音人直接听所研究句子。同一语料由较长时间参加韵律标注的 10 位年轻人做这两个试验，其结果只有一位听音人对相同语段在有意义和无意义两种情况下，听辩结果相关程度较小( $r=0.733$ )，而其他 9 位的听辩结果在两种情况下的相关性很高。这说明，受过韵律标注训练人如认真标注，其结果能反映实际语流存在的间断，这样的结果可用于研究句法和韵律之间关系。这也说明我们的韵律标注结果总的来说是正确可靠的。

研究韵律短语和韵律词，我们用了两个语料，作了两个试验。两个语料都取自一语料库 (Zu, 1999)，符合 Pijper & Sanderman (1994) 提出的作韵律研究的两个要求。这些句子短的包含两个语法词，最长的有九个语法词。一个语料是发音人 F00 念的 44 个语句 (utterance)，另一个语料是三位发音人念的 49 句 (M00 和 M01 各 21 句，F00 的 17 句)。发音人 F00 念的 44 个语句，用于第一个试验，其目的是为确定韵律短语的边界特点。第二个试验用三位发音人念的 49 句，让听音人判断韵律词的边界。

本研究请没有语言学知识的人参加听辨试验，由她们判断语句中的韵律词和韵律短语边界。她们没有任何听力上的毛病。虽然她们没有语言学专门知识，但她们至少都具有中学文化水平，都用过词典。为了最大限度地减少听音人把词汇知识带入听辨判断，以达到比较准确划分韵律词的目的，听音人事先不知道我们做的研究跟词汇有关，我们要求听音人作大小间断的判断。

参加第一个试验的听音人共 52 位，香港的 17 位，北京的 35 位，他们的任务是确定语句的大小间断，给出语句中韵律大短语和韵律小短语的边界位置。参加第二个听辨试验是 20 位北京年轻人，其中有三分之一参加过第一个试验。听音人在第二个试验中要确定的是语句里哪些音节是连在一起的。这两个试验相隔一年半。听音人对本研究目的一无所知。从间断 (break) 观点看，零级间断指话语中的“附着群” (Beckman & Ayers, 1994)，它里头词与词之间没有任何间断，从音联学说看，闭音联 (close juncture) 指话语内部音段与音段之间的紧密过渡。把紧密连在一起的几个音节的语音片段称为韵律词，跟零级间断和闭音联定义相一致。

## 1.2 语句 $F_0$ 规整和时长规整

本文对我们前两年用的语句  $F_0$  规整公式 (Lin Maocan, 1999; 林茂灿 2000) 作了修改，现用下面公式：

$$J=5 \times (F_0 - F_{0 \min}) \div (F_{0 \max} - F_{0 \min})$$

公式中  $F_{0 \max}$  代表某个发音人某句话的  $F_0$  最大值， $F_{0 \min}$  代表同一个发音人同一句话的  $F_0$  最小值， $F_0$  是待计算的基频值。本文  $F_0$  是对某发音人某句话进行规整的。图 1-3 中各小图的上部为  $F_0$  规整曲线。

把音高语音与  $F_0$  音域之间关系看作是线性的而不是对数的，因为 E. Zwicker 和 R. Feldkeller (1967, 转引自林达悃 (1985)) 对主观音高感受与频率之间关系作了进一步研究，发现以美为单位的 Z 音高，在 500 赫兹以下与频率呈线性关系，这是修改的理由之一。修改的理由之二是，看来仍用五度值度量  $F_0$  与语音音高的相对关系是合适的。

用下面的公式，计算音节时长 T 的规整值：

$$t = (t - t_{ave}) / sd \cdot \dots (7 \cdot 5)$$

式中， $t_{ave}$  代表某人所念语句 (或其他语音单元) 中所有

音节的韵母时长平均值，sd 为其标准差， $t$  代表待规整某音节韵母的时长值， $t$  代表这个音节韵母的时长规整值。本文各音节韵母的规整，是对某发音人念所有音节韵母进行的，以便对不同句子进行比较。图 1-3 各小图的下部为 T 规整，其值大于零的画在横轴的上面，横轴之上的竖线越长，表示规整值越大，即时长越长；小于零的的规整值画在横轴的下面，横轴之下的竖线越长，表示规整值越小，即时长越短；横轴之上的时长，当然大于横轴之下的。规整得到的音节时长 T 表示了各音节时长的相对长度，也就是说，规整得到的音节时长 T 体现了说话人念各个语句对音节时长做的相对分配 (timing)。

把音节时长的相对分配 (relative timing) 跟  $F_0$  曲线画在同一个图上如图 1-3 那样，便于观察和研究语句中声调  $F_0$  变化与音节时长 T 的关系。语音韵律现象不仅跟基频  $F_0$  和时长 T 有关，有时振幅 (A) 还起一定作用，有些情况还要考虑共振峰。关于语句振幅 (A)，将在以后研究中给出。

## 2. 韵律短语和韵律词

### 2.1 韵律短语 (Lin, 1999; 林茂灿, 2000)

我们用孤立语句语料库的语句，通过声学分析和知觉试验看到，普通话语句 (utterance) 中有两种人们可以感觉到间断：无声波间断 (break with silent pause) 和有声波间断 (break with filled pause)。无声波间断是由停顿 (pause) 引起的；有声波间断不是由停顿引起的，而是由间断之前音节时长拉长等所致。无声间断和有声间断之后音节  $F_0$  相对于其前面音节都发生了程度不同的跃变。无声波间断有大小之分，有声波间断也有大小之分。无声停顿长度小于 100ms 时，往往给人以小的无声间断感觉。把大的有声波间断与小的有声波间断区分开不是主要靠其间断前后音节的时长及  $F_0$  跃变，而是主要靠其  $F_0$  范围。韵律短语有大小之分。韵律大短语主要由大的无声停顿及边界前音节市场拉长引起的，韵律小短语主要由小的有声停顿和 / 或无声停顿引起。

### 2.2 韵律词

普通话韵律词有简单韵律词，复杂韵律词和复合韵律词等三种类型。本研究认为，词汇词是通过简单韵律词，复杂韵律词和复合韵律词组成韵律短语以至语句的。韵律词在韵律组织中是至关重要的一种韵律成分。

#### 2.2.1 韵律词内部的 $F_0$ 表现

简单韵律词：简单韵律词指两音节和三音节等的词汇词，以及强调念的单音节词，如副词和介词等。听音人都会 (100%) 认为它们所包含的音节是紧密地连在一起念出来的。人名，地名，专有名词及外来词等作简单韵律词处理。赵元任先生在《语言问题》(1980) 第四讲“词汇跟语法”指出，停顿不能发生于词 (word) 内。石基琳 (Shi, 1986,

p.136)也谈到,词汇完整性(lexical integrity)是语言学的一般原则,没有任何规则能作用于词汇项的内部结构。本研究结果也表明,人们说话时不会破坏或改变词汇信息。简单韵律词的 $F_0$ 表现是:1)阴平后面音节的 $F_0$ 往往下降;阳平后面音节的 $F_0$ 位置多数下降,也有抬高的;“上声”后面音节的 $F_0$ 位置总比其前面的高,也就是说,上声之后音节的 $F_0$ 不会立即下降;去声后面音节的 $F_0$ 位置有的下降,有的抬高。

复杂韵律词和复合韵律词:参加本试验20位听音人中,90%以上听音人认为是紧密连在一起念的那些音节组,定为复杂韵律词;60%的听音人认为是紧密连在一起念的音节组,定为复合韵律词。复合韵律词和复杂韵律词在声学性质上有不同的表现,从而引起听音人对它们有不同的反应。复杂韵律词由两个以上词汇词组成。1)助词“的”、“着”和“了”等等,往往跟其前面的词连在一起念。2)副词“也”和“不”(表示否定)等,如果轻读,往往跟其后面的词连在一起读成复杂韵律词。3)两个以上多音节组(如两个双音节组)紧密连在一起念时,组成复杂韵律词。复杂韵律词的特点:1)一个词汇词 $F_0$ 音域比另一个词汇词的小,而且一个词汇词 $F_0$ 位置比另一个词汇词的低。2)有一个音节时长较长,其它音节往往较短。3)音节间既可能出现连续变调,也可能发生 $F_0$ 协同发音。

由于简单韵律词和复杂韵律词内部 $F_0$ 的这种声学表现,所以称它为“ $F_0$ 变化组”。

复合韵律词的特点:多数是由两个以上简单韵律词和/或复杂韵律词组成,它的一部分 $F_0$ 音域比其另一部分的小,但其 $F_0$ 位置不比另一部分的低,它们之间还可以发生某种 $F_0$ 重设,这种 $F_0$ 重设往往是就后一个成分相对其前面的末了音节而言的。

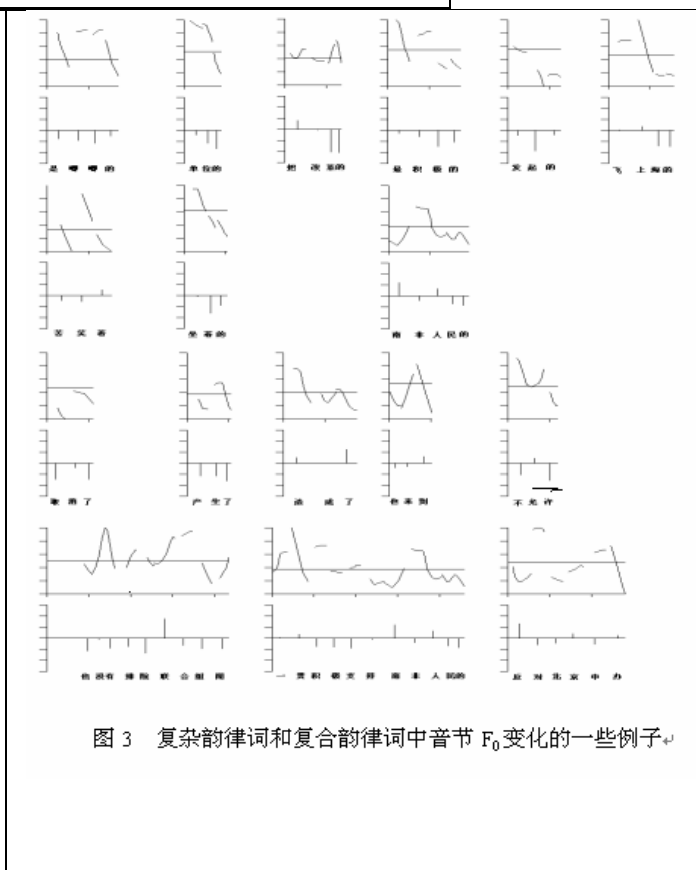
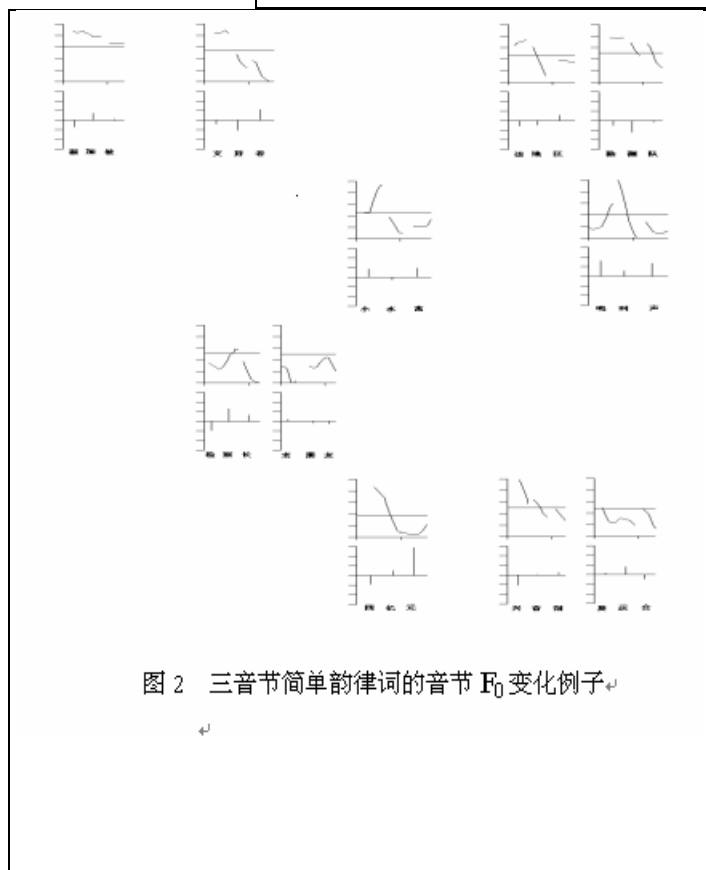
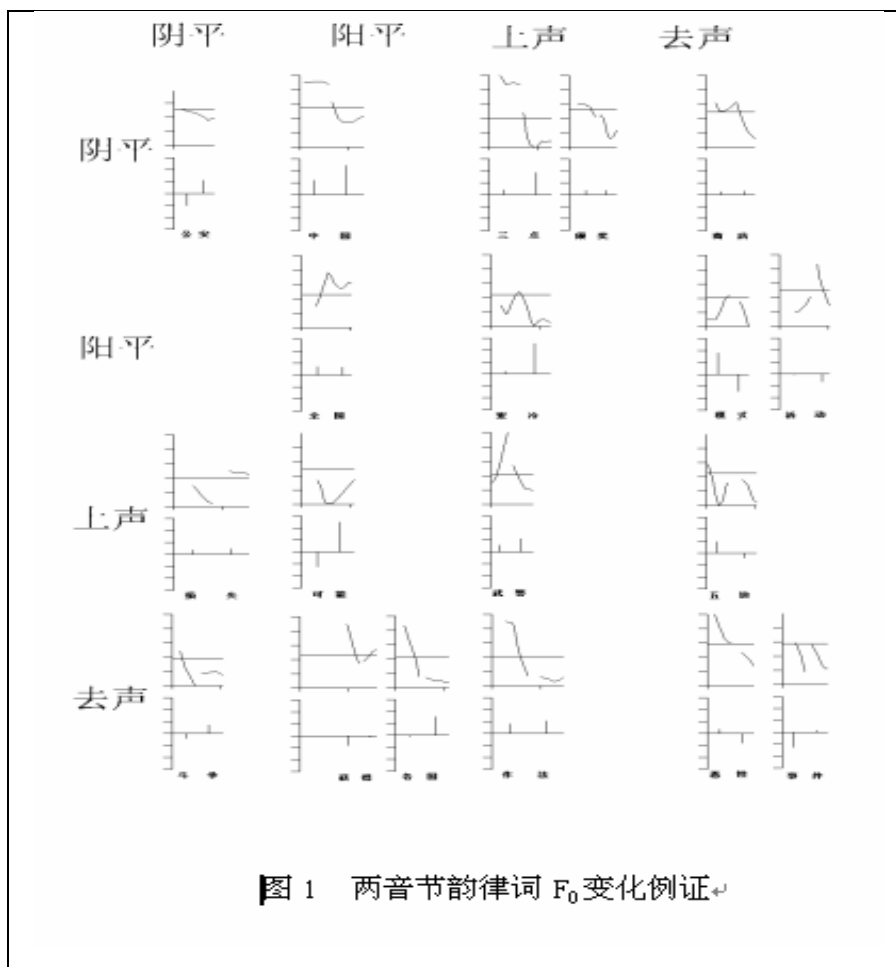
这里必须指出的是,汉语里的复合韵律词是任意的、非绝对的,它决定于说话人怎么念,如“中国人民志愿军”既可以读作“中国|人民|志愿军”,也可以读作“中国人民|志愿军”,还可以读作“中国|人民志愿军”;在“中国人民”或“人民志愿军”中,各音节 $F_0$ 既可以读成是逐步下降的,也可以读成前部分 $F_0$ 位置和 $F_0$ 音域比后面的高和大。对语音片段中各音节 $F_0$ 呈逐步下降的,绝大多数听音人都会觉得它的各个音节是紧密地连在一起念出来,也就是说,这种语音片段的各音节间没有什么间断,而对前部分 $F_0$ 位置和 $F_0$ 音域比后面的高和大的语音片段,就有一小部分听音人认为前后两部分不是紧密连在一起念出来的,也就是说,这种语音片段的前后两部分之间有一点点间断。

图1给出了两音节的简单韵律词 $F_0$ 变化例证,这些例子来自本实验用的59句。两音节组四个声调共有16个声调组合,图1缺少“阳平+阴平”组合。图1有19个小图,小

图上部为用声调五度值表示的 $F_0$ 曲线,下部为规整的时长T。在“阴平+上声”组合中,“三点”和“颁奖”的 $F_0$ 曲线形状大体相同,可是其 $F_0$ 音域和位置都不一样,这是因为它们来自不同句子,“三点”念得重。在“阳平+去声”组合中,“模式”的“模” $F_0$ 位置略高于“式”,而“活动”的“动” $F_0$ 位置大大高于“活”的,这是由于“活动”的“动”念得比“活”重所致。“恶性”和“事件”都是“去声+去声”组合,“恶性”中“性”的 $F_0$ 随着“恶”的下降而下降,可是在“事件”中,“件”的 $F_0$ 起点跟“事”的基本一样高,这大概因为“件”不比“事”念得轻。在“去声+阳平”组合中,“获得”的“得” $F_0$ 在“获”之后变成上升,是由于它念得很轻并跟其后接词“全国”发生了 $F_0$ 协同发音;而“各国”中“国”的 $F_0$ 没有上升,因为“各”比“国”念得重而且其后有一个间断。

图2给出了三音节简单韵律词 $F_0$ 变化的一些例子。“鸣叫声”中的“叫” $F_0$ 位置比其前后的都高,这是因为发音人念这三个音节时把“叫”音节强调了。“兴奋剂”三个去声的 $F_0$ 逐步下降,可是,“奥运会”也是三个去声,第三个音节“会” $F_0$ 位置反而比“运”高,这是“运”比其前后都轻。“检察长”和“老朋友”都是“上声+阳平+上声”组合,“检察长”中“察”重读了,而在“老朋友”中,“老”被重读了。

图3中有几个带附着成分的复杂韵律词 $F_0$ 变化的例子。先看“的”的 $F_0$ 表现。助词“的”附着在其前面成分上一同念了出来,组成复杂韵律词,可是“的”的 $F_0$ 表现却随其前面音节声调的不同而变化:“是唧唧的”中“的” $F_0$ 在其前面阴平音节 $F_0$ 之后才下降,“单位的”的“的” $F_0$ 随其前面“位” $F_0$ 下降而下降,“改革的”中“的” $F_0$ 在“革” $F_0$ 上升之后下降,上声之后的“的” $F_0$ 曲线被抬高了。本研究语料中,只有“苦笑着”和“坐着的”两个“着”的例子。在韵律词中,“了”的 $F_0$ 表现有点象“的”。“也来到”中副词“也”和“不允许”中否定词“不”,这位发音人没有给以强调,它就跟后面的连在一起念,其 $F_0$ 变化跟同一声调的相似。图3还有三个复合韵律词的例子。在“反对北京申办”中,“申办”的“申”相对于“北京”的“京”发生了 $F_0$ 重设,使“申办”跟其前面的成为不同的韵律成分:“北京” $F_0$ 音域比“反对”的小,但是“北京” $F_0$ 位置不比“反对”的低,因而,“反对北京”组成复合韵律词,“申办”是简单韵律词。在“也没有排除联合组阁”中,“联合”的“联”相对于“排除”的“除”发生了 $F_0$ 重设,使得“也没有排除”和“联合组阁”成为两个韵律成分,而这两个韵律成分分别是复合韵律词。“积极支持”是复合韵律词,“南非人民的”是复杂韵律词,而“一贯”是简单韵律词。



### 2.2.2 韵律词的边界特点

韵律词的边界可以从两方面考察，一是考察韵律词的末音节韵母时长是否拉长了，另一个是研究韵律词与韵律词之间的  $F_0$  重设。

#### 2.2.2.1 韵律词的末音节韵母时长

研究语句中韵律词末音节时长是否拉长，只涉及不是紧挨“无声停顿”和“有声停顿”之前的韵律词。因为处于“无声停顿”和“有声停顿”之前而且又紧挨它们的韵律词，或者就是韵律短语，或者是韵律短语的末了一个韵律词。关于韵律短语末音节时长拉长问题，请参见拙作(林茂灿, 2000)。

发音人	韵律词数目	首音节韵母时长 (T) 平均值和标准差 (ms)	末音节韵母时长 (T) 平均值和标准差 (ms)	t 检验
F00	37	$T_{b\text{ave}}=163\text{ms}$ $T_{b\text{sd}}=44\text{ms}$	$T_{e\text{ave}}=164\text{ms}$ $T_{b\text{sd}}=50\text{ms}$	$P>0.1$
M00	37	$T_{b\text{ave}}=149\text{ms}$ $T_{b\text{sd}}=47\text{ms}$	$T_{e\text{ave}}=163\text{ms}$ $T_{b\text{sd}}=48\text{ms}$	$P>0.1$
M01	40	$T_{b\text{ave}}=153\text{ms}$ $T_{b\text{sd}}=47\text{ms}$	$T_{e\text{ave}}=155\text{ms}$ $T_{b\text{sd}}=42\text{ms}$	$P>0.5$

表 1 三位发音人念的不紧挨“无声停顿”和“有声停顿”前韵律词的末音节和首音节韵母 T 平均值及其标准差

表 1 给出了三位发音人念的不紧挨“无声停顿”和“有声停顿”前韵律词的末音节和首音节韵母 T 平均值及其标准差。韵律词包括简单韵律词，复杂韵律词和复合韵律词。单音节如果被强调了，就成为一个韵律词，表一没有计算这类韵律词。从表 1 可以看到，三位发音人的不紧挨“无声停顿”和“有声停顿”前韵律词的末音节韵母 T 平均值都比首音节的大几毫秒或十几毫秒，但其平均值在统计上没有意义，因为表一告诉我们，它们的 t 检验的显著性水平都大于 0.1。例如，发音人 M00 不紧挨“无声停顿”和“有声停顿”前韵律词的末音节韵母时长平均值  $T=163\text{ms}$ ，首音节韵母时长平均值  $T=149\text{ms}$ ，其时长差平均为 14ms，其差值在三位发音人中是最大的，但其平均值缺乏统计意义，因为其平均值偏差更大。因而，不紧挨“无声停顿”和“有声停顿”前韵律词末音节韵母时长不是总比其首音节的长。

#### 2.2.2.2 韵律词与韵律词之间的 $F_0$ 重设

表 2 给出了三位发音人韵律词与韵律词之间的  $F_0$  重设情况。从表 2 可以看到，三位发音人念的句子中，韵律词的首音节  $F_0$  与前一个韵律词的末音节  $F_0$  之间，大多数发生了正  $F_0$  重设，少数发生负  $F_0$  重设。三位发音人中，韵律词之间发生的正  $F_0$  重设值最大的是 86Hz，最小的 30Hz；韵律词之间发生的负  $F_0$  重设值最大的是 43Hz，最小的 23Hz。不过它们的偏差都很大。

韵律词边界特点是，韵律词的末音节韵母时长不是总比其首音节的长，但是韵律词之间发生了程度不同的  $F_0$  重设。韵律词与韵律词之间的  $F_0$  重设是后面韵律词相对于前面韵

律词后一个成分而言的。韵律词在由词汇词组成更大韵律成分中起重要作用。韵律词与韵律词之间的  $F_0$  重设，和韵律词内部各音节  $F_0$  的特有表现，是口语在韵律方面形成的十分重要特点。它们对口语的自然悦耳，起了很大作用。

	$F_0$ 正重设的平均值，标准差和个数	$F_0$ 负重设的平均值，标准差和个数
F00	$F_{0\text{ave}}=86.3\text{Hz}$ $F_{0\text{st}}=59\text{Hz}$ N=45	$F_{0\text{ave}}=43.3\text{Hz}$ $F_{0\text{st}}=8.15\text{Hz}$ N=3
M00	$F_{0\text{ave}}=30.4\text{Hz}$ $F_{0\text{st}}=15.4\text{Hz}$ N=39	$F_{0\text{ave}}=26.3\text{Hz}$ $F_{0\text{st}}=11.8\text{Hz}$ N=9
M01	$F_{0\text{ave}}=35.4\text{Hz}$ $F_{0\text{st}}=20.6\text{Hz}$ N=46	$F_{0\text{ave}}=23.4\text{Hz}$ $F_{0\text{st}}=17.4\text{Hz}$ N=12

表 2 三位发音人韵律词与韵律词之间的  $F_0$  重设情况

## 3. 语句中重音研究

### 3.1 语句重音研究进展

“语句重音分布模式直觉”(杨玉芳,《心理学报》, 28卷 3 期, 1996)一文指出,在汉语中有介于句子重音和词重音的短语重音存在。“汉语语句重音对音高和音长的影响”(许洁萍、初敏、贺琳、吕士楠, 2000)一文以连续的广播语言为研究对象,对汉语中语句重音对韵律特征参数的影响进行了研究。他们在分析了不同重音条件下音长和音高的变化及其相互关系后指出:(1)音高是语句重音的基本表达手段,随着语句重音级别的提高,音高分布曲线向高频方向推移;(2)在连续语流中词被‘重读’和‘轻读’的情况下,音长分布出现双峰,表示它们的音长有的受语句重音影响,有的不受语句重音影响;(3)在‘正常’、‘重读’和‘轻读’三种情况下,音高和音长之间的相互关系是:不相关、负相关和正相关,这说明汉语语句重音音高和音长之间存在互补关系。

“关于普通话韵律短语重音的实验研究”(钟晓波、郑波、杨玉芳, 2002)一文为研究韵律短语重音作了三个逻辑上紧密联系的实验:实验一利用重叠词短语对普通话短语重音进行研究,实验二进行句子重音和短语重音相互作用的研究,实验三利用实验二得到的结果对韵律短语重音作进一步研究。这篇论文的结论是,“汉语普通话中存在韵律短语重音,而且这个短语重音落在短语的语义焦点所在词上;词的重读音节时长的延长是短语重音的一个重要声学表现。”

“汉语语句重音中重读音节音高变化模式研究”(王培、吕士楠、杨玉芳, 2002)一文从重音知觉实验、问答匹配实验和语料库分析三个方面研究重音。“重音知觉实验结果表明,音域平移和高音点提高是重音知觉线索,但是高音点的提高对词重音知觉的作用更明显。重音音节音高变化模式的两项研究表明,重读音节的音高在低音线一高音线上的渐降是汉语语调模式上变化,高音点的提高是重读音节音



高变化的主要声学表现,低音线的变化更多的受到低音线渐降的限制,变化幅度不十分明显,而且不是必须提高。高音线—低音线双线语调模型中,高音线起落的变化,前后音节高音点的对比关系表明句中音节的重读程度。

### 3.2 我们对语句重音的研究

韵律词是我们研究重音的出发点。我们认为语句重音是分层次的,从韵律词重音出发来构建语调短语重音和语句重音。 $F_0$  高低线也是具有层次性的。

#### 3.2.1 语句重音的层次性

简单韵律词中有一两个音节听起来较为尖锐、有力,这一两个音节是韵律词的突显部分。在简单韵律词中,尖锐、有力音节的  $F_0$  位置离这个词的  $F_0$  平均值远和 / 或  $F_0$  音域大。这一两个音节时长如果也长,听起来更突显。复杂韵律词的组成成分在  $F_0$  音域和  $F_0$  位置上不一样,  $F_0$  位置离这个

词  $F_0$  平均值较远和/或  $F_0$  音域大的音节或音节组往往被感觉为尖锐、有力。复合韵律词中  $F_0$  音域大和  $F_0$  位置高的部分,听音人认为比较尖锐、有力。

韵律词中听起来突显的一两个音节,称作词重音。韵律短语重音指它所包含的几个韵律词里凸显度最大的一两个音节。语句重音指它所包含的几个韵律词短语里凸显度最大的一两个音节。因而,语句重音具有层级性。

#### 3.2.2 语句的 $F_0$ 高线和 $F_0$ 低线

Thorsen (1985)研究了丹麦语语句的  $F_0$  高低线,他的  $F_0$  低线指语句中第一个和最后一个重读音节 (accented syllables)  $F_0$  最小值的连线,而  $F_0$  高线指句中第一个和最后一个重读音节后的音节 (post-tonic syllables)  $F_0$  最大值的连线。Ladd (1984) 强调地说,“所拟合  $F_0$  高线或低线的各个点在音系上应是等价的”。本文从韵律词重音出发拟合韵律短语  $F_0$  高低线,以及语句  $F_0$  高低线。

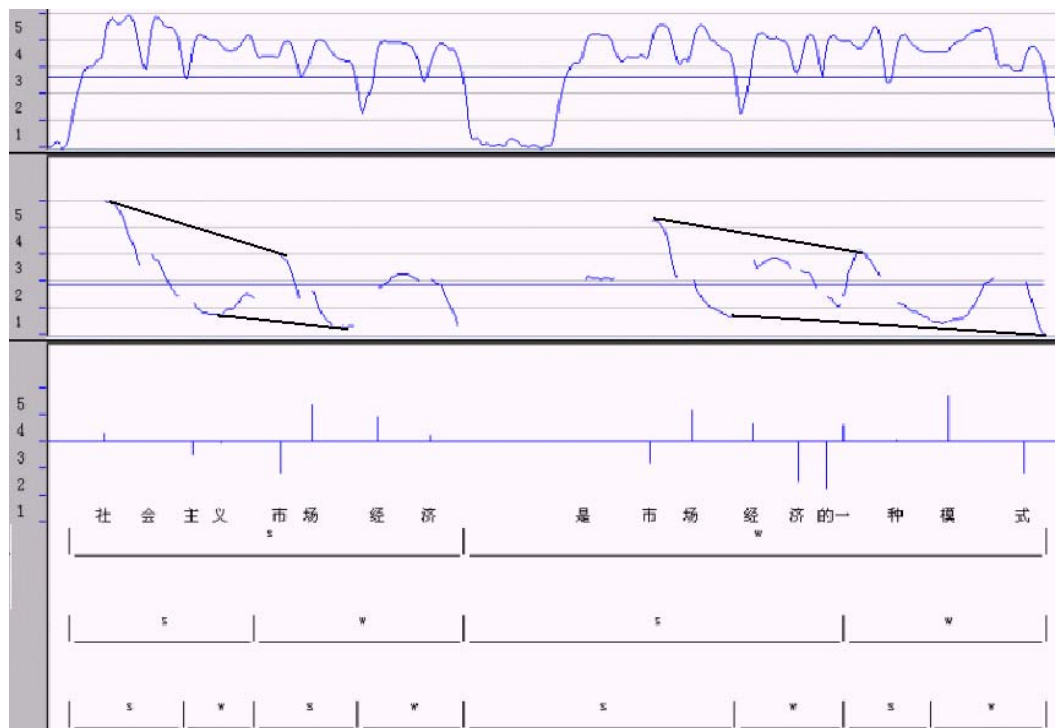


图4 发音人 M01 念的“社会主义市场经济是市场经济的一种模式”各音节的规整基频 ( $F_0$ ), 时长 (T) 和  $F_0$  高低线 (中部), 振幅曲线 (上部) 及这句话的重音分布 (下部)

由于语句重音的层次性,我们认为  $F_0$  高低线也有三个层次,它们是韵律词的  $F_0$  高点和低点 (或  $F_0$  高低线), 韵律短语的  $F_0$  高低线,以及全句的  $F_0$  高低线。韵律词  $F_0$  高点和低点 (或  $F_0$  高低线) 分别指韵律词中重读音节  $F_0$  的最大值 (或上声的  $F_0$  最小值) 和这个词的  $F_0$  最小值 (或这个词的  $F_0$  最大值) 或其  $F_0$  最大值点和最小值点的连线。韵律短语  $F_0$  高低线是分别把所包含韵律词的  $F_0$  最大值点和最小值点连接而成的。以同样的道理,把一个句子中所包含韵律短

语的  $F_0$  最大值点和最小值点分别连接起来的两条直线,就成为这个句子的  $F_0$  高线和  $F_0$  低线。

图4 给出发音人 M01 念的“社会主义市场经济是市场经济的一种模式”各音节的规整基频 ( $F_0$ ), 时长 (T) 和  $F_0$  高低线 (图中部), 振幅曲线 (图上部) 及这句话的重音分布 (图下部)。发音人在“社会主义市场经济”和“是市场经济的一种模式”之间有个停顿,使这句话分成两个韵律大短语。在“社会主义”中,“主义”的  $F_0$  随前面“社会”

的下降而下降，“主义”与“社会”之间没有出现 $F_0$ 重设，“主义”的 $F_0$ 范围和位置也明显地比“社会”的小和低，绝大部分听音人认为这两个词是紧密连着一一起念的，所以，“社会主义”是个复杂韵律词。“市场经济”中“经济” $F_0$ 范围和位置相对于“市场”的小和低，但是“经”的 $F_0$ 相对于“场”（不是相对“市场”）发生了重设，只有约60%听音人认为“市场经济”是紧密连着一一起念出来，因而，“市场经济”是复合韵律词。“社会主义”和“市场经济”所以成为两个韵律词，因为“市场”的“市” $F_0$ 相对于“主义”的发生了 $F_0$ 重设。“是市场经济的”是复合韵律词，因为“是”非重读，其 $F_0$ 位置处于这句话 $F_0$ 平均值附近，时长也不长，它跟“市场”一起念；还因为“经济” $F_0$ 范围和位置比“市场”的小和低，“经” $F_0$ 相对于其前面“场”的发生重设。“一种模式”是复杂韵律词。“是市场经济”和“一种模式”的边界特点是，“一种”的“一”相对于其前面“经济”发生了 $F_0$ 重设。

根据图4下部的重音分布，给出这句话两个韵律短语的 $F_0$ 高低线（用直线分别画在规整的各音节 $F_0$ 曲线上（图中部））。第一个韵律短语是前重型，因为它包含的第一个韵

律词的 $F_0$ 范围和位置比第二个的大和高。第二个韵律短语也是前重型的。全句的 $F_0$ 高线和低线，是把它所包含两个韵律短语的 $F_0$ 最大点和最低点连接而成（图中没画出），它是前重型。

图5给出发音人M00念的“村民也指点他上喇叭里介绍一下自己的经验”各音节的规整基频（ $F_0$ ），时长（T）和 $F_0$ 高低线（图中部），振幅曲线（图上部）及这句话的重音分布（图下部）。“村民”和“也指点他上喇叭里介绍一下自己的经验”为两个韵律大短语，因为“村民”之后除了约有375ms的无声停顿，“村民”的“民”时长还明显拉长，因而引起给人以更大间断的感觉。“也指点他上喇叭里”和“介绍一下自己的经验”为两个韵律小短语，因为它们之间只有约350ms的无声停顿。“村民”是一个简单韵律词，在这句话中它成了韵律大短语。“也”念作“半上”，时长很长，它是一个简单韵律词。“指点”跟“他”连在一起念，“指点他”是一个复杂韵律词。“上喇叭里”也是一个复杂韵律词。“也指点他上喇叭里”这个韵律小短语由三个韵律词组成。韵律小短语“介绍一下自己的经验”由两个复杂韵律词“介绍一下”和“自己的经验”组成。

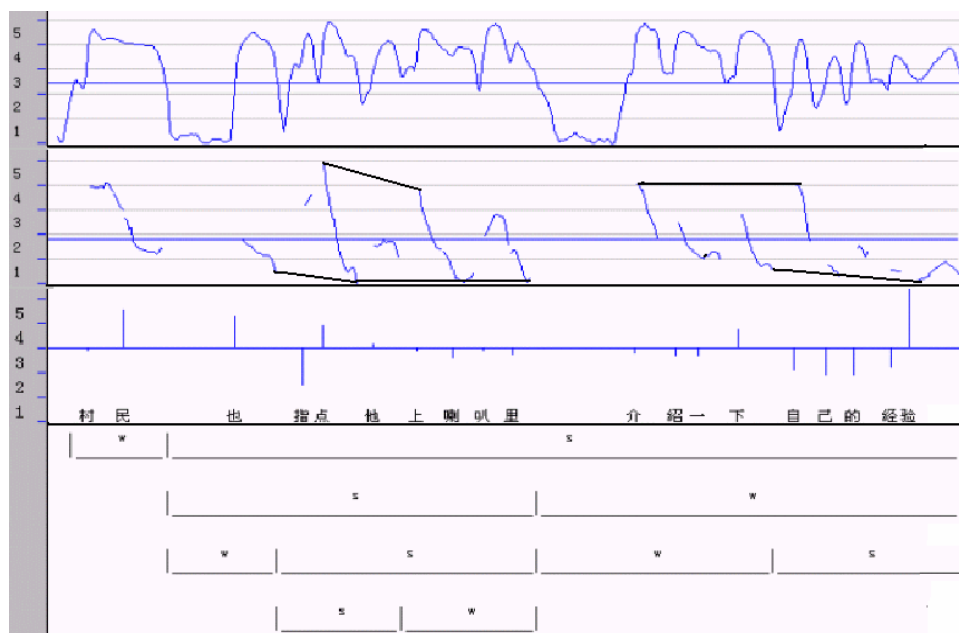


图5 发音人M00念的“村民也指点他上喇叭里介绍一下自己的经验”各音节的规整基频（ $F_0$ ），时长（T）和 $F_0$ 高低线（图中部），振幅曲线（图上部）及这句话的重音分布（图下部）

根据图5下部的重音分布，给出这句话三个韵律短语的 $F_0$ 高低线（用直线分别画在规整的各音节 $F_0$ 曲线上（图中部））。“也指点他上喇叭里”这个短语是中重型，因为“点”在这儿读成 $F_0$ 高降，虽然它本身是上声。“介绍以下自己的经验”是后重型，后一个韵律词比其前面的稍微重一点。全句的 $F_0$ 高低线，是把三个韵律短语的最大点和最低点连接起来（图中没画出），是中重型。

### 3.2.3 韵律短语的 $F_0$ 高低线总体趋势

图6和图7给出三位发音人(m00,m01和f00)韵律短语的 $F_0$ 高低线模式。图6和图7横座标为时长，其不同在于纵坐标，图6纵坐标是以频率为单位，图7纵坐标是以声调五度值为单位。

图6和图7有6个小图，上左小图为（SW）<sub>1</sub>，上右为（SW）<sub>2</sub>，中左为（WS<sub>非上</sub>）<sub>1</sub>，中右为（WS<sub>非上</sub>）<sub>2</sub>，下左为



(WS<sub>上</sub>), 下右为 WSW, 各个小图 F<sub>0</sub> 高低线时长都画成相等的。三位发音人的这些句子, 有的无停顿, 有的有一个停顿, 最多有两个停顿。对一个停顿的语句, 把其前面韵律短语记为 (SW)<sub>1</sub> 和 (WS)<sub>1</sub>, 后面的记为 (SW)<sub>2</sub> 和 (WS)<sub>2</sub>, 对两个停顿的语句, 这儿只考察第一个和最后一个韵律短语的情况。对没有停顿的语句, 把它算作第二韵律短语。(SW) 表示前重 (第一个韵律词比其后面的重), (WS) 表示后重 (第二个韵律词比其前面的重), WSW 表示中重 (中间韵律词念得比其前面和后面的都重)。

本研究用 59 句子最短 6 个音节最长 20 个音节, 三位发音人念各个句子速度不一样, 各个韵律短语所占时长也不一样, 重音在各个句子中位置也不一样。图 6 和图 7 各个小图的韵律短语 F<sub>0</sub> 高线和 F<sub>0</sub> 低线是对同一轻重格式各个韵律短语相应 F<sub>0</sub> 高线和 F<sub>0</sub> 低线的 F<sub>0</sub> 值分别作平均得到的 (对 SW 和 WS 取前后两个点的 F<sub>0</sub> 值分别作平均, 对 WSW 取前中后三个点的 F<sub>0</sub> 值分别作平均)。图 6 和图 7 各个小图的 F<sub>0</sub> 高线和 F<sub>0</sub> 低线取相同长度。

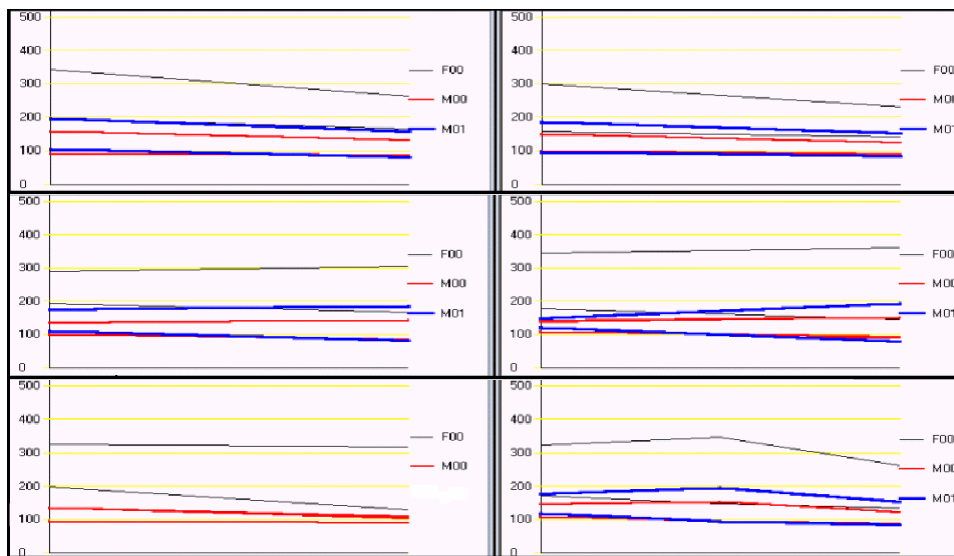


图 6 三位发音人韵律短语的以频率为纵坐标 F<sub>0</sub> 高低线

比较图 6 的 F<sub>0</sub> 的高低线可以看到, F00 的 F<sub>0</sub> 值及其范围 (F<sub>0</sub> 高线和低线之间距离) 都比 M00 和 M01 的大, F00 的 F<sub>0</sub> 低线位置几乎就是 M00 和 M01 的高线位置, F00 是位女同志, 她念这些句子较为强调, 所以音域也宽。从图 7

看到, 以声调五度值为纵坐标的三位发音人 F<sub>0</sub> 高线的起点或终点绝大多数处于声调五度值的“5”, F<sub>0</sub> 低线的起点或终点绝大多数处于五度值的“1”。

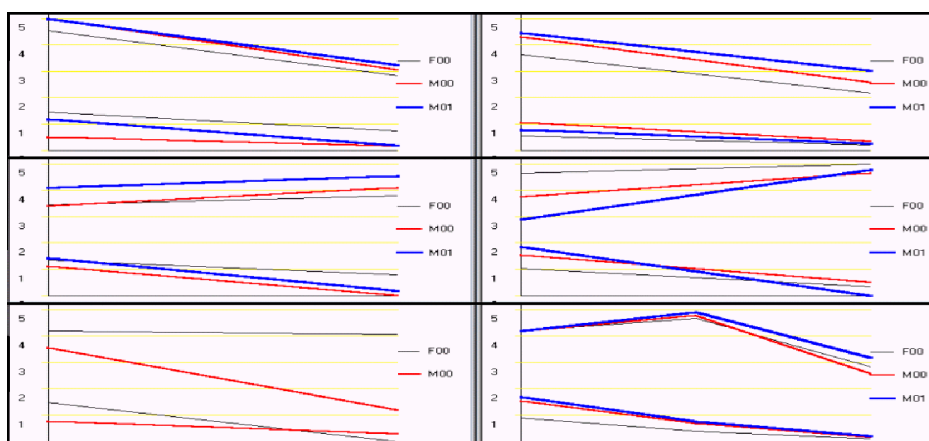


图 7 三位发音人韵律短语以声调五度值为纵坐标 F<sub>0</sub> 高低线

三位发音人韵律短语的前重 (SW) 型占 50% (F00 占 35%; M00 占 48%; M01 占 67%), 后重 (WS) 型占 28% (F00 占 38%; M00 占 32%; M01 占 14%), 中重 (WSW) 型占 19% (F00

占 27%; M00 占 20%; M01 占 19%)。在韵律短语中, SW 型和 WS 型的 F<sub>0</sub> 低线总是下降的, WSW 型 F<sub>0</sub> 低线呈折线下降, 前一段下降快后一段下降慢一些。SW 型 F<sub>0</sub> 高线, 无论

头一个

韵律短语还是末了一个韵律短语都是下降的。WS 型  $F_0$  高线，无论头一个韵律短语还是末了一个韵律短语都是上升的，但 M00 的 (WS<sub>上</sub>) 例外。WSW 型  $F_0$  高线呈升降型， $F_0$  先升后降在重音处最大。

从图 6 和图 7 看到，无论 (SW)<sub>1</sub> 还是 (SW)<sub>2</sub> 在 S 处的  $F_0$  音域和  $F_0$  高线位置都比 W 的大和高。 $F_0$  音域指  $F_0$  高线和低线之间的距离， $F_0$  高线位置指  $F_0$  高线在图 6 和图 7 纵坐标上的数值。图 6 和图 7 中部 (WS<sub>非上</sub>)<sub>1</sub> 和 (WS<sub>非上</sub>)<sub>2</sub> 也是 S 处的  $F_0$  音域和  $F_0$  高线位置都比 W 的大和高。图 6 和图 7 下左小图中，F00 的 (WS<sub>上</sub>) 也是 S 处的  $F_0$  音域和  $F_0$  高线位置都比 W 的大和高，但 M00 的 (WS<sub>上</sub>) 只有一例，它是 S 处的  $F_0$  音域和  $F_0$  高线位置都比 W 的小和低。在 WSW 的 S 处  $F_0$  音域和  $F_0$  高线位置都比其前后 W 的大和高。从绝大部分的韵律短语来说，重音与  $F_0$  音域和  $F_0$  高线位置密切相关，重音处的  $F_0$  音域大而且  $F_0$  高线位置高（上声的  $F_0$  低线位置低）。

总的说来，在韵律短语中， $F_0$  低线总是下降的，而  $F_0$  高线只在重音后下降；重音处的  $F_0$  音域大而且  $F_0$  高线位置高（上声的  $F_0$  低线位置低）。也就是说， $F_0$  位置越高（上声的  $F_0$  位置越低）和/或  $F_0$  音域越大的音节或音节组，给人以更突显、更重的感觉，如果时长也长，重的感觉更明显。但是也有这样的情况，韵律词中一个音节即使其  $F_0$  值大但时长很短，而邻近音节  $F_0$  值虽然小但时长长，后者却给人以较重的感觉，这可能就是许洁萍、初敏、贺琳、吕士楠（2000）文章中说的音高与时长之间成负相关的情况。总的说来，韵律短语或语句中的音节  $F_0$  曲拱（包括时长）越接近单字调的，就给人以越重的感觉。

语句中非重音的  $F_0$  曲线比起单念的要复杂得多，有多种多样表现，从图 4 和图 5 就可以看出来。在  $F_0$  五度制空间，普通话非重读阴平，阳平和去声的  $F_0$  曲线位置有处于‘4’，‘3’，也有‘2’甚至‘1’（度）的，上声  $F_0$  曲线位置都很低，非重读的多处于‘二度’左右。至于非重音节  $F_0$  曲线方向，阴平的可以上升也可以下降，阳平也可以下降，上声有略升的也有略降的，去声可以略升。语句中重读音节发音时用力足，声带足够紧张，使得  $F_0$  曲拱饱满，而非重读音节发音时用力相对不足，声带相对不够紧张，加上受连读变调和协同发音的影响和生理上的限制，使得  $F_0$  曲拱偏离了‘预定’位置和方向。

许毅对普通话声调在一定语境条件下的变化作了系统、深入研究（Xu, 1994, 1997, 1999）。他看到声调  $F_0$  曲拱随其前面声调末端  $F_0$  发生了很大变化，特别当其声母是浊的时候。例如，低调后面的高调  $F_0$  曲线明显成为上升的，同样，高调后面的低调  $F_0$  曲线明显成为下降的。他和他的同事（Xu

and Wang, 2001）指出，在这种条件下出现如此大的  $F_0$  过渡，是由于从一个声调到另一个声调需要花一定时间的缘故。最近他们看到（Xu and Sun, 2002），普通话说话人完成六个半音的音高上升平均要花 142ms。许毅（Xu, 1999）一项研究得到的音节平均时长为 180ms，他们指出（Xu and Sun, 2002）即使说话人用了最快速度从低调转到高调，这种转变所花时间占了音节时长的一大半。他们指出，过渡时间如此的长只能从发音生理作解释。许毅研究了实验句中焦点问题（即将在 *Journal of Chinese Linguistics* 刊出），他说跟焦点有关的有三种不同的音高范围：焦点词（非句末）的音高范围要扩展，焦点之后词的音高范围又低又小，其他位置上的词的音高范围是中性的。

## 4. 汉语语调

陈述句、疑问句、祈使句、感叹句等属于语法范畴的功能句调，语调还包括说话时所表现的情感，如喜欢、愤怒、悲伤、恐惧、爱慕和厌恶等，当然还有语义和语用引起的一般口气问题。由于语调的复杂性，人们研究语调往往从功能语调入手。这儿介绍普通话疑问句和陈述句的声学表现，而疑问句也是先探讨一般疑问句（不带“吗”是非问句）。

### 4.1 汉语语调的研究概况

赵元任先生为汉语语调提出了精辟看法和形象比喻，他从“北平语调的研究”（1929）和“国语语调”（1932）到《汉语口语语法》（1968）都提出中性语调或正常语调问题。他说“中性语调里有两种变化：一种是字与字相连所发生的变化，一种是因字音的轻重而发生的声调上的变化”。赵元任把“音节的声调和句子的语调比作小波浪跨在大波浪上面”，使人们看到声调和语调是两个层面的音系事件。赵元任指出，“语调里最要紧的变化是音程和时间的放大跟缩小”，他说“这种变化最好拿一个机械的比方来解释”，这种比方就是“橡皮带”理论。音程的放大或缩小，就是音高音域的放大或缩小。我们从赵元任的这些论述看到，他把重音跟语句中音节的音域变化紧密联系在一起。

赵元任在“汉语的字调跟语调”（1933）提出“语调跟字调的叠加”问题，并举了一些例子，他在《汉语口语语法》中举了许多“代数和”的例子。我觉得赵元任所说的“代数和”多用于句末。

他在“英语语调（附美语变体）与汉语对应语调初探”（1933）提出了语调层级构造（hierarchical structure），语调由调冠（或句首前，prehead），调头（主句首，main head），调核（核心，nucleus）和调尾（句尾，tail）构成。

吴宗济先生提出变调块及其移动说（1982, 1990, 1998），他对普通话双音节组、三音节组和四音节组变调进行了研究，认为变调组是语调基元。他说“普通话中构成语句的意

群单元，除两字和三字的组合外，以四字组的用途最多。”

“把四字组的韵律变量分析好，就为研究成句的语调、轻重和时长奠定了基础。”他又说，“语句中变调组会整块地向上或向下移动”，“在不同语气下调域会展宽或缩小，调阶会抬高或降低。”

沈炯（1985，1992，1994a，1994b）首先用实验看到音高音域的实际表现，他说“语调是由一连串声调音域组织起来的音高调节形式，声调是在声调音域中滑动的曲拱。”他说“在许多音节组成的语句里，把这些高音点和低音点连接起来，就得到全句的高音线和低音线”。他认为“语势重音的音理是声调音域高音线上移，它使声调音域向上移动。声调音域低音线与语势重音没有明显关系。”他看到“陈述语调是高音线聚落形式和低音线下延形式两种特征的组合，疑问句语调是高音线渐落形式和低音线上敛两种特征的组合，高音线聚落和低音线上敛是普通的祈使句，全句高音线渐落和低音线下延是一种重要的感叹语调。”

沈晓南（Xiaonan Susan Shen，1990）认为汉语北京话语调有三个基本调子(tune)，即 tune-1, tune-2 和 tune-3，她给出了这三个调子的音高(F<sub>0</sub>)曲线。tune-1 的 F<sub>0</sub> 曲线从中调开始，上升高到中高调，然后下降到低音阶；tune-2 从中高调开始，向上升高到高调，然后下降一些，终止于高音阶或高中音阶；tune-3 从中高调开始，上升到高调，然后逐步下降到低音阶。tune-1 用于陈述句，tune-2 用于选择问句和特殊问句，tune-3 用于是非问句。

Nien-Chuang Chang（1958）看到汉语疑问句是句末的音高表现，他说“成都话疑问句末了音节声调受到了‘扰动’，原来高升调的仍然为高升，但终点更高；低降调的变成低平；高降的变成高平；低降升的变成低升。”

吕叔湘先生说(1985)，“是非问一般是对整个命题的疑问，但有时也会集中在某一点，形成一个焦点。这个焦点在说话中可以用对比重音来表示。”

胡明扬先生（1987）认为语调可以指全句的语调，也可以专指句终的语调。句终语调指的是全句末了从最后一个重音算起的这部分语调。关于句终语调，他进一步指出，“北京话语调的音高问题不是‘音高变化’，即‘升’和‘降’的问题，而是字调的起点的高低问题。他认为，‘高’是整个字调定调定得高，‘低’是整个字调定得低。”

劲松（1992）从实验中看到，北京话的语调是出现在句子末一节奏单元中重读音节的音高变化。

曹剑芬通过声学语音学的实验分析，研究汉语声调与语调的关系(2002)，她指出“在语流中，声调和语调同时并存，两者的关系是调阶叠加的‘代数和’，而不是调形叠加的‘代数和’。”

## 4.2 我们对语调的观察

### 4.2.1 研究语料，方法和思路

我们使用自然电话对话语料（即 973 语料），观察各种语调现象。这儿介绍不带“吗”是非问句的研究结果，涉及其中的 65 个自然句。为了得到句末一两个音节 F<sub>0</sub> 变化模式和时长与边界调可能存在的关系，还分析了四位发音人自然念出我们自编的对话材料。对所用句子的基频(F<sub>0</sub>)作了五度值规整，每个音节声韵母时长也作了规整。我们作了听辨试验，观察声学表现与主观听辨之间关系。还作了合成实验，以验证声学分析结果的可靠性。

### 4.2.2 疑问信息与边界调

语调是口语问题，我们研究语调从韵律词（韵律词并不总等于词汇词）为基本单元的韵律结构出发。我们认为一般疑问句除了其韵律结构以外有两个重要特点，一是重音，二是疑问信息，其他疑问句的特点也是如此。我们关于重音的基本看法就是前面所说的，不过，我们正在用研究疑问句的语料进一步研究其重音的声学表现。对一般疑问句来说，其疑问信息由其语段末了韵律词的末了一（两）个音节所携带，有的还由这个起头音节携带，所以把一般疑问句中携带语段疑问信息的音节叫做边界调。

#### 4.2.2.1 韵律大短语与语调短语

下面是来自973电话语料库中一个对话：

服务员：一张票四十块钱手续费，可以吗？

客人：一张票四十块钱手续费？

服务员：对。

在这个对话中，“一张票四十块钱手续费？”是客人对服务员说的“一张票四十块钱手续费”产生疑问而说的。服务员说的“一张票四十块钱手续费”这部分语段，人们听起来是陈述性质的，而客人说的“一张票四十块钱手续费？”是疑问语气。

图8是“一张票四十块钱手续费，可以吗？”这句话的 F<sub>0</sub> 规整（上部）和音节时长规整（中部和下部）。从图8看到，这句话在“手续费”之后有个停顿，加上“手续费”的“费”时长很长，人们听起来在“手续费”与“可以”之间有明显间断，因而这句话分成两个韵律大短语“一张票四十块钱手续费”和“可以吗？”在“一张票四十块钱手续费”中，“钱”和“手”时长相对都长，而且“手续费”的“手”相对其前面音节“钱”有个 F<sub>0</sub> 跃变，所以韵律大短语“一张票四十块钱手续费”至少有两个韵律词“一张票四十块钱”和“手续费”。“可以吗”既是一个韵律大短语也是一个韵律词。对这两个韵律大短语分别划出其 F<sub>0</sub> 高低线，它们都是逐步下降的。这句话的韵律大短语“一张票四十块钱手续费”系陈述口气，而韵律大短语“可以吗？”为疑问口气。

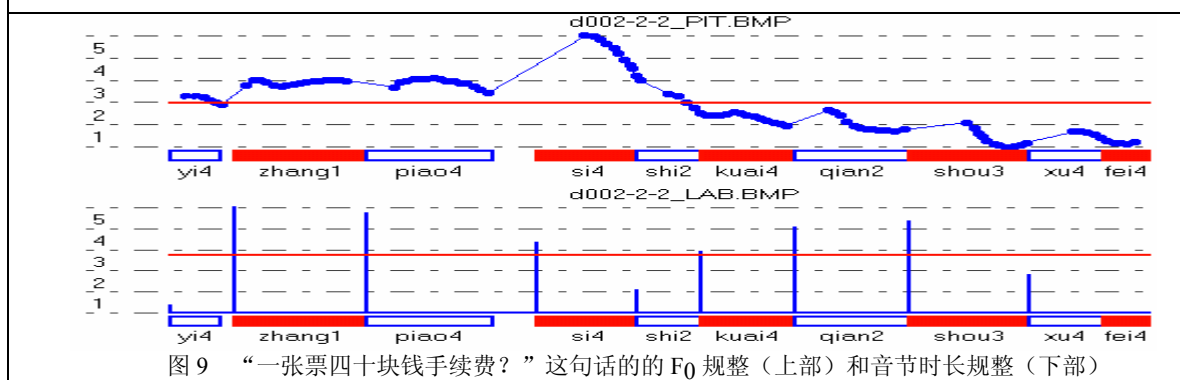
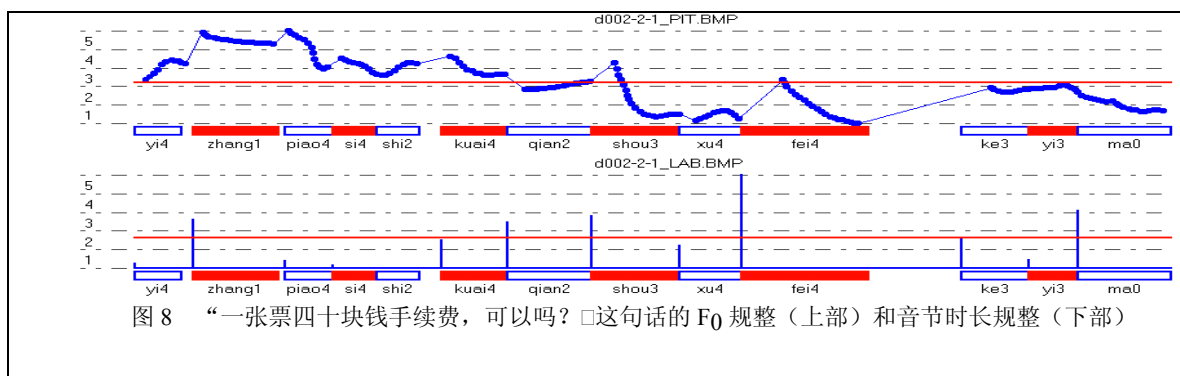


图9是“一张票四十块钱手续费？”这句话的F<sub>0</sub>规整(上部)和音节时长规整(中部和下部)。从图9看到，“四十块钱手续费”的“四”相对于“一张票”的“票”有个明显的F<sub>0</sub>跃变，同时“一张票”的“票”时长长，人们听起来觉得在“一张票”与“四十块钱手续费”之间有个有声间断，因而这个问句的“一张票”和“四十块钱手续费”分别是韵律大短语。

客人以疑问口气说出“一张票四十块钱手续费？”，服务员答“对”，因而这个疑问属于是非问(回声问)。请一些人听这句话，他们觉得“票”的调子没有下降而很平，象“飘”的调子；“四”念得很重，既是疑问焦点，又是疑问信息携带者，因为似乎这句话之所以为疑问也来自这个音节，也就是说疑问信息来自“四”。在这句话中，韵律大短语“一张票”是陈述语气，而韵律大短语“四十块钱手续费”是疑问口气。

从图8和图9两句话的韵律分析看，韵律大短语是语调短语，因为疑问信息存在于这样的韵律短语中。“一张票四十块钱手续费？”的疑问信息由这个“四十块钱手续费”韵律短语的起始音节所携带。这也是疑问信息由语调韵律短语起始音节携带的一个例子。

### 4.2.3 边界调

疑问信息指疑问句靠什么(及在句中什么地方)给人以疑问的感觉。一般疑问句(不带‘吗’疑问句)更需要研究这个问题，因为这种疑问句既不带疑问词也不包含疑问结构等。973 电话对话语料中得到F<sub>0</sub>曲线的句子有上千句，但这样的是非问句只找到三个，服务员问：“一个人住的？”，客人答：“是，一个人住。”，即答话人除了说是否之外，还重复一下对方问的内容；回声问也不到二十句，而且句末多数是去声，没有上声的。于是，我们编了下面这样的句子，其中乙的是非问句属于回声问，由5对发音人念这些对话：

甲：某先生要去某地方。

乙：某先生要去某地方？

甲：是，某先生要去某地方。

句中主语分别为金先生，胡先生，马先生和魏先生等，主语所用的姓氏有不同声调。宾语的地名由16个两音节声调组合组成，如西安，成都，广州，湛江；包头，合肥，沈阳，大同；青岛，南宁，海口，上海；深圳，重庆，武汉，大庆。5对说普通话发音人各担任甲和乙一次说以上16组对话。本文给出对一对发音人的研究结果。

#### 4.2.3.1 听辨试验1:

	主语 (某先生)	谓语 (要去)	宾语 (某城市)
C	21%(11%)		
	8%(6%)		
	96%(6%)		
S	20%(16%)		
	11%(8%)		
	93%(8) (不计“包头”, “青岛”“湛江”和“大庆”4个句子)		

表3 听音人分别判断主语部分, 主谓部分和整个句子(主谓宾)为疑问时的平均百分数(和标准差)

用 Kay 3700 和 Prrat 分析软件中的切音装置分别得到这些“主语+谓语+宾语”疑问句的主语部分,“主语+谓语”部分, 加上原来的疑问句, 每个疑问句共有三个语段。所得到的所有语音片段由八位说北京话的中年人作判断, 他们在是疑问句还是陈述句中做强迫判断。表1给出听音人听两位发音人(C和S)念的语音片段的听辨试验结果, 其他发音人的听辨结果与这两位相似, 另文给出。表1表明, 绝大多数听音人如果只听到包含这种句子的主语以及主语加谓语部分, 不会认为具有疑问语气, 仅仅当听了整句即包含句末宾语部分时, 才认为是疑问句。

#### 4.2.3.2听辨试验2:

在听辨试验1基础上, 我们作了第二个听辨试验, 进一步确定疑问信息跟实验句最后两个音节的哪个音节有关。这个试验的语料当然得用第一个试验两位发音人念的句子, 不过这个试验让听音人听的一个是整个句子, 另一个是把句子最后一个音节切去即不包含最后一个音节的语音片段。表2给出第二个听辨试验的结果。从表2看到: 1) 少量实验句的疑问信息有句首音节携带, 因为100%和86%听音人分别把发音人s念“吴先生要去海口”和“胡先生要去合肥”听为疑问的, 而且还有75%和63%听音人分别把发音人s念“吴先生要去海口”中的“吴先生要去海”和“胡先生要去合肥”中的“胡先生要去合”也听为疑问的; 2) 多数实验句的疑问信息由句末音节携带: 听音人没有把其他语句切去最后一个音节的语音片段判断为疑问句, 也就是说, 大多数听音人只把包含最后一个音节的语音片段判断为疑问句, 这说明这些句子的疑问信息主要由最后一个音节所携带, 跟句子倒数第二音节关系不大。这里要指出的是, 表2中发音人s念的谓语为“包头”, “青岛”, “湛江”和“大庆”的句子, 听音

人判断为疑问口气的得分较低, 因为这位发音人念这4句没有念好, 所以在表1中就没有统计。

#### 4.2.3.3声学分析:

对这两位发音人各念的16个疑问句和相应的肯定句及陈述句作了声学分析, 得到其以频率和以声调五度值为纵坐标的 $F_0$ 曲线。图10是两位发音人说的并被8位听音人中绝大多数判断为疑问句的最后两个音节 $F_0$ 曲线, 同时给出相应肯定句和陈述句里的两个音节 $F_0$ 曲线。从 $F_0$ 曲线看到, 携带疑问信息最后音节不同声调的 $F_0$ 曲拱位置比叙述句的抬高了, 而且其时长也长; 携带疑问信息最后音节 $F_0$ 曲拱象单念时的那样, 而相应叙述句和肯定句的末了音节不论什么声调, 其 $F_0$ 曲线大多是下降的。

我们还初步考察了973电话对话语料65句携带疑问信息音节的 $F_0$ 表现, 首先看到在自然语句中一般疑问句的疑问信息大多说也是由句末音节携带, 只有两句由句首携带; 自然语句中一般疑问句疑问信息的声学表现要比我们自编对话语料中的复杂; 由于所用自然语句的是非问句很少, 没有看到是非问跟回声问在其它方面有什么差别;

总的说来, 陈述句末了音节不论什么声调, 其 $F_0$ 曲线大多是下降的, 而携带疑问信息音节的 $F_0$ 曲拱不仅象单念时的那样, 而且阴平、阳平和去声的终点位置(音阶)要抬高; 如果其时长短于200毫秒, 其 $F_0$ 曲拱位置要比陈述句的抬高约一度半(声调五度值); 疑问信息音节的时长如果较长, 阴平、阳平和去声 $F_0$ 曲拱位置不抬高但其终点抬高, 上声 $F_0$ 曲拱整个抬高; 携带疑问信息音节如果 $F_0$ 曲拱位置抬高, 而且阴平、阳平和去声的终点位置也抬高, 时长也长, 其疑问意味就很浓。至于最后第二音节的 $F_0$ 曲线位置的抬高, 如果它跟疑问信息无关的话, 它就是由 $F_0$ 协同发音引



起。

#### 4.2.3.4 合成验证:

我们用 praat39 以合成方法改变语句最后一个音节的  $F_0$  曲线位置和时长, 把疑问语气变成陈述语气, 和把陈述语气改成疑问语气。所合成语句样品由十位说北京话中年人作是疑问语气还是陈述语气的听辨判断。绝大多数听音人对所合成语句样品的陈述语气和疑问语气都能作出正确反映。这个合成语句的听辨结果表明, 语句最后一个音节是能携带疑问信息的, 携带疑问信息音节  $F_0$  曲线像单念时的那样, 而相应叙述句末了音节相不论什么声调, 其  $F_0$  曲线大多是下降的; 疑问语气的  $F_0$  曲拱位置(音阶)比陈述语气的高, 而且其时长也长; 最后第二音节的  $F_0$  曲线没有加以改变, 说

明它的  $F_0$  曲线抬高跟携带疑问信息关系不大。

#### 4.2.3.5 关于句调的结论

我们用自然和朗读语句研究了一般是非问句的边界调和重音两方面问题。本实验结果表明, 疑问信息主要由语调短语末了韵律词中最后一个非轻声音节所携带, 也可以由句首音节携带。所以, 我们把携带疑问信息的音节称作边界调。携带疑问信息音节的  $F_0$  曲线位置(音阶)比陈述的约抬高一度半(声调五度值), 而且其时长也长; 携带疑问信息音节  $F_0$  曲拱不仅象单念时的那样而且阴平、阳平和去声的终点抬高, 而相应叙述句的末了音节象不论什么声调, 其  $F_0$  曲线大多是下降的。至于最后第二音节的  $F_0$  曲线, 疑问的也比肯定和陈述的高, 这是由于协同发音所引起。

		西安		包头		青岛		深圳		成都		合肥		南宁		重庆		广州		沈阳		海口		武汉		湛江		大同		上海		大庆	
		c	s	c	s	c	s	c	s	c	s	c	s	c	s	c	s	c	s	c	s	c	s	c	s	c	s	c	s	c	s	c	s
听音人	当	2	3	3	2	3	4	3	2	2	6	2	2	2	3	2	3	1	8	4	7	2	4	1	0	0	4	5	5				
	调	0	8	3	1	3	2	3	9	5	3	1	9	1	8	1	3	6	9	5	1	9					6						
	长	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	%	
	四																																
听音人	当	9	8	1	1	1	8	7	1	8	8	8	9	9	1	1	7	1	9	1	1	1	1	8	9	8	1	1	1				
	调	0	3	0	0	0	6	0	0	6	6	9	5	1	0	0	9	0	1	0	0	0	0	0	5	0	0	0	0	0			
	长	%	%	0	0	0	%	%	0	%	%	%	%	%	0	0	%	0	%	0	0	0	0	%	%	0	0	0	0	0	0		
	四			%	%	%			%						%	%	%	%	%	%	%	%	%	%			%	%	%	%	%		

表 4 听音人分别对表 1 整个句子和这个句子不含倒数第一音节的语音片段判断为疑问的百分数

本研究看到, 语句重音是以韵律词为基本节奏单元按层次分布的, 韵律词是一个基频 ( $F_0$ ) 变化组。  $F_0$  高低线也是有层次的。在各个韵律层次中,  $F_0$  位置越高(上声的  $F_0$  位置越低)和/或  $F_0$  音域越大的音节或音节组, 给人以更加突显、更重的感觉。语调短语末了韵律词有时出现这样的情况: 一个音节  $F_0$  值大但时长很短, 而邻近音节  $F_0$  值虽小但时长较长,  $F_0$  值虽小但时长较长的音节却给人以较重的感觉。总的说来, 音节  $F_0$  曲拱(包括时长)越接近单字调的, 就给人以越重的感觉(正常重音)。

## 5. 句调对字调的非线性作用

一般是非问句有边界调和重音两个特征, 我们从这两个特征来讨论句调对字调的作用。从是非问边界调的声学表现看, 句调有个特征加到单字调上, 因为携带疑问信息音节  $F_0$  曲拱不仅像单念时的那样, 而且有的阴平、阳平和去声终点抬高, 有的  $F_0$  曲线位置(音阶)抬高, 而且其时长可能长, 而相应叙述句末了音节如果不强调, 不论什么声调的  $F_0$  曲线大多是下降的。

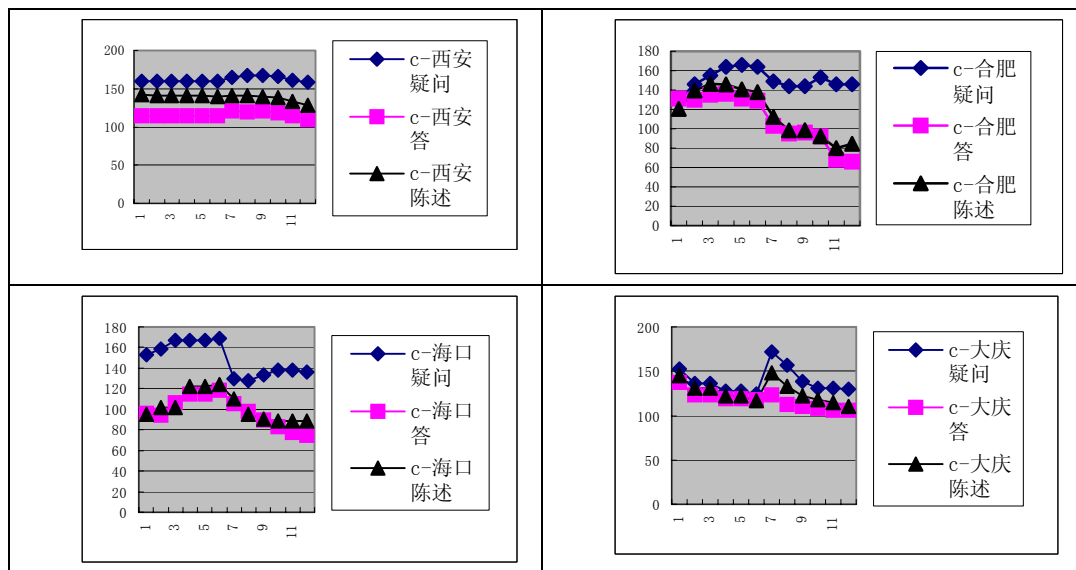


图 10 疑问句和叙述句最后两个音节 F<sub>0</sub> 曲线

再从语调短语中重音的声学表现看，音节念得越重（正常重音），音节 F<sub>0</sub> 曲拱（包括时长）就越接近单字调，而在 F<sub>0</sub> 五度值空间，普通话非重读阴平，阳平和去声的 F<sub>0</sub> 曲线位置有处于‘4’，‘3’，也有‘2’甚至‘1’的，上声 F<sub>0</sub> 曲线位置都很低，非重读的多处于‘2’左右。至于非重音节 F<sub>0</sub> 曲线方向，阴平的可以上升也可以下降，阳平也可以下降，上声有略升的也有略降的，去声可以略升。语句中重读音节发音时用力恰当，声带紧张度大，使得 F<sub>0</sub> 曲拱饱满，而非重读音节发音时用力相对小，声带相对不够紧张，加上受连读变调和协同发音的影响和发音生理上的限制，使得 F<sub>0</sub> 曲拱偏离了‘预定’（本调）位置和方向。语调短语中各个音节 F<sub>0</sub> 有不同表现，主要受重音的作用，还受连读变调和协同发音的影响和发音生理上的限制。

我们认为，在句调和字调的关系方面，主要是句调对

### 参考文献

曹剑芬 (2002), “汉语声调与语调的关系”, 《中国语文》2002 第三期。  
 初敏和吕士楠, 1996, 一种高清晰度高自然度的汉语文语转换系统, 《声学学报》, 21 卷 4 期。  
 胡明扬(1987), “关于北京话的语调问题”, 《北京话初探》, 商务印书馆。  
 劲松(1992), “北京话的的语气和语调”, 《中国语文》1992 年第二期。  
 李爱军, 2002, “普通话自然口语的韵律标注研究”, 《中国语文》创刊 50 周年学术讨论会论文 (江西大学)。  
 林茂灿、颜景助, 1980, 《普通话轻声的声学性质》, 《方言》1980 年 3 期。

声调的作用，这是因为音节念得越重（正常重音），音节 F<sub>0</sub> 曲拱（包括时长）就越接近单字调，而非重读音节在 F<sub>0</sub> 位置上和 / 或 F<sub>0</sub> 音域上主要因重音才有种种变化。我们还认为句调对字调的这种作用是非线性的，这不仅因为重音是层次性的，而且边界调与 F<sub>0</sub> 曲拱及时长的关系也不是一对一的。

## 6. 下步工作

以上是我们对普通话句调的初步看法，写出来跟听取大家意见，跟大家讨论。对 973 电话对话语料和自编语料还在继续分析和观察，研究自然语句中重音的声学表现，边界调的声学模式，以及是非问句与回声句有什么差异等，期望这些研究能更深入，能得到扎实的定量结果。

林茂灿 (1999), “语句 F<sub>0</sub> 曲线的构建和其作用---兼谈语句 F<sub>0</sub> 和时长规整”, 《现代语音学论文集》金城出版社。  
 林茂灿(2000), “普通话语句中间断和语句韵律短语”, 《当代语言学》, 2000 年第四期。  
 林茂灿(2001), “普通话孤立句的韵律结构和 F<sub>0</sub> 下倾”, 《2001 年中国语音学会论文集》。  
 林茂灿、颜景助 (2001), “普通话语句的韵律结构和重音层次性--汉语语调特征之一”, 《纪念李荣先生八十华诞文集》, 商务印书馆。  
 林茂灿 (2002), “普通话孤立句的韵律结构和它的基频(F<sub>0</sub>) 高低线”, 拟在《当代语言学》发表。  
 吕淑湘 (1985), “疑问·否定·肯定”, 《中国语文》1985 年第四期。  
 邵敬敏 (1996), 《现代汉语疑问句研究》, 华东师范大学出

- 版社。
- 沈炯 (1985), 《北京语音实验录》1985。
- 沈炯(1992), “汉语语调模型刍议”, 《语文研究》1992 年第 4 期;
- 沈炯 (1994a), 汉语语势重音的音理 (简要报告), 《语文研究》1994 年第 3 期。
- 沈炯(1994b), “汉语语调构造和语调类型”, 《方言》1994 年第 3 期。
- 吴宗济(1982), “普通话语句中的声调变化”, 《中国语文》, 1982 年第 6 期。
- 王培、吕士楠、杨玉芳, “汉语语句重音中重读音节音高变化模式研究”, 《声学学报》, Vol. 27, No. 3, 2002。
- 吴宗济(1990), “汉语普通话语调的基本调型”, 《王力先生纪念文集》, 商务印书馆。
- 吴宗济(1998), “普通话四字组韵律变量的处理规则”, 《语音研究报告》, 中国社会科学院语言研究所语音研究室刊。
- 许洁萍、初敏、贺琳、吕士楠, “汉语语句重音对音高和音长的影响”, (《声学学报》, Vol.25, No.4, 2000。
- 杨玉芳, 1997, 句法边界的韵律征兆, 《声学学报》, 22 卷 4 期, 414-422。
- 杨玉芳, “语句重音分布模式直觉”, 《心理学报》, 28 卷 3 期, 1996。
- 赵元任 (1929), “北平语调的研究”, 《最后五分钟》附录, 中华书局, 1929; 《赵元任语言学论文集》(吴宗济、赵新那编)。
- 赵元任 (1932), “国语语调”, 赵元任先生 1932 年 2 月 8 日的演讲词, 《赵元任语言学论文集》(吴宗济、赵新那编)。
- 赵元任(1933), “汉语的字调跟语调”, 《赵元任语言学论文集》(吴宗济、赵新那编)。
- 赵元任 (1933), “英语语调 (附美语变体) 与汉语对应语调初探”, 《中央院史语所集刊》第 4 本第 3 分册; 《赵元任语言学论文集》(吴宗济、赵新那编)。
- 赵元任(1968), 《口语语法》, 吕叔湘译, 商务印书馆, 1997。
- 赵元任, 1980, 《语言问题》, 71-91 页, 台湾: 台湾商务印书馆。
- 钟晓波、郑波、杨玉芳, “关于普通话韵律短语重音的实验研究”, 《声学学报》, Vol.27, No.2, 2002。
- Beckman, Mary E., And Gayle M. Ayers(1994), *Guidelines for ToBI Labelling* (Version, 2.0.)
- Boogaart, J., and Silverman, (1992), Evaluation the overall comprehensibility of speech synthesizers. *Proc. of the 1992 Inter. Conference on Spoken Language Processing*, Canada: University of Alberta, 1207-1210.
- Chang, Nien-Chuang (1958), Tones and intonation in the Chengtu dialect (Szechuan, China), *Phonetica*, Vol. 2, Pp. 59-84.
- Hockett, C. F. (1947), *Peiping Phonology*, Journal of the American Oriental Society, 67, 253-267.
- Ladd, D. Robert, (1996), *Intonation Phonology*, Cambridge: Cambridge University Press.
- Ladd, D. Robert (1984), “Declination: A review and some hypotheses”, *Phonology Yearbook*, Vol. 1, pp. 53-74.
- Li, Aijun, (1998), Durational characteristics of the prosodic phrase in Standard Chinese. *The Proc. of the Conference on Phonetics of the Languages in China*, ed., by Eric Zee and Lin Maocan, Hong Kong: City University of Hong Kong, 65-68.
- Lehiste, I & Ann Arbor Mich (1960), An acoustic- phonetic study of internal open juncture. *Supplement to Phonetica* 5, pp. 1-55.
- Lihiste, L. And W. S.-Y. Wang(1977), Perception of sentence and paragraph boundaries with and without semantic information. *Phonologica* 1977, ed., by Dressler and Pfeiffer ( Institute fur Sprachwissenschaft der Universitar Innsbruck, Innsbruck), 277-283.
- Pijper, Jan Roelof and Angelien A. Sanderman (1994), On the perceptual strength of prosodic boundaries and its relation to suprasegmental, *Journal of the Acoustical Society of America*, 96, 4, 2037-2047.
- Price, P. J., M. Ostendorf, S. Shattuck-Hufnagel & C. Fong, (1991), The use of prosody in syntactic disambiguation, *Journal of the Acoustical Society of America*, 90, 6, 2956-2969.
- Sanderman, A. A. And R. Collier (1996), Prosodic rules for the implementation of phrase boundaries in synthetic speech, *Journal of the Acoustical Society of America*, 100, 5, 3390-3397.
- Shen, Xiaonan Susan (1990), *The Prosody of Mandarin Chinese*, University of California Press.
- Shi Chinlin (1986), The prosodic structure of tone sandhi in Mandarin Chinese. Unpublished Ph. D. dissertation. University of California at San Diego.
- Thorsen, Nina Gronnum (1985), Intonation and text in Standard Danish, *J. Acoust. Soc. Am.*, Vol. 77, Num. 3, pp. 1205-1216.
- Trager, G. L. and B. Bloch (1941), The syllable phonemes of

- English. *Language*, vol. 17, pp. 223-246.
- Xu, Y. (1994), "Production and perception of coarticulated tones", *J. Acoust. Soc. Am.*, 95, 2240-2253.
- Xu, Y. (1997), Contextual tonal variation in Mandarin, *J. Phonetics*, 25, 61-83.
- Xu Y. (1998), Consistency of tonal-syllable alignment across different syllable structures and speaker rates, *Phonetica*, 55, 179-203.
- Xu, Y. (1999), Effect of tone and focus on the formation and alignment  $F_0$  contours, *J. Phonetics*, 27, 55-105.
- Xu, Y. (2001), Fundamental frequency peak delay in mandarin, *Phonetica*, 58, 26-52.
- Xu, Y. and Sun, X. (2000), How fast can we really change pitch? Maximum speed of pitch change revisited, *Proceeding of the 6<sup>th</sup> International Conference on Spoken Language Processing*, Beijing, pp. 666-669.
- Xu, Y. and wang, Q. E. (2001), Pitch targets and their realization: Evidences from Mandarin Chinese, *Speech Communication*, 33, 319-337.
- Xu, Y. and Sun, X. J. (2002), Maximum speed of pitch change and how it may related to speech, *J. Acoust. Soc. Am.*, 1399-1413.
- Xu, Yi (2002), Source of tonal variation in Mandarin Chinese, *Journal of Chinese Linguistics*.
- Wu Jiongji (1996), A new method of intonation analysis for Standard Chinese: frequency transposition processing of phrasal contours, edited by Gunner Fant.

## PROSODIC STRUCTURE AND INTONATION IN STANDARD CHINESE

*Lin Maocan*

Institute of Linguistics, CASS

linmaocan@163.com

### ABSTRACT

It was found in our two works that there are two features in Chinese intonation: stress and boundary tone. The acoustical manifestations of these two features were examined. The following results were drawn from prosodic structure of read speech and spontaneous speech in Standard Chinese. And that the acoustical manifestations of these two features was test by speech synthesis is reliable.

In the first work, 59 sentences were used to seek the boundary signals of prosodic word and prosodic phrase, to study the acoustic characteristics of stress, and to construct  $F_0$  top line and  $F_0$  bottom line. These sentences were taken from a data-base of isolated declarative sentences that were used for Chinese recognition.

The results of the first experiment were as follows. 1) Prosodic word is a  $F_0$  variation group, one or two syllables in it are more prominence than others, and  $F_0$  reset always occurs between prosodic words. 2) The boundary of prosodic phrase is demarcated not only by break with silent pause, but also by break with filled pause. Major prosodic phrase and minor prosodic phrase are differentiated. 3)  $F_0$  top line and  $F_0$  bottom

line are constructed based on stress-principle. The lines of  $F_0$  top and  $F_0$  bottom is hierarchical. In prosodic phrase, there are three kinds of  $F_0$  top line and  $F_0$  bottom line, they are:  $F_0$  top line descends, ascends and ascend first and then descend, but  $F_0$  bottom line is always declined. They correspond to WS (the last prosodic word is more stronger than proceeding one(s)), SW (the first prosodic word is stronger than the following one(s)), WSW (the mid prosodic word is stronger than preceding and following one (s)). 4) It is known from  $F_0$  5-point scale,  $F_0$  scale,  $F_0$  range (distance between  $F_0$  top line and  $F_0$  bottom line) and duration are related to stress. In general, more higher  $F_0$  scale and/or more larger  $F_0$  range in some syllable or syllable group of a utterance, more stress (normal stress) is perceived; Namely,  $F_0$  curve in some syllable of a utterance more approximates to  $F_0$  contour in citation form, more stress (normal stress) of the syllable is perceived. Stress in utterance is hierarchical too.

It was clarified that stress is one element of Chinese intonation. A wide variations of  $F_0$  scale,  $F_0$  range and duration in the syllable related of a utterance are conditioned by stress, and some of them are lead to by tone sandhi and tonal

coarticulation and articulatory constraints.

Yes-no question without final particle “吗” was studied in the second work. In this kind of interrogative sentence, no any interrogative word or interrogative structure is appeared. 28 read dialogues and 63 telephone dialogues were used to study the boundary tone. In 28 dialogues, the sentence pattern is ‘a certain mister will go to a certain city’. Two men who are native Beijing Mandarin speakers read these dialogues. And 63 telephone dialogues take from a data-bank.

The results of the second experiment are as follows: 1) the interrogative information in yes-no question without final particle is carried by the last one (or two) syllables in the final prosodic word, with some exceptions when it is carried by the first one syllables, Therefore, question information in yes-no question is known as boundary tone. 2) Major prosodic phrase is intonation phrase, because major prosodic phrase bears boundary tone; 3) As for the acoustical manifestations of boundary tone, they are those that  $F_0$  curve in boundary tone very likes  $F_0$  contour of the syllable in citation form, but  $F_0$  curve of the last syllable in declarative sentence is falling whatever it is any tone when it is not emphasized; The  $F_0$  scale of boundary tone is raised about 1.5 degree ( $F_0$  5-point scale), when the duration of the syllable is shorter than about 200ms, or The  $F_0$  scale of the syllable does not raised, but the ending part of the  $F_0$  curve for tone-1, tone-2 and tone-4 is raised,

when the duration of it is longer; when  $F_0$  scale of the boundary tone is raised and the duration of the syllable is longer, interrogative mood is very strong. 4) As for the raising of the  $F_0$  scale of second syllable in an utterance, it is caused by  $F_0$  tonal coarticulation, when the second syllable is not related to the question information.

Interrogative mood was become into declarative mood, or inverse by changing the scale of whole  $F_0$  curve, or the scale of the ending of  $F_0$  curve and duration in the last syllable of two utterances using the synthesis system in Praat39. And  $F_0$  scale and/or  $F_0$  range of some syllable in two utterances were changed to get different stress. These synthetic samples were judged by ten listeners who speak Standard Chinese. This judgment test by listeners indicated that the acoustical manifestation of boundary tone and stress were reveal was reliable.

As for the interrelation between tone and intonation, it is advocated it is that intonation act on tone, and this action is non-linear, because it is hierarchical that stress act on  $F_0$  scale,  $F_0$  range and duration, and the relation between boundary and the scale of  $F_0$  curve and duration is not one-to-one.

The results above are tentative, because the results about the representation of boundary tone and stress were drawn from two different speech data. Stress in telephone dialogues is being studied;  $F_0$  pattern of boundary tone will be given.